BELGRADE METROPOLITAN UNIVERSITY

eLearning

16th International Conference on e-Learning 2025

16th International Conference on e-Learning 2025

www.elearning.metropolitan.ac.rs

PROCEEDINGS

The Sixteenth International Conference on e-Learning

eLearning

16th International Conference on e-Learning 2025

Editors
Dragan Domazet
Nemanja Zdravković

Belgrade Metropolitan University

Belgrade, 25-26 September 2025

www.metropolitan.ac.rs

Publisher

Belgrade Metropolitan University
Tadeuša Košćuška 63, Belgrade, Serbia
http://www.metropolitan.ac.rs

For Publisher

Prof. Dr. Dragan Domazet, Belgrade Metropolitan University, Serbia

Editors

Prof. Dr. Dragan Domazet, Belgrade Metropolitan University, Serbia
Assoc. Prof. Dr. Nemanja Zdravković, Belgrade Metropolitan University, Serbia

Chair of Organizing Committee

Assoc. Prof. Dr. Nemanja Zdravković, Belgrade Metropolitan University, Serbia

Organizing Committee

Assoc. Prof. Dr. Nemanja Zdravković, Belgrade Metropolitan University, Serbia
Asst. Prof. Dr. Aleksandra Stevanović, Belgrade Metropolitan University, Serbia
Olga Pavlović, Belgrade Metropolitan University, Serbia
Ana Ročkomanović, Belgrade Metropolitan University, Serbia

Design

Petar Cvetković

Printing

Control P, Niš

Circulation

20

DOI

https://doi.org/10.46793/eLearning2025
ISBN 978-86-89755-37-4
eISBN 978-86-89755-38-1

Belgrade, 2025.

This international conference is co-financed under the Contract on Co-financing from Budgetary Funds for the Organisation of a Scientific Conference in 2025, under reference number: 451-03-3635/2025-03)

CONTENT

ORGANIZERS

PARTNERS:

CO-SUPPORTED

COORDINATORS OF THE INTERNATIONAL PROGRAMME COMMITTEE:

Nemanja Zdravković, Belgrade Metropolitan University, Serbia
Miguel Ángel Conde, Universidad de Salamanca, Spain

MEMBERS:

Mohammed Saqr, University of Eastern Finland
Sonsoles López-Pernas, University of Eastern Finland
Miguel Ángel Conde, Universidad de Salamanca, Spain
Dragan Domazet, Belgrade Metropolitan University, Serbia
Emilija Kisić, Belgrade Metropolitan University, Serbia
Aleksandra Stevanović, Belgrade Metropolitan University, Serbia
Miljan Milošević, University of Kragujevac, Institute for Information Technologies
Igor Saveljić, University of Kragujevac, Institute for Information Technology
Đorđe Kadijevich, Institute for Educational Research, Belgrade, Serbia
Krassen Stefanov, Sofia University, Bulgaria
Alok Mishra, Molde University College-Specialized University in Logistics, Norway
Nikola Vitković, Faculty of Mechanical Engineering, University of Niš, Serbia
Nemanja Zdraković, Belgrade Metropolitan University, Serbia
Mirjana Ivanović, University of Novi Sad, Serbia
Laura Fedeli, University Macerata, Italy
Sofoklis A. Sotiriou, Head of Research and Development Ellinogermaniki Agogi, Greece
Eliza Stefanova, St. Clement of Ohrid University of Sofia, Bulgaria
Michaele Moro, The University of Padova, Italy
Miloš Kostić, Belgrade Metropolitan University, Serbia
Mirjana Ivanović, University of Novi Sad, Serbia
Behnaz Norouzi, University of Oulu, Finland
Nikola Kadoić, Faculty of Organization and informatics, University of Zagreb, Croatia
Dijana Oreški, Faculty of Organization and informatics, University of Zagreb, Croatia
Nikola Vitković, Faculty of Mechanical Engineering, University of Niš, Serbia
Ilijazi Venezija, Belgrade Metropolitan University, Serbia
Jovana Jović, Belgrade Metropolitan University, Serbia

REVIEW COMMITTEE:

Nemanja Zdravković, Belgrade Metropolitan University, Serbia
Aleksandra Stevanović, Belgrade Metropolitan University, Serbia
Jovana Jović, Belgrade Metropolitan University, Serbia
Danijela Ljubojević, Institute for Educational Research, Serbia
Emilija Kisić, Belgrade Metropolitan University, Serbia
Sonsoles López-Pernas, University of Eastern Finland Miguel
Miloš Kostić, Belgrade Metropolitan University, Serbia
Mirjana Ivanović, University of Novi Sad, Serbia
Behnaz Norouzi, University of Oulu, Finland
Nikola Kadoić, Faculty of Organization and informatics, University of Zagreb, Croatia
Dijana Oreški, Faculty of Organization and informatics, University of Zagreb, Croatia
Nikola Vitković, Faculty of Mechanical Engineering, University of Niš, Serbia
Ilijazi Venezija, Belgrade Metropolitan University, Serbia

Competences-driven and GenAI-supported hybrid personalized learning

Dragan Domazet[1 [0000−0001−8095−5146]]

Belgrade Metropolitan University
Tadeuša Košćuška 63, 11000 Belgrade, Serbia
dragan.domazet@metropolitan.ac.rs

Abstract. The paper presents a conceptual model of the system for personalized learning of students, in which lessons and their topics are divided into mandatory and op-tional, and are searched according to the competencies they provide to the student. The area of competences is divided by the depth of the competences (the depth of knowledge and skills they provide) and by the breadth (the number of teaching topics of each teaching unit). During implementation, the area of competences can change both in depth and breadth, depending on the needs of students and em-ployers. Each student has his own specific repository of links to learning objects, located in the university's repository, which is used for learning, as well as for verification and evaluation of what has been learned.

A hybrid procedure for the preparation of teaching materials is applied. The teacher prepares the application of the GenAI tool in accordance with the univer-sity's standards and checks the generated results. Two approaches to the applica-tion of generative artificial intelligence are used for the preparation of teaching materials: 1) application of prompt engineering for GenAI tools; 2) application of automated preparation of prompts for GenAI tools. All forms of verification of achieved competences (tests, assignments, exams, etc.) are personalized because each student can have different optional topics in his competence portfolio of each course.

The goal is to significantly increase the efficiency of preparing teaching materi-als and verifying what has been learned, which is necessary in the case of per-sonalized learning because it requires a larger number of optional topics and dif-ferent depths of their study, and more work for their implementation Therefore, the application of GenAI tools is a necessity in the case of personalization of learning and verification of what has been learned.

Keywords: Personalized learning · AI in education · Hybrid preparation of teaching materials

1 Introduction

New technologies are developed faster and more often. Due to competition, companies have to offer new, innovative products or services. Their lifespan is getting

Research Paper - Invited Paper
DOI: https://doi.org/10.46793/eLearning2025.008D
Part of ISBN: 978-86-89755-37-4

shorter due to the growth of new technologies and new market demands. The challenge for higher education is how to offer curricula that offer a wide range of competencies to students to meet the needs of companies and the desires of students. Academic curricula quickly become outdated, offering outdated competencies and becoming unattractive to students and employers. Students and employers are increasingly interested in short and fast courses that address their new needs. One of the most relevant answers to this great challenge for higher education institutions, in our opinion, is the personalization of learning, because it can provide the necessary flexibility of curricula. But personalizing education requires:

- smaller teaching groups or even individual consultations and guidance of students,
- development of many different digital learning materials for students.

These two factors increase the cost of education. Higher education institutions are now under pressure to reduce their costs, but with the personalization of education, this task is now a major challenge for them. Using GenAI is the solution to this challenge. This can significantly reduce the cost of developing digital learning materials and the time required for this development, that is, increase the efficiency of this development. This paper presents a conceptual solution for the personalization of education supported by GenAI, which was prepared and accepted by the Belgrade Metropolitan University (BMU).

The paper builds on the researches listed at the end of the paper, but the paper does not aim to provide their analysis here. The paper presents a proposal for a new concept for the development and application of personalized learning for students, guided by their needs for acquiring new competencies, supported by the application of generative artificial intelligence for the preparation of all kinds of teaching materials, which students can also prepare, but with the active participation of their teachers.

2 Required prerequisites for successful personalized education at BMU

2.1 Development of the curriculum and the syllabus

Before we determine the curriculum, we analyze global and local development trends in the economy, industries and technologies trying to predict what competencies employers will need in the next 5 to 10 years. These analyzes have a major impact on our curriculum design decisions. We want to prepare our students for the future needs of employers and avoid offering curricula and courses based on the interests of our teachers, who may have some special interests. As the dynamics of change are increasingly significant, it is important to anticipate the future needs of employers, and not use only current needs and requirements. Education must be carefully planned to be ready for the future.

Another important actor in the design of our curricula are the recommendations of various professional groups and their professional organizations. All

BMU study pro-grams must be based on the recommendations of internationally recognized professional organizations in a particular field of science (such as IEEE and ACM in computer science). They periodically publish their Body of Knowledge (BOK) recommendations. We select the most appropriate body of knowledge (BOK) for each of our study programs. Each BOM defines its own objectives, learning outcomes and competencies. We map them into the curriculum we want to design based on these recommendations. Taking into account the recommended BOK units of study, we list all the courses that should implement the selected BOK (Fig. 1).

These BOK-to-curriculum mappings make our curricula internationally recognizable. In cases where BOK is not specific enough in a field of science, we expand it to more specific fields and sub-fields, bearing in mind the specific requirements of our study programs. We also take into account similar curricula of the most recognized and prestigious universities in the world, when determining the compulsory and optional courses of the curriculum. After that, we map the already defined objectives, learning outcomes and competencies of each curriculum to the objectives, learning outcomes and competencies of each of its courses. All their extensions and implementations provide a deeper and wider insight into the foundations determined by the BOK and the curriculum as its specialization and extension.

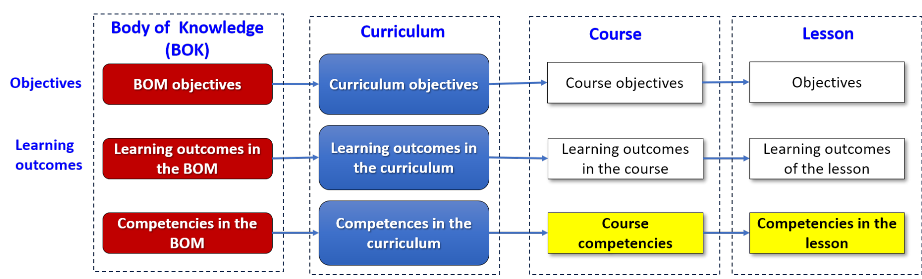


Fig. 1. Mapping the objectives, learning outcomes and competencies from the original BOK to our relevant curriculum, its subjects and their lessons.

This process of mapping objectives, learning outcomes and competencies ensures their top-down consistency and two-way linkage. No course or its lesson should be irrelevant or independent of the curriculum and its BOK. This ensures consistency of all courses and its lessons with the BOK and its curriculum. For example, if we add a competency to a lesson that is not supported by any course-level competencies, then a warning message can be sent to the course or curriculum creator, asking for their approval and verification of this expansion of the competency list. Therefore, our curricula, courses and their lessons are guided by the BOK competencies, but can be adapted and expanded according to the future needs of our target labor markets. This is an important feature that will be explained later.

For each course, we analyze the available and relevant international textbooks and for each course we select several of them, which are used by the most prominent universities and publishers. The curriculum designer collaborates with the course designers to determine appropriate lessons based on selected textbooks for each course. In this way, we map the goals, learning outcomes and competencies of each course into the goals, learning outcomes and competencies of each of its lessons (Fig. 2).

If the objectives, learning outcomes and competencies are well planned in the curriculum, this curriculum development process reduces the risk that the curriculum or some of its courses or lessons will become obsolete when the first generation of students graduates. But that is not enough. We must be ready for surprises and for new demands of the labor market that we did not expect and plan for. The solution to this challenge is fast, adopting new competencies into our curricula on the fly, and the "recipe" for this challenge is to personalize our curricula.

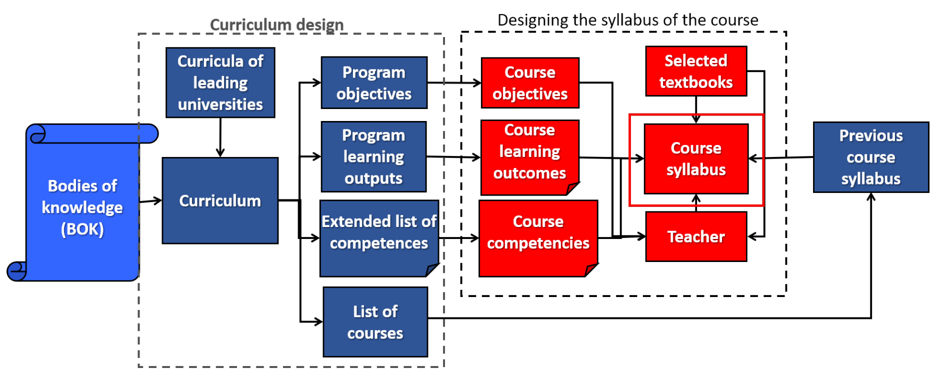


Fig. 2. Mapping the corpus competencies into the curriculum and into its subject programs.

3 Application of learning objects

Before we explain the concept of personalization of education, we need to explain the role of the so-called learning objects that we use in all our subjects and curricula. Learning object (LO) is the smallest unit of knowledge that has 4 components (Fig 3):

1. A concept that reveals new knowledge
2. An example of how new knowledge can be used
3. A task that the student must solve by applying new knowledge
4. A test used to assess understanding of a concept

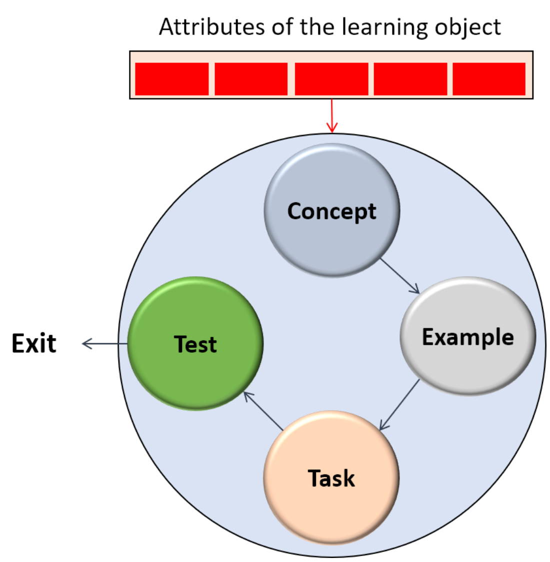


Fig. 3. Structure of the learning object.

The learning object has sections that contain a description of these LO components: Concept, Example, Task and Test. These sections are actually carriers of knowledge and skills of learning objects. Sections can have a textual, graphic and visual type of knowledge presentation. They may also have external references to different software services or products. Learning objects can have their own hierarchy if needed when modeling learning activities and learning materials. This allows a learning object to have one or more sub-objects. Only the lowest in this hierarchy is the smallest possible learning object that provides detailed learning, or the smallest piece of knowledge. A learning material (LM) consists of a set of learning objects that have different structures. In the context of higher education, a common lesson structure is shown in Figure 4.

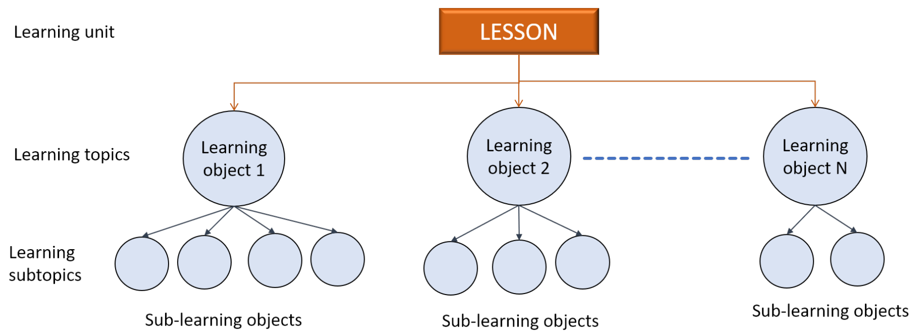


Fig. 4. Learning objects and their sub-objects of a lesson.

Learning objects in Figure 4 can be topics of learning units, i.e. lessons. Sub-objects can be sub-topics.

Learning objectives, learning outcomes and learning competencies can be deter-mined for each learning unit, i.e. lesson, but also for each learning topic, i.e. learning object. They can be specified as learning object attributes and can also be used to search for learning objects with a specific value of one of these three characteristics.

As a competence, it has two components: knowledge and skill of applying this knowledge (here we ignore its third component of competence – the attitudes of the person who has the competence, because it is not of interest for our analysis). For our further analysis, we will use competencies as the main classifier of learning objects representing learning units (lessons), learning topics or learning sub-topics.

3.1 Use of three levels of competence of learning objects

For the learning object that represents the learning topic, we can determine three levels of competence:

1. Level 1 (L1): This compulsory competence that every student must acquire
2. Level 2 (L2): This is an optional competence (above level 1) that the learner can choose to acquire.
3. Level 3 (L3): Optional competence above level 2, which reflects the student's ability to think critically, i.e. for rational, analytical and creative thinking and the ability to synthesize new knowledge based on given facts.

Figure 5 shows the structure of learning objects of a lesson (learning unit). The number and content of learning object sections determine the depth of competence, and the number of topics (sub-learning objects) reflects the breadth of competence. The green color represents the mandatory (obligatory) competencies of the lesson, and the yellow color represents the optional competencies.

As shown in Figure 6, the lesson learning process for Level 1 may use a mostly fixed, sequential sequence of activities that perform the same actions on learning

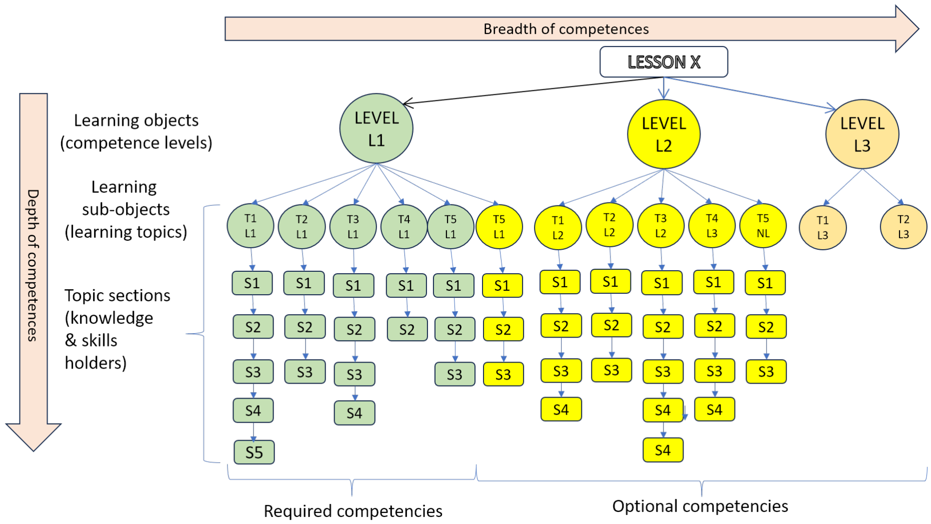


Fig. 5. Lesson presented with three learning objects representing three levels of competence with their sub-objects representing lesson topics with their sections.

objects, i.e. on its mandatory topics. This is the most commonly used learning path for the acquisition of mandatory (or required) competences (level 1).

For optional competences (level 2) learning paths can be more complex, as they can also use branching learning processes. For level 3, there are several activities that the student per-forms in order to prepare an essay to present in a debate with other students If the task is to develop a learning unit and its learning objects, he/she can apply problem-oriented learning at L1 and L2 competence levels, and in some cases at L3.

Figure 7 shows these learning processes as sub-processes applied at the L1, L2 and L3 competence levels.

According to the accreditation rules, in Serbia, the course must provide students with up to 100 points, of which 30 to 70 points can be earned by pre-exam obligations (assignments). BMU decided to apply the 70:30 model, i.e. that the pre-exam assignments provide up to 70 points, and the exam, up to 30 points. In this way, the concept of active learning is supported, which implies greater permanent and active learning of students during the semester, which enables constant contact with students and direct insight into their progress in learning during the semester.

With this in mind, Figure 8 a shows a competency model for their courses. Students must have at least 35 points to be eligible for the exam. They must win 28 points from mandatory (key) competencies and at least 7 points for optional points. This gives students a better chance of getting at least 35 points, as the model offers more electives than core competencies. Teachers are obliged to provide teaching materials for basic competencies for each basic topic of at

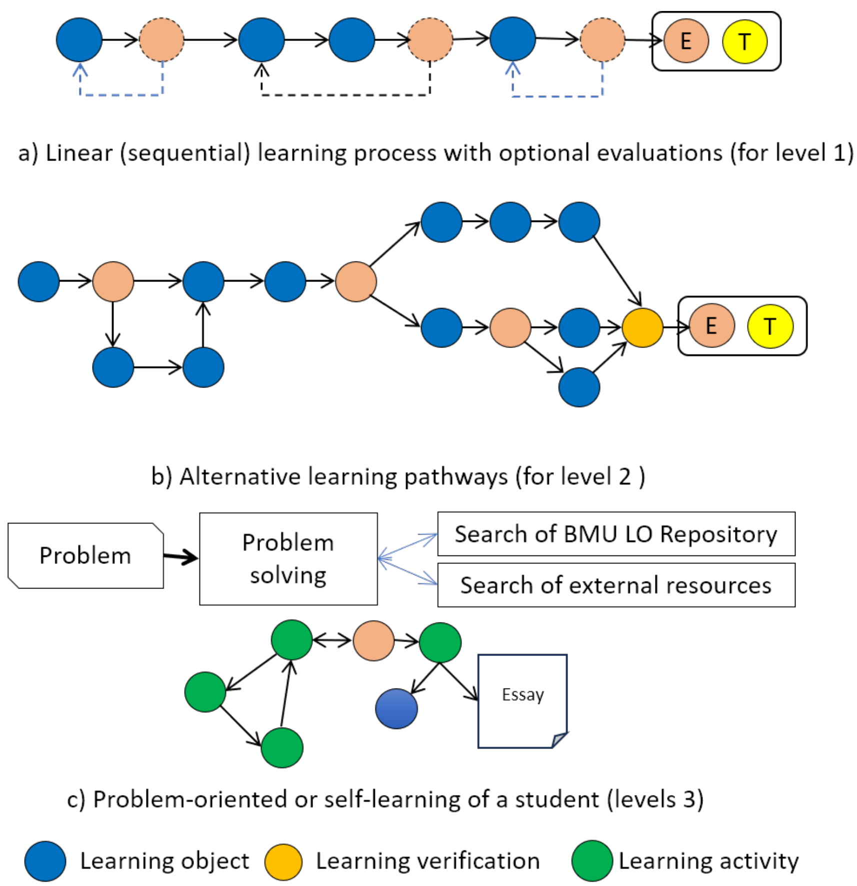


Fig. 6. Learning process in the case of three levels of competence.

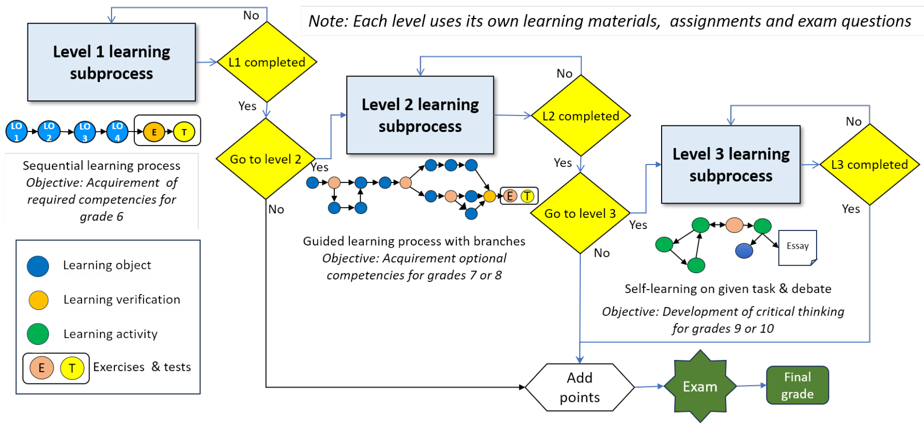


Fig. 7. Typical learning sub-processes applied at L1, L2 and L3 competence levels.

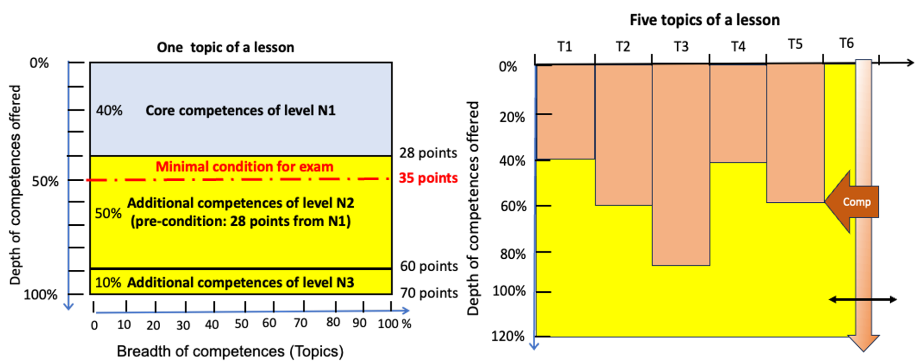
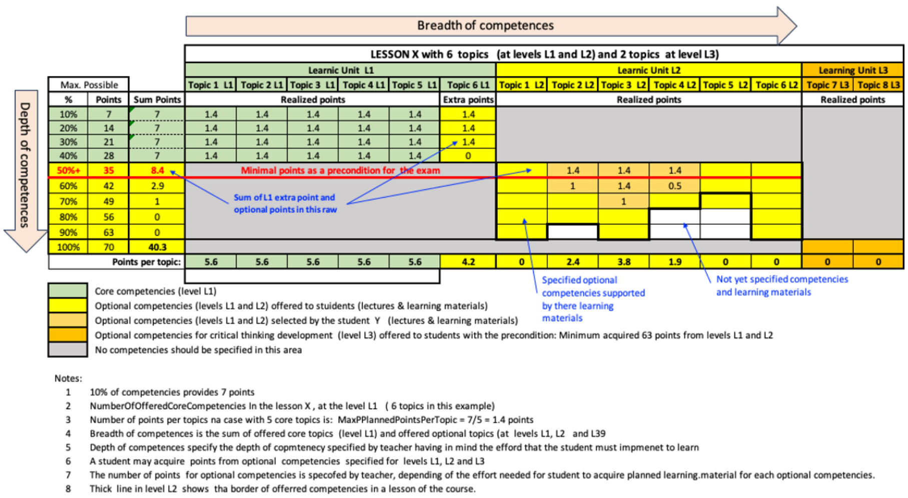


Fig. 8. Setting the competence level of the planned topics of one lesson with 5 mandatory and one optional topic.

least 28 points (40% of all competencies). Each topic is planned for more than 28 points, because students must be able to obtain a maximum of 70 points. In the example in Figure 8.b, topic 6 is optional and can be left for the subsequent specification of its competencies, after the students' request.

A more detailed Excel model of competencies is shown in Figure 9. In this example, 10% of the depth of competencies corresponds to 7 points. The last 10% of competencies (for level 3) correspond to the last 7 points of competencies. The analyzed lesson of the course has 5 basic topics, and the 6th is an optional topic. This model is used by the teacher to plan the organization of his/her lectures and exercises.

Students also use an Excel model to select optional competencies. An Excel model is created for each lesson of the SE201 Introduction to Software Engineering course and is created for each student. The student can use it to select,



Max. Possible			LESSON X with 6 topics (at levels L1 and L2) and 2 topics at level L3)														
			Learnic Unit L1						Learnic Unit L2						Learning Unit L3	
			Topic 1 L1	Topic 2 L1	Topic 3 L1	Topic 4 L1	Topic 5 L1	Topic 6 L1	Topic 1 L2	Topic 2 L2	Topic 3 L2	Topic 4 L2	Topic 5 L2	Topic 6 L2	Topic 7 L3	Topic 8 L3
%	Points	Sum Points	Realized points					Extra points	Realized points						Realized points	
10%	7	7	1.4	1.4	1.4	1.4	1.4	1.4								
20%	14	7	1.4	1.4	1.4	1.4	1.4	1.4								
30%	21	7	1.4	1.4	1.4	1.4	1.4	1.4								
40%	28	7	1.4	1.4	1.4	1.4	1.4	0								
50%+	35	8.4	Minimal points as a precondition for the exam							1.4	1.4	1.4				
60%	42	2.9								1	1.4	0.5				
70%	49	1									1					
80%	56	0														
90%	63	0														
100%	70	40.3														
Points per topic:			5.6	5.6	5.6	5.6	5.6	4.2	0	2.4	3.8	1.9	0	0	0	0

Fig. 9. Excel model of topics and their competencies with levels L1, L2 and L3 of one lesson of the course SE201 Introduction to Software Engineering.

plan and manage their competencies during the semester. The same possibility is provided by the inter-face of the iMET system, which manages work and data related to students and all other processes and services at the university.

Figure 10 shows the acquired competencies of three students. In this example, the course has up to 8 topics, 6 are compulsory and 2 are optional and can be offered at all three competency levels. Each of the three students has different acquired elective competencies.

3.2 Organization of learning objects

Figure 11 shows the organization of learning elements in the case of lesson 1 of the course SE201 Introduction to Software Engineering planned for the first week of the semester (which lasts 15 weeks). The lesson has 5 compulsory learning topics (planned for 90 minutes of teaching), one of which is intended for exercises lasting 90 minutes. The L1 and L2 competence levels have the same structure. In this case, the only difference is that the L1 level offers only compulsory competences and for which learning material is provided, and the L2 level offers only additional optional teaching materials of the same compulsory competences. Level 3 obliges the student (if he wants to gain the points that level N3 offers) to choose two specific topics planned for the development of critical thinking. Within each lesson topic, one or more suggested essay topics and appropriate references for their preparation are offered. For each topic, the student must prepare an essay on a given assignment, which is offered each week. The student chooses two of the several essay topics offered during the semester and presents them to other students at a specially organized debate. The teacher must provide a list of

Max. Possible		Student 1	Learnic Unit L1						Learnic Unit L2						Learning Unit L3	
			Topic 1 L1	Topic 2 L1	Topic 3 L1	Topic 4 L1	Topic 5 L1	Topic 6 L1	Topic 1 L2	Topic 2 L2	Topic 3 L2	Topic 4 L2	Topic 5 L2	Topic 6 L2	Topic 7 L3	Topic 8 L3
%	Points	Sum Points	Realized points					Extra points	Realized points						Realized points	
10%	7	7	1.4	1.4	1.4	1.4	1.4	1.4								
20%	14	7	1.4	1.4	1.4	1.4	1.4	1.4								
30%	21	7	1.4	1.4	1.4	1.4	1.4	1.4								
40%	28	7	1.4	1.4	1.4	1.4	1.4	0								
50%+	35	8.4	Minimal points as a precondition for the exam							1.4	1.4	1.4				
60%	42	2.9								1	1.4	0.5				
70%	49	1									1					
80%	56	0														
90%	63	0														
100%	70	40.3														
Points per topic:			5.6	5.6	5.6	5.6	5.6	4.2	0	2.4	3.8	1.9	0	0	0	0

Max. Possible		Student 2	Learnic Unit L1						Learnic Unit L2						Learning Unit L3	
			Topic 1 L1	Topic 2 L1	Topic 3 L1	Topic 4 L1	Topic 5 L1	Topic 6 L1	Topic 1 L2	Topic 2 L2	Topic 3 L2	Topic 4 L2	Topic 5 L2	Topic 6 L2	Topic 7 L3	Topic 8 L3
%	Points	Sum Points	Realized points					Extra points	Realized points						Realized points	
10%	7	7	1.4	1.4	1.4	1.4	1.4	1.4								
20%	14	7	1.4	1.4	1.4	1.4	1.4	1.4								
30%	21	7	1.4	1.4	1.4	1.4	1.4	1.2								
40%	28	8.4	1.4	1.4	1.4	1.4	1.4	1.4								
50%+	35	12.3	Minimal points as a precondition for the exam						0.9	1.4	0.8	1.4	1	1.4		
60%	42	3.3								0.8		1.3		1.2		
70%	49	3.2								0.9		1.3		1		
80%	56	2.5										1.2		1.3		
90%	63	2.5										1.2		1.3		
100%	70	53.2														
Points per topic:			5.6	5.6	5.6	5.6	5.6	5.4	0.9	3.1	0.8	6.4	1	6.2	0	0

Max. Possible		Student 3	Learnic Unit L1						Learnic Unit L2						Learning Unit L3	
			Topic 1 L1	Topic 2 L1	Topic 3 L1	Topic 4 L1	Topic 5 L1	Topic 6 L1	Topic 1 L2	Topic 2 L2	Topic 3 L2	Topic 4 L2	Topic 5 L2	Topic 6 L2	Topic 7 L3	Topic 8 L3
%	Points	Sum Points	Realized points					Extra points	Realized points						Realized points	
10%	7	7	1.4	1.4	1.4	1.4	1.4	1.4								
20%	14	7	1.4	1.4	1.4	1.4	1.4									
30%	21	7	1.4	1.4	1.4	1.4	1.4									
40%	28	7	1.4	1.4	1.4	1.4	1.4									
50%+	35	7	Minimal points as a precondition for the exam						1.4	1.4	1.4	1.4				
60%	42	4.1							1.2	1	1.4	0.5				
70%	49	3.1							1	1.1	1					
80%	56	2.2							1		1.2					
90%	63	0.5							0.5							
100%	70	44.9														
Points per topic:			5.6	5.6	5.6	5.6	5.6	1.4	5.1	3.5	5	1.9	0	0	0	0

Fig. 10. Acquired basic and optional competencies in the case of 3 students and their results.

references for each proposed essay topic, but the student may suggest his own topic and may use additional references.

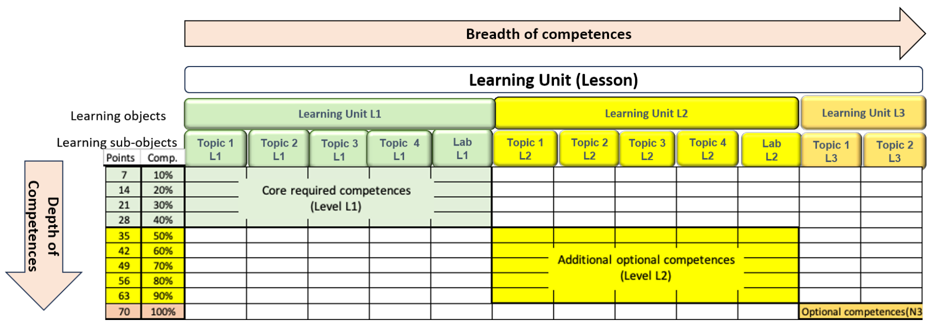
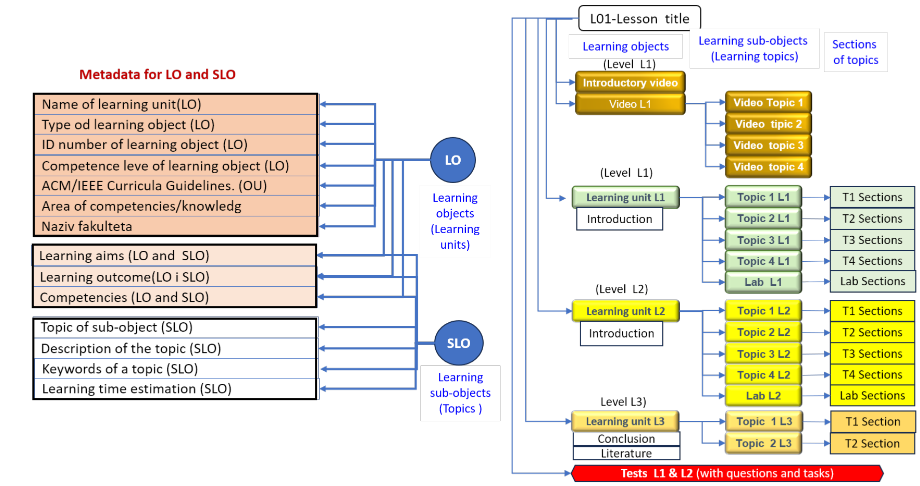


Fig. 11. An example of the structure of learning objects used in lesson 1 of the course SE201 Introduction to Software Engineering.

Figure 12.a shows the learning object (LO) and sub-learning object (SLO) metada-ta currently used to search for lessons in the Learning Object Repository (LOR). Figure 12.b shows the structure of learning objects (levels of competence L1, L2 and L3) and their sub-objects. At the L1 competency level, the learning material also offers videos for each L1 topic. This structure of LOs and SLOs provides students with all the learning material needed for their learning. There is some overlap between L1 and L2 sections (usually L1 sections are shorter), but since the learning material is in digital form (web and PDF forms), it is feasible for students to use it. Tests are given for each lesson and for L1 and L2 levels separately.

As already mentioned, the tests provide one question (out of at least 5 possible) for each topic. Therefore, the teacher must prepare at least 5 alternative questions for each lesson topic and competency level. The test can be multiple choice or essay type. These tests are used to assess the learned knowledge acquired in each lesson. The course SE201 has 13 weeks of study and uses 13 tests to assess the knowledge of its students. In addition to these tests with randomly asked questions, the course also uses tests for randomly selected short tasks from the database of tasks related to one topic, which the student must solve. In this way, the teacher assesses the level of ac-quired skills of the student, i.e. his/her ability to apply learned knowledge. These tests are given after teaching a group of related lessons. In the case of SE201, 3 tests with tasks are used. Similar to question tests, task tests provide at least 5 alternative tasks, for each group of lessons in the course. Students do not have access to test questions and tasks. They are obtained through a random selection of testing software and are only done during a limited testing time under controlled conditions. The questions and tasks for both types of tests are different for L1 and L2 competency levels.



a) Metadata used for learning objects and sub-learning objects b) Basic structure of learning objects at 3 competency levels

Fig. 12. Meta data of LO and SLO and basic structure of competences of 3 levels for the example given in Fig. 11.

Figures 13 and 14 show the headings of sections and sub-sections used by sub-objects, at the L1 level and the L2 level, separately. Section titles are given in bold font, and sub-section titles in normal font.

4 Repository of learning objects

Our goal is to track the mapping of learning outcomes (LOUT), learning competencies (COMP) and their knowledge and skill components (KNOW and SKILL), for each created topic at all competency levels for each lesson (learning unit) and each course at the university (Fig. 15). Obviously, this is not an easy task, but it is an important requirement for implementing personalized learning, because we need to control the goals, learning outcomes, and competencies of each new learning object we create when developing new optional lesson topics. Figure 15 shows an example of the created metadata (TC) for the topic T3 of the L2 competency level containing pointers to the LOUT, COMP, KNOW and SKILL components of the associated learning ob-ect (LO) stored in the university's learning object repository (LOR).

Figure 16.a shows the developed compulsory and optional topics of the teaching material of the course SE201-L04. The course has five basic topics and teaching material (LM) must be developed for them. Learning material for other elective top-ics (T5, T7 and T8) can be developed later, for the next academic year, if a student is interested in a new elective topic. Figure 16.b shows the compulsory and optional topics chosen by student 1. He/she also decided to develop learning material (LM) for the optional topic T6 for all three competency levels.

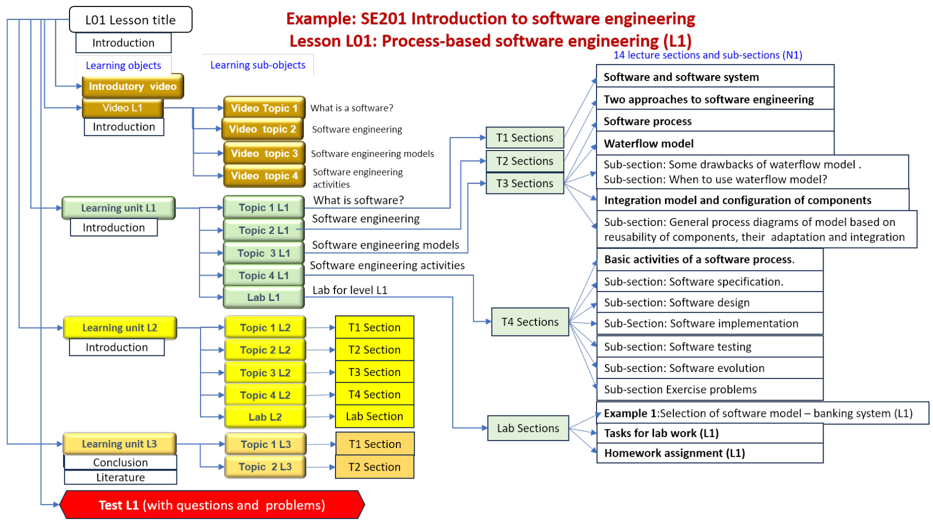


Fig. 13. Example of knowledge content of learning objects and their L1 sections of lesson SE201-L01.

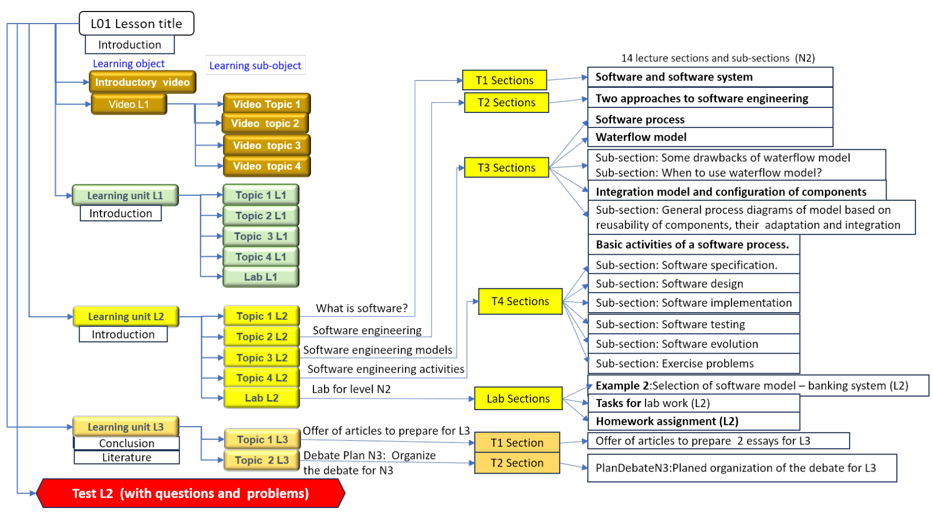


Fig. 14. Example of knowledge content of learning objects and their L2 sections of lesson SE201-L01.

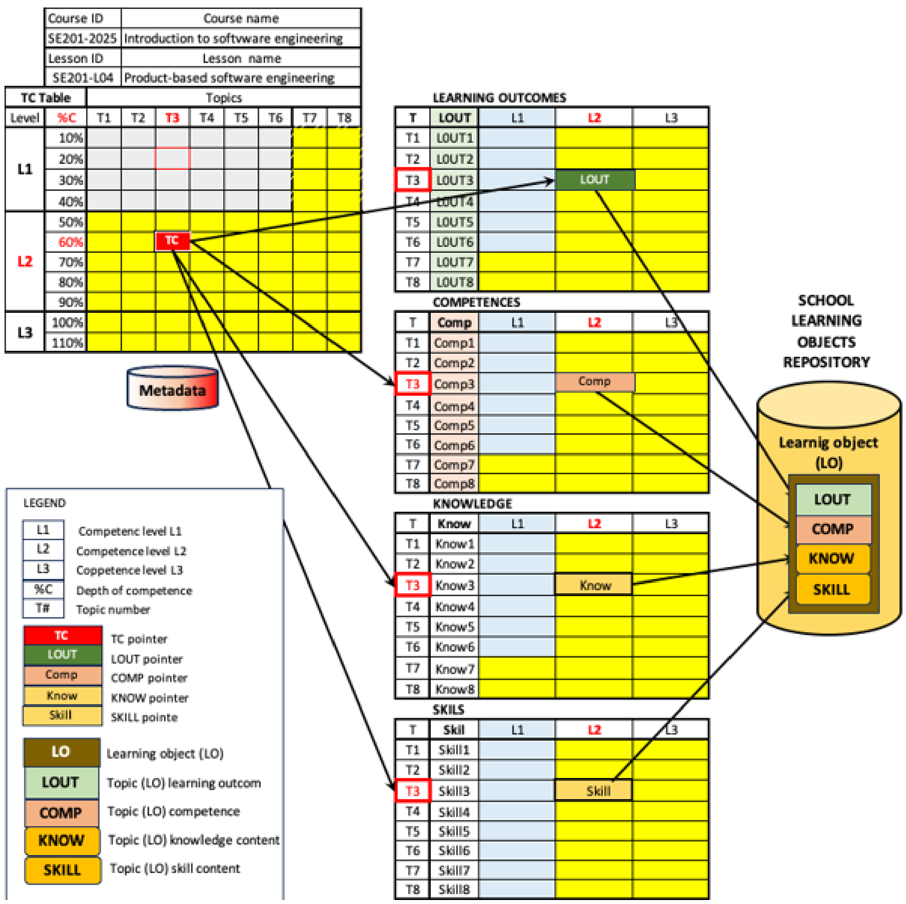
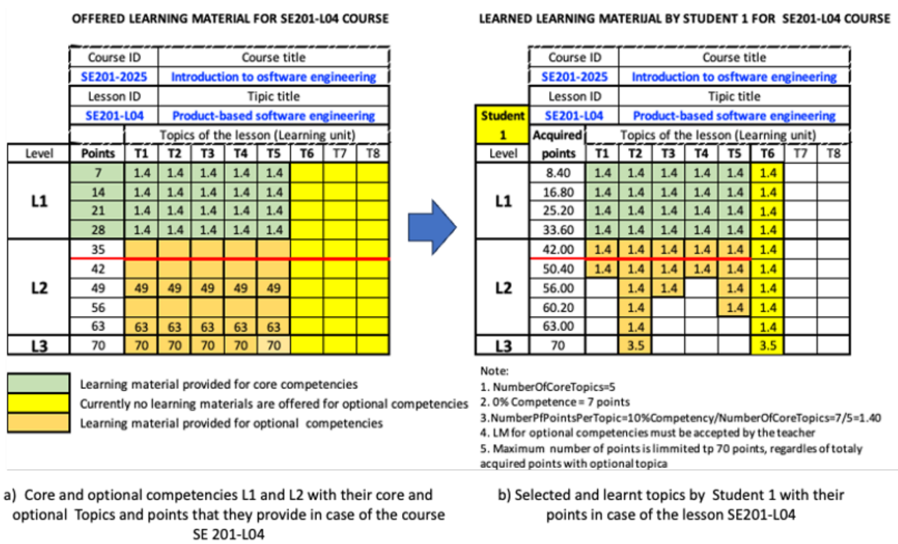


Fig. 15. Conceptual data model for creation of learning topics of a lesson by selecting their level of offered competences.

As explained in the notes on Figure 16, the 7-point competencies are represented as table rows. Each new row shows a new (deeper) competency that LM offers. Each field in a row gives 1.4 points (out of a total of 70 points). This number is the ratio between the 7 points related to each row (10% of 100% of the total planned competence) and the number of core topics (5 in the case of lesson SE201-L04). When new optional topics are added (such as T6 in Figure 16b), the same number of points are used as for core competencies (1.4 in the case shown). Since the course can provide a maximum of 70 points for all pre-exam tasks, the learned competences and acquired points of the offered learning material must not provide more than 70 points for the pre-exam tasks even in cases where they offer more than 70 points.



Course ID	Course title
SE201-2025	Introduction to osftware engineering
Lesson ID	Tipic title
SE201-L04	Product-based software engineering

Level	Points	T1	T2	T3	T4	T5	T6	T7	T8
		Topics of the lesson (Learning unit)							
L1	7	1.4	1.4	1.4	1.4	1.4			
	14	1.4	1.4	1.4	1.4	1.4			
	21	1.4	1.4	1.4	1.4	1.4			
	28	1.4	1.4	1.4	1.4	1.4			
L2	35								
	42								
	49	49	49	49	49	49			
	56								
	63	63	63	63	63	63			
L3	70	70	70	70	70	70			

a) Core and optional competencies L1 and L2 with their core and optional Topics and points that they provide in case of the course SE 201-L04

Student 1	Course ID	Course title
	SE201-2025	Introduction to osftware engineering
	Lesson ID	Tipic title
	SE201-L04	Product-based software engineering

Level	Acquired points	T1	T2	T3	T4	T5	T6	T7	T8
		Topics of the lesson (Learning unit)							
L1	8.40	1.4	1.4	1.4	1.4	1.4	1.4		
	16.80	1.4	1.4	1.4	1.4	1.4	1.4		
	25.20	1.4	1.4	1.4	1.4	1.4	1.4		
	33.60	1.4	1.4	1.4	1.4	1.4	1.4		
L2	42.00	1.4	1.4	1.4	1.4	1.4	1.4		
	50.40	1.4	1.4	1.4	1.4	1.4	1.4		
	56.00		1.4	1.4		1.4	1.4		
	60.20		1.4			1.4	1.4		
	63.00		1.4				1.4		
L3	70		3.5				3.5		

b) Selected and learnt topics by Student 1 with their points in case of the lesson SE201-L04

Fig. 16. Offered (a) and selected and learned (b) student 1 learning material for subject lesson SE201-L04.

5 Use of Prompt Engineering for preparation of teaching materials by teachers

5.1 Preparation of learning materials for the mandatory topics of each lesson at all three levels of competencen

As seen in Figure 17, the teacher, before starting to use Prompt Engineering for the preparation of learning materials, must have in mind the competencies that the course and each of its lessons should provide to the student (Fig. 1), he analyzes text-books used for similar courses by renowned universities and publishers in the world, analyze reference syllabi, used by renowned universities. In addition, he searches for information sources on the web, as well as other learning resources.

Based on this preparation, the teacher creates the curriculum of his/her course and defines the teaching topics for each lesson (as a learning unit) and, in accordance with Figure 5, creates the structure of each lesson with its topics at all three competence levels. For all mandatory topics, which are defined for level L1 (Fig. 16.a), the teacher should provide deeper competences for each mandatory topic at levels L2 and L3 as well. In this way, he/she enables the student to obtain a maximum of 70 points on each topic. Based on this preparation, the teacher creates the syllabus for his course, defines teaching topics for each lesson and, in accordance with Figure 5, creates the structure of each lesson with its topics at all three levels of competence, and students, following the learning procedure shown in Fig. 17 can earn up to 70 points.

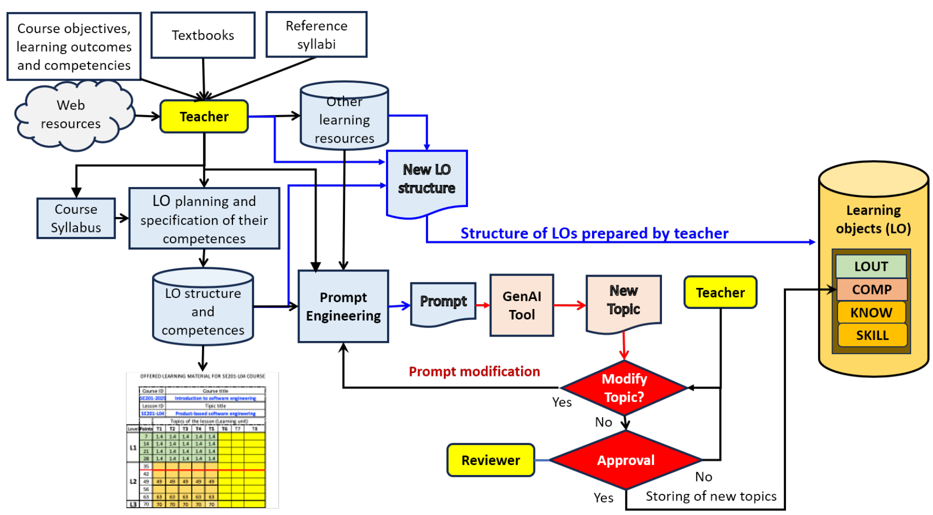


Fig. 17. Hybrid procedure for development of core L1 and supported optional topic in levels L2 and L3.

If, due to the specificity of mandatory topics, the teacher cannot provide the possibility of obtaining 70 points in this way, then he/she is obliged to provide learning mate-rial for optional topics at the L1, L2 and/or L3 level (Figure 16.b). It is important that the sum of points on all mandatory and optional topics, and all levels of competence, is at least 70. If new optional topics are added over time, the total sum of points is greater than 70. In these cases, the student does not have to earn all possible points (70 or less) on mandatory topics, because by working on optional topics, he/she earns. in total, the required 70 points, or less, if he/she did not manage to achieve the maximum 70 points.

Table 1 shows the learning materials and pre-exam assignments that the teacher must provide for each lesson to students in SE201 Introduction to Software Engineer-ing. As each lesson can be gradually expanded with new optional topics to the level of L1, L2 and L3, the work of the teacher is too much, and it becomes unacceptable. In order for it to be acceptable, in addition to teachers, it is planned that students who want to earn new and specific competences, they can develop appropriate learning materials under the supervision of teachers.

The teacher and students may use generative intelligence, according to university rules, in order to reduce their workload, and to significantly increase the efficiency of development of new, optional learning objects at levels L1 and L2, and essays at level L3. The university's internal standards define the technical, ethical and legal requirements that the author of learning materials must apply when preparing learning materials. Compliance with these requirements in case the author is a student is checked by teacher, and when the author is a teacher, it is checked by a specially appointed reviewer. More than 60 criteria for evaluating the quality of the prepared teaching material have been defined, adapted to the specifics of the course. Depending on the quality assessment, the financial compensation to the teacher, as the author of the learning material, is determined every year, in the event that he has re-vised the learning material at least 5%.

Table 1. Learning materials, pre-exam assignments and exam points for SE201.

Learning material offered to students and pre-exam assignments	Number of instances	Points	Total points
Online learning material offered to all students			
1. Video introduction L1	1		
2. Video lecture L1	1		
3. Multimedia web lecture L1	13		
4. Multimedia web lecture L2	13		
5. PDF lecture L1	13		
6. PDF lecture L2	13		
7. PPT lecture presentation L1	13		
8. PPT lecture presentation L2	13		
Pre-exam assignments			
1. Question tests L1	13	1	13
2. Question tests L2	13	1	13
3. Task test L1	3	1	3
4. Task test L2	3	1	3
5. Project (1, part) L1	1	8	8
6. Project (2, part) L2	1	12	12
7. Essays L3	2	4	8
8. Student activity	1	10	10
Total pre-exam points:			**70**
Examination			
Examination: 13 questions and 3 tasks	1	30	30
TOTAL POINTS:			**100**

When preparing the learning material, it is taken into account that it is based on the recommended textbooks that have been selected for each subject. The scope of the learning material is taken into account, depending on the specifics of the subject, because students today are looking for as short learning materials as possible to read or watch video clips. The volume of learning material for compulsory topics at the L1 level, as a rule, is at least two times smaller than the corresponding learning material prepared for the L2 level, which also includes the learning material prepared for the L1 level. A special edition of teaching material at the L1 level is prepared for students who want to just pass the exam, while securing the minimum required 35 points. Within these 35 points, at least 80% of the points must be provided by learning on compulsory topics (which provide up to 28 points). The difference of up to 35 points students should earn from the optional teaching material intended for the L2 level, of his own choice.

In addition to the summary presentation of the learning material for which the teacher used the selected textbooks, the teacher can add his own contributions to the content of the teaching material, as well as processed parts generated by the GenAI tool. When creating online teaching material, the teacher is obliged to indicate the use of sources, and to emphasize especially the parts that are not contained in the recommended textbooks (if there are such parts), i.e., that he created himself or was created by the GenAI tool. He is obliged to carefully study and approve the learning material generated by GenAI, which are not contained in the selected textbooks. This part of learning material is very limited in relation to the one that is generated on the basis of the specified textbooks, which should be major source for learning material.

After preparation, the teacher defines the structure of each lesson (as in Figures 5, 10 and 16), and in accordance with the recommendations for writing prompts (prompts), specifies a prompt on the basis of which the selected generative artificial intelligence tool GenAI generates the text (with images) for the topic for which the prompt is specified. It usually takes several iterations of improvement of the prompt to get acceptable learning material In this way, texts are generated for all topics of one lesson, and then the procedure is repeated for each lesson. If the teacher is well trained in preparing prompts for his course, the application of generative intelligence can drastically reduce the time for preparing learning materials, i.e. learning objects for all three levels of competence.

In accordance with the university's standards, the final approval for the teaching material is given by a specific reviewer who thus checks whether the teacher has done his work in accordance with the internal standards for the preparation of learning materials and with the prescribed criteria for evaluating the quality of teaching materials. Every academic year, the process of applying and evaluating the revised learning material is repeated. Within 7 years at most, the teacher is required to prepare completely new teaching material (at least 80% of the content should be new). In the event that due to some reasons (new technologies, new requirements, etc.) there is a need for significant changes to the existing learning material, then after the dean's approval, the teacher prepares a new version of the learning material. Changes in the content of the learning material are controlled by software. In addition to learning materials, the teacher is required to prepare tests and pre-exam assignments for stu-dents for each learning topic of a lesson and for all course lessons, using appropriate prompts and GenAI tools (Fig. 18).

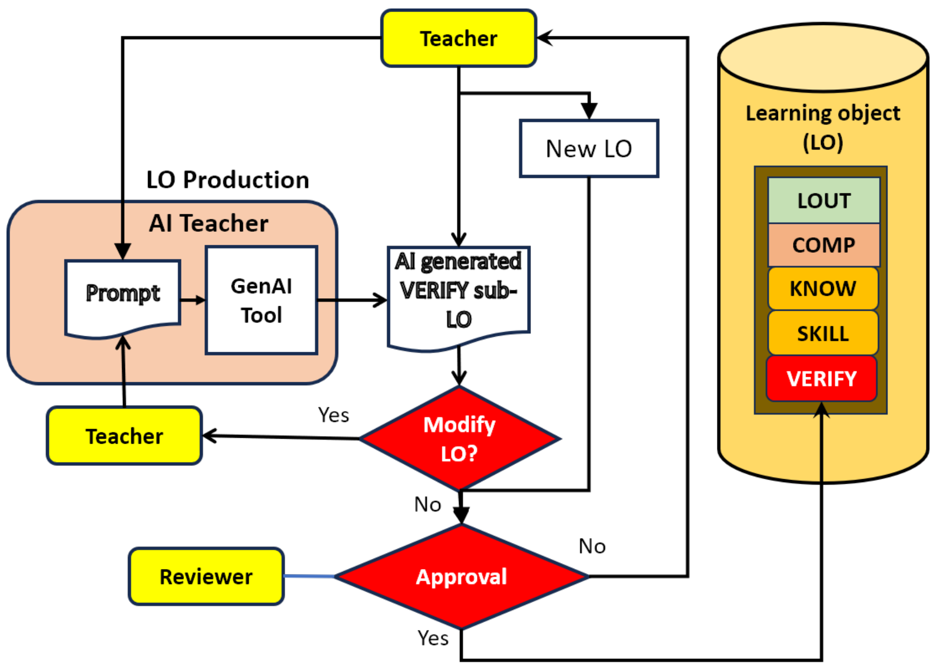


Fig. 18. Specifying of VERIFY elements of learning objects and their contents for L1 and L2 competence level.

In a similar way, the teacher must prepare at least five times more exam questions and tasks than there are topics in all lessons of the course (Figure 19). On the exam, each student receives randomly generated special questions and tasks, it depends on the portfolio of competences acquired in the subject, i.e., the topics and their depth, which he chose and studied, taken from student learning object repository. In this way, the repository of all learning objects contains all the necessary learning materials, for verifying what has been learned and the database of questions and tasks for the exam.

All learning materials are available to students, but they cannot access learning verification materials (e.g. question or task tests, exam questions and tasks). They only can access the obtained grades and possible explanations of the obtained grades.

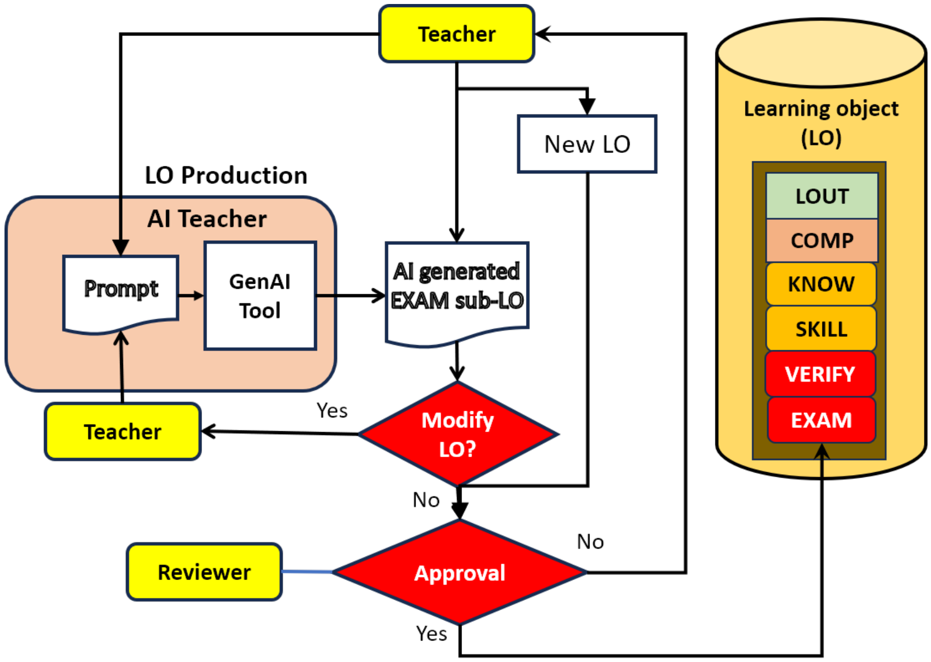


Fig. 19. Specification of the database of exam questions and tasks.

5.2 Preparation of learning materials for optional topics of each lesson

As already mentioned, in addition to compulsory courses, optional courses can be offered at L1, L2 and L3 levels. For some of them, the teacher can prepare the necessary learning contents, and in most cases, the student is expected to prepare them, under the teacher's supervision, according to his wishes. Based on the student's proposal for new competences, the teacher should determine optional topics for each lesson that the student has to develop (Figure 20).

In a demonstrative way, the structure of each lesson (teaching unit) is determined, which is the basis for creating learning materials. As you can see, in addition to the mandatory topics, optional topics for levels L1, L2 and L3 are defined for each les-son. Over time, that number of optional topics will probably increase, in accordance with the wishes of students and employers to acquire some specific competencies.

Now the student can create the missing mDita lessons with their own topics and use their learning objects to prepare tests and the exam questions and tasks. It is important to note that each student must receive different exam questions and tasks in the exams, depending on the optional competences he has chosen and studied, which contain his special lessons and topics, as well as special exam questions and tasks. That is why there are student learning facilities that support all these specificities and are necessary during the studies

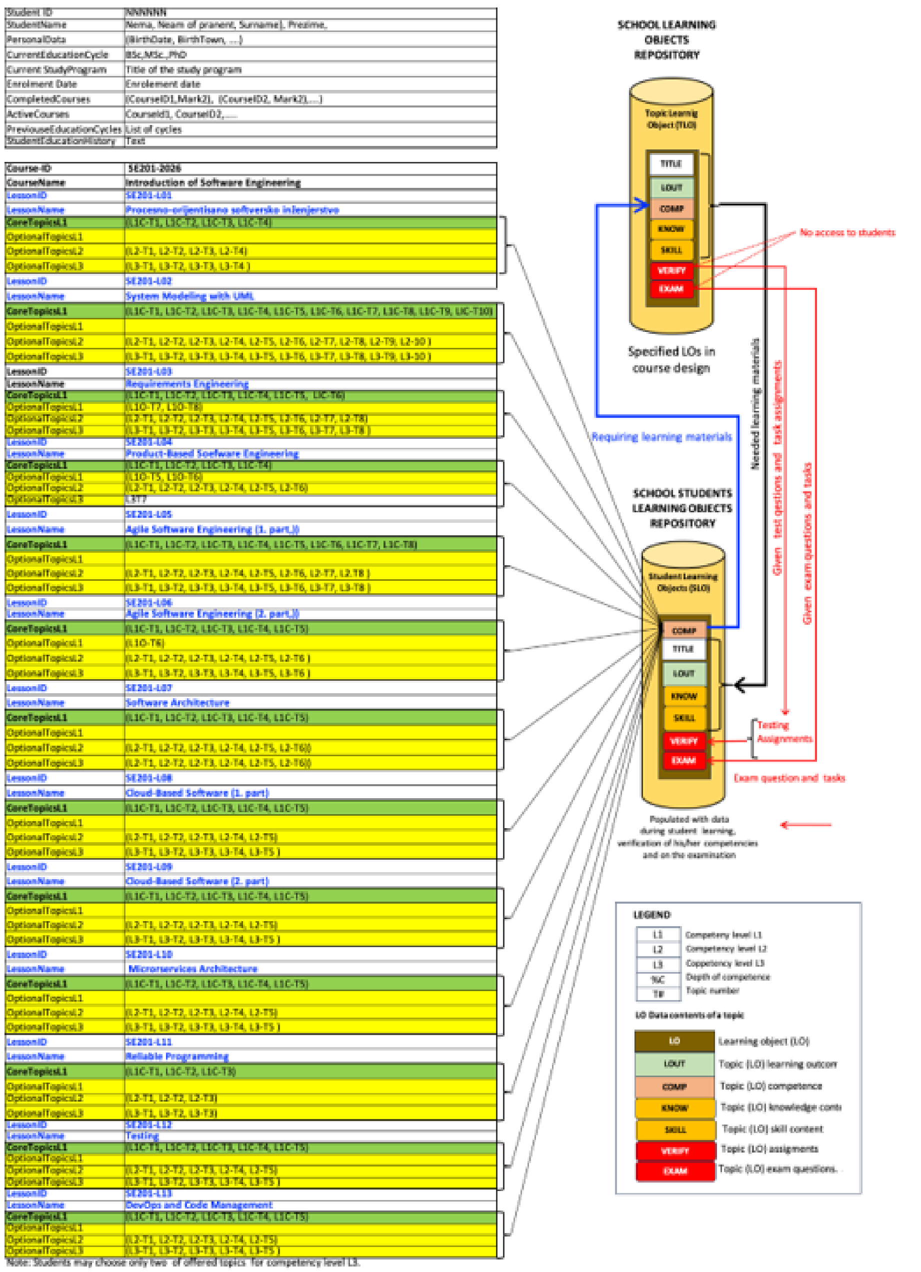


Fig. 20. Creation of Student Learning Objects.

of every student. Of course, created learning objects that have been prepared and placed in the university repository of learning objects can be shared and used by other students who have chosen the same optional competencies and their learning objects later during their studies.

Figure 21 shows, in a simplified way, the hybrid process of preparing and submit-ting learning material for optional and specific topics of each lesson in the course. Teaching materials are placed in the repository of learning objects of the university, but links to these learning materials are placed in a separate student's repository of links to learning materials, which each student has and which is filled during the entry of new learning materials. In this way, it is known exactly which optional teaching materials the student chose during his studies in each subject, in each lesson.

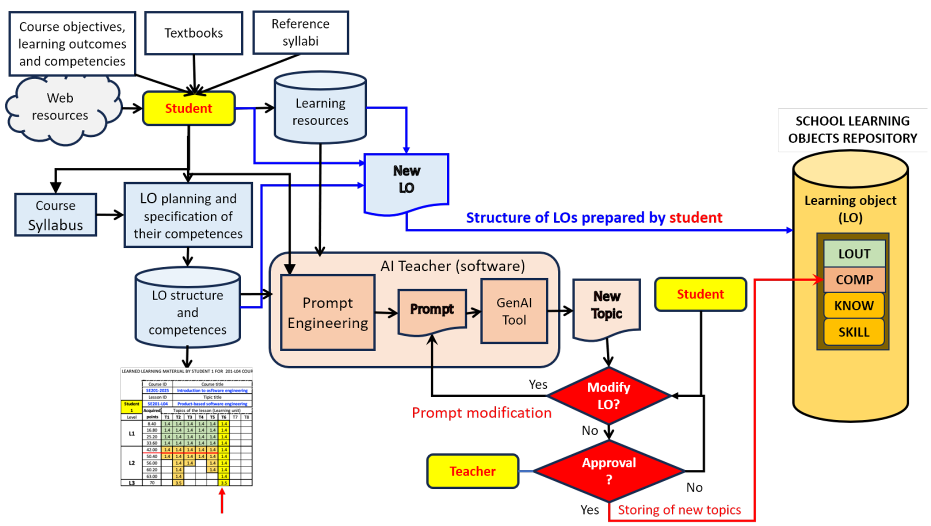


Fig. 21. Hybrid development procedure of optional teaching materials of specified optional competencies (L1, L2, L3).

Figure 22 shows, in a simplified way, a hybrid process in which a student prepares learning material for optional competencies for one or more courses, their lessons and topics. He also prepares tests and pre-exam assignments for these lessons and topics. The teacher verifies the defined pre-exam tasks that are part of the learning material and are given for further use by all students who choose the same specific competency.

Figure 23 shows a similar procedure for preparing all possible exam questions and tasks (with correct solutions) that the student prepares for the optional competencies and topics he has chosen. After verification by the teacher, the tests with exam ques-tions and tasks are placed in the repository of the university's learning facilities, and they become inaccessible to students. The student who

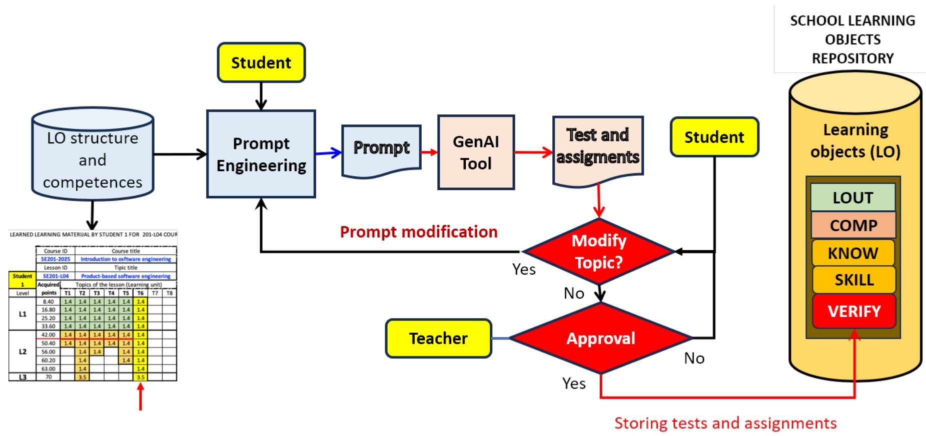


Fig. 22. Preparation of pre-examination requirements for optional competencies and topics, and their placement in the repository of the university's learning facilities.

prepared them is obliged to sign a statement on the confidentiality of the exam questions and answers he pre-pared (as part of his learning process).

6 Automatic generation of prompts for GenAI tools when preparing teaching materials

Applying prompt engineering increases work efficiency, but it is highly dependent on the skill of the user of the GenAI tool. Depending on the problem he wants to solve by applying it, he must apply the appropriate prompt structure. Even so, it often takes several iterations to perfect it and provide application results of the chosen GenAI tool that are good for practical application in a particular field. The question arises, is the field of application of the GenAI tool in the preparation of learning materials possible for the creation of software that will itself, without human intervention, generate the appropriate prompt for the application of the selected GenAI tool in certain narrower areas of education, and especially in certain narrower areas of higher education?

These researches are now in the preparatory phase at our university. Figure 24 shows a conceptual model of the process of possible automatic obtaining of instructions in the preparation of learning materials in higher education. The goal is not to completely remove the teacher from the process, but to assign him an appropriate role that slows down the efficiency of the work as little as possible, and can significantly affect the quality of the final results of the application of the selected GenAI tool.

As you can see, the preparatory part of the process is the same as in the case of applying prompt engineering. The difference is that instead of using prompt engineering, here an AI software module is introduced to automatically generate the prompts. It should provide fully automatic generation of the needed prompt,

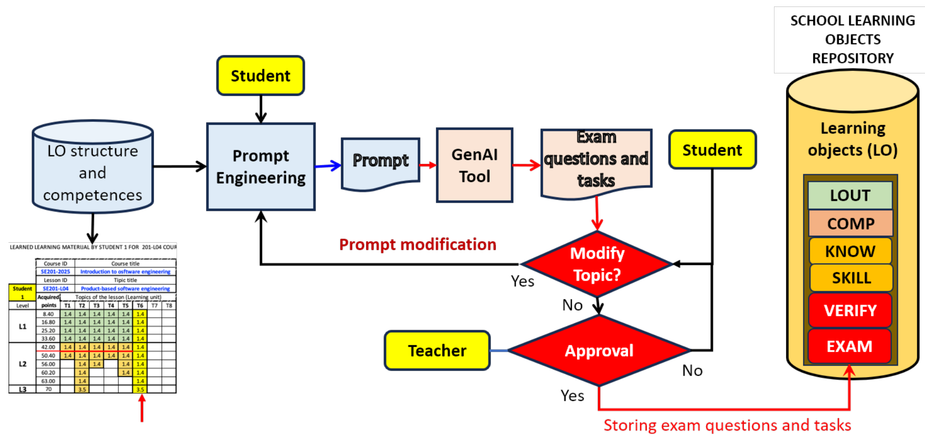


Fig. 23. Preparation of exam questions and assignments for optional competencies and topics, and their placement in the repository of the university's learning facilities.

with which GenAI will generate the desired learning material. The prompt is the query text based on which the GenAI tool produces, in our case, the text of certain topics. So, we have two cycles of iterative improvement of the query, but also the software which, now, instead of the teacher, should produce the appropriate text of the query (prompt).

The acceptability of the text of the material for the topic is assessed by the teach-er. If he is not satisfied, he now does not make direct changes in the text of the prompt, but in the software that should generate a new prompt that will produce either an acceptable prompt or a new modification of the software, until an acceptable result is obtained - acceptable teaching material in the opinion of the teacher. The teacher's decisions and the obtained result - new teaching material, are checked by a reviewer whose role is to ensure, in addition to the teacher, the obtained quality of the teaching material, but also to check whether the teacher carries out the prescribed procedure according to the given rules and. restrictions. The human factor is often an element of risk in maintaining the quality of the process of preparing teaching mate-rials, so this additional control is desirable and is standardly applied in the preparation of learning materials at BMU.

If the reviewer has objections to the quality of the generated learning material, he submits his objections to the teacher, who then once again modifies the prompt gen-eration software that generates a new prompt and repeats with which the GenAI tool generates new teaching material with a single topic. If, after that, the reviewer has no more objections, the obtained learning material is archived in the university's repository of learning materials At the same time, the software for automatically generating prompts is also archived.

The question arises: what is gained by specifying software to generate prompts in-stead of directly and iteratively specifying the prompts that will produce ac-

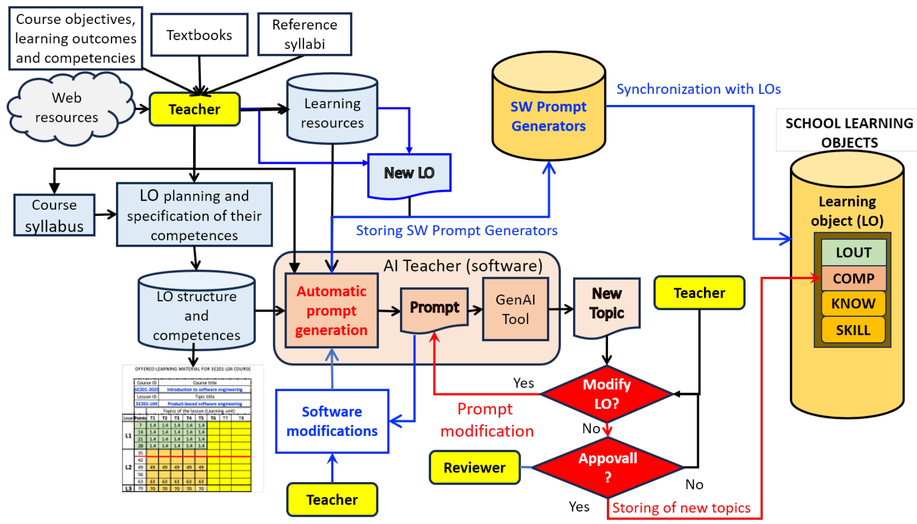


Fig. 24. Hybrid procedure of automatic development of learning materials for specified L1 and L2 competencies.

ceptable learning material with a selected GenAI tool ? This research is currently in its infancy, but the smart software is expected to provide a much greater speed of producing the needed prompt than when using manually and iteratively specification of a needed prompt. Future research will hopefully confirm these expectations. Similarly, as for generating teaching materials for learning, teaching materials can be generated for verification of what has been learned (tests, tasks, and other pre-exam home-work assignments of students), as well as tasks and questions for exams (Figure 25).

When a student decides to create optional learning material for optional specific competencies, he/she uses the same procedure for automatic prompts generation as the teacher (Fig. 24 and Fig. 25). The role of teacher is assumed by the student, and the role of reviewer is assumed by the teacher.

7 Using the system for personalized learning by students

The student uses the personalized learning system in two ways:

- Use as an author of teaching material, if he has been assigned to prepare, with the help of the teacher, teaching material for a specific competence for which he wants to acquire competence, and
- Use as a student who has chosen an optional competency for which there is already teaching material, noted by the teacher, or a student who first re-quested the acquisition of that specific competency.

Figure 26 shows the procedure for preparing the teaching material in the case that the student has the role of the author of the teaching material, which

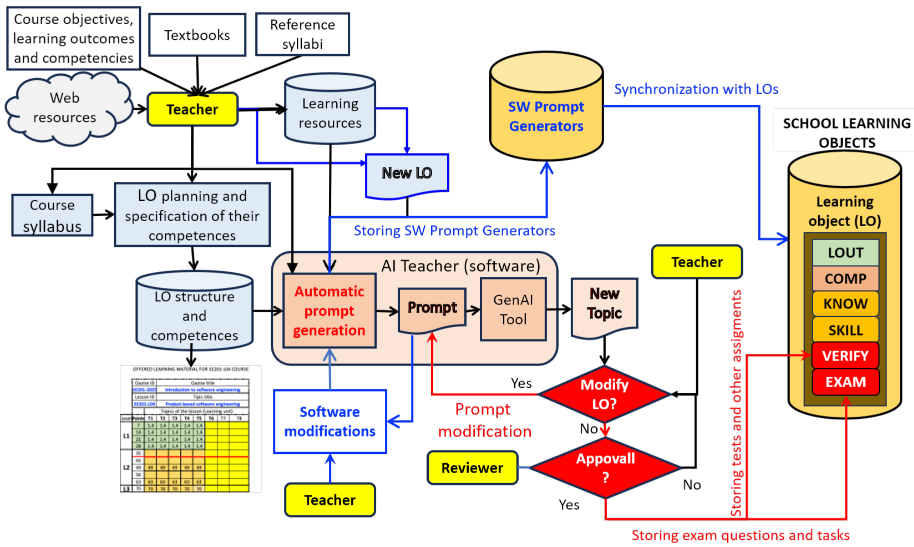


Fig. 25. Procedure for automatic generation of prompts for generating tests and tasks for pre-exam assignments, as well as test questions and tasks for exams.

includes the preparation and evaluation of the students' pre-exam assignments, as well as their exam questions and assignments. As already mentioned, the procedure is the same as in the first case of the teacher. The only difference is that the teacher, in addition to his advisory role, also has the role of a reviewer of the received teaching materials.

Figure 27 shows the procedure of using the system for personalized learning in the case that the student has his usual and most common role, i.e. to learn and use developed learning materials. In this case, through the student interface for using the iLearn system, he searches, finds and access the lesson he needs to learn.

As each student has his own specific portfolio of competencies that he acquires (due to the use of optionally offered competencies and learning materials), the student must have his own special repository, which contains links to learning materials that are stored in the university repository. These links are created when the student determines, for each year of study, which optional competencies he wants to ac-quire. Most often, students can access prepared learning materials, for which the necessary links are then created. If they made a request to acquire optional specific competences for which the needed learning material was not developed yet, then they need to develop learning materials for new optional topics according the procedure of creating learning material described in Figure 26. Students also may prepare pre-exam assignments and, exam questions and tasks, verified by their teacher. On exams, students may receive only exam questions and tasks that correspond to their portfolios of acquired competencies.

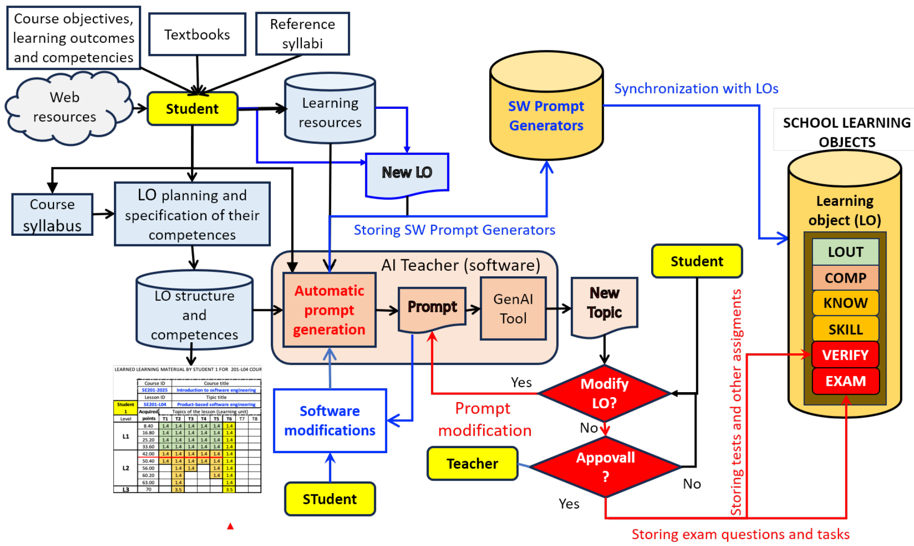


Fig. 26. Hybrid procedure of automatic development of learning materials for specified L1 and L2 competenciess.

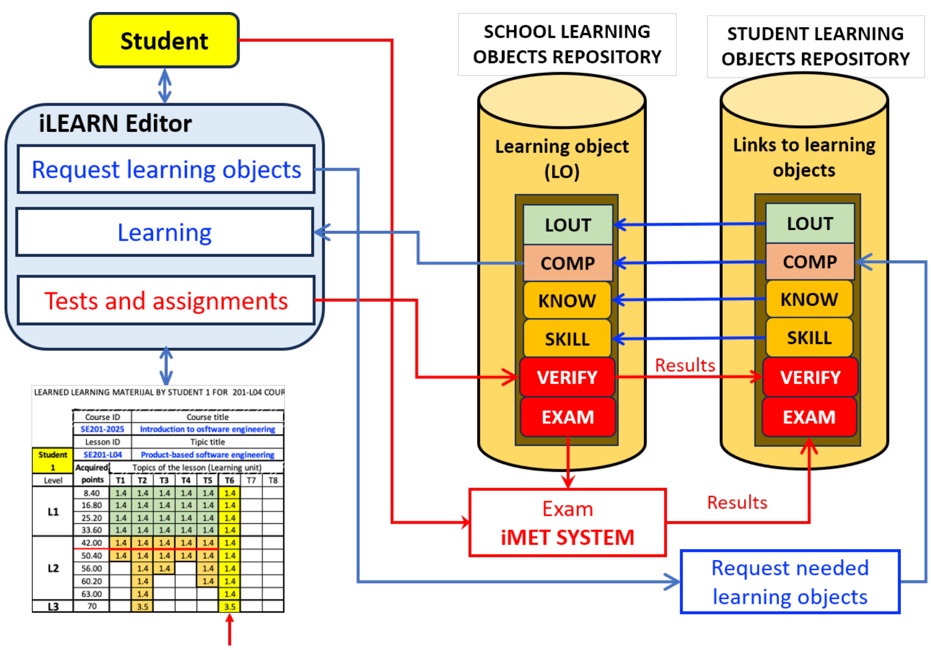


Fig. 27. Student needed learning objects, gets and learn their content, get and submit tests and assignments and do exams.

8 Conclusion

The paper presents a conceptual model of a system for hybrid personalized learning guided by competences and supported by the GenAI tool of the generative artificial intelligence system. The aim of the paper was not to provide an overview of research in the field of research and development of systems for personalized learning, but to devote itself entirely to a more detailed exposition of the concept of such a system that is under development and partially implemented at the Metropolitan University in Belgrade.

The paper presents the concept of the system as a whole, i.e. which provides conceptual solutions to the most important research and development challenges for the application of personalized learning in higher education. In the last two academic years, the concept was tested in the case of the application of prompt engineering using several models of generative intelligence, but GenAI tool of the CharGPT4.0 system was used. The results of the experimental application on the SE201 course showed that a significant increase in efficiency is achieved in the preparation of teaching materials, including the preparation of tests, but that the teacher must be well trained in the preparation of prompts. The goal is to train the students to use GenAI tools for the needs of their future job, but to do so with a full understanding of the problem they are solving, i.e. to acquire the appropriate competencies that are defined for each topic studied in the course.

The challenge is also how to hold F2F teaching in classrooms when students study different teaching topics, within their personalized learning in each course. The development of pedagogical models of F2F teaching on personalized syllabi is also necessary. The current F2F lectures can only cover the compulsory part of the program, and for the personalized parts, the practice of student consultation with the teacher and his assistants must be introduced.

By applying engineering prompts, useful experiences were gained, and encouraging results were achieved, especially in the efficiency of creating learning materials and tests. This enables rapid adaptation of syllabi and their courses to the demands of students and employers.

In the next phase of the research, research into the automatic generation of prompts will be carried out using intelligent software for generating appropriate prompts. The following research and development activities are planned:

1. Development of a software system that implements the set conceptual model of learning personalization using prompt engineering
2. Research and development of an intelligent software generator of the required prompts for obtaining learning materials necessary for the acquisition of new and required competencies.
3. Pilot application of the developed system for personalized learning based on the conceptual model presented here in different syllabi and groups of specific courses, in order to determine new challenges in application and find new solutions based on new experience in application.

These researches are demanding and extensive, because BMU always strives to apply the results of its research projects in teaching in all university programs. As of this academic year, all learning materials (except in the education field of art and design) support the application of three levels of competence. In the preparation of learning materials, the concepts presented in this paper are applied, and for each teaching unit, the goals, learning outcomes and competencies that the teaching unit must provide to students are clearly defined, while their mapping is carried out starting from the reference corpus of knowledge, as described in this paper.

BMU is open to cooperation in further research in the field of application of artificial intelligence in the personalization of learning with other universities, with the support of appropriate funds to support planned research.

Acknowledgments: This study was not funded by any external source.
Disclosure of Interests: The author has no competing interests to declare that are relevant to the content of this paper.

References

1. Ayeni, O.O., Al Hamad, N.M., Chisom, O.N., Osawaru, B., Adewusi, O.E.: Ai in education: A review of personalized learning and educational technology. GSC Advanced Research and Reviews **18**(2), 261–271 (2024)
2. Cardone, M., Rodriguez, R., Ishmad, K.: Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations. Tech. rep., U.S. Department of Education, Office of Educational Technology, Washington, D.C. (2023)
3. Chan, C.K.Y., Colloton, T.: Generative AI in higher education: The ChatGPT effect. Taylor & Francis (2024)
4. Chen, L., Zhang, Y.: The impact of delay on user engagement in educational technology. IEEE Transactions on Learning Technologies (2024)
5. Cheng, B., Wang, M., Yang, S.J., Peng, J., et al.: Acceptance of competency-based workplace e-learning systems: Effects of individual and peer learning support. Computers & Education **57**(1), 1317–1333 (2011)
6. Cooper, G.: Examining science education in chatgpt: An exploratory study of generative artificial intelligence. Journal of science education and technology **32**(3), 444–452 (2023)
7. DeepSeek: DeepSeek (2025), https://www.deepseek.com/en, [Online; accessed 11-Sep-2025]
8. Adinda, D.: A competency-based approach to support e-learning during the covid-19 situation. In: 20th European Conference On E-Learning (2021)
9. Domazet, D., Gavrilović, N.: Personalization of e-learning by using of learning processes with branching. In: EDULEARN19 Proceedings. pp. 10432–10442. IATED (2019)
10. Domazet, D., Gavrilović, N., Jović, J.: Efficiency of development of e-learning materials based on learning objects. In: Proceedings of the 10th International Conference on eLearning (eLearning-2019). pp. 26–27 (2019)

11. Sluijsmans, D.M., Prins, F.J., Martens, R.L.: The design of competency-based performance assessment in e-learning. Learning environments research **9**(1), 45–66 (2006)
12. EU Commision: Law on artificial intelligence. Tech. rep.
13. Farahani, M.S., Ghasmi, G.: Artificial intelligence in education: A comprehensive study. In: Forum for Education Studies. vol. 2, pp. 1379–1379 (2024)
14. García-Peñalvo, F.J., Llorens Largo, F., Vidal, J., et al.: The new reality of education in the face of advances in generative artificial intelligence. In: Asociación Iberoamericana de Educación Superior a Distancia (AIESAD) (2023)
15. Giannini, S.: Generative ai and future of education. Tech. rep., UNESCO (2023), https://unesdoc.unesco.org/ark:/48223/pf0000385877
16. Gocen, A., Aydemir, F.: Artificial intelligence in education and schools. Research on Education and Media **12**(1), 13–21 (2020)
17. Goel, A.K., Polepeddi, L.: Jill watson: A virtual teaching assistant for online education. In: Learning engineering for online education, pp. 120–143. Routledge (2018)
18. GROK-DeepSearch: GROK-DeepSearch (2025), https://grok.com/, [Online; accessed 11-Sep-2025]
19. Harry, A.: Role of ai in education. Interdiciplinary Journal & Hummanity (INJURITY) **2**(3) (2023)
20. Holmes, W., Tuomi, I.: State of the art and practice in ai in education. European journal of education **57**(4), 542–570 (2022)
21. Hsu, Y.C., Ching, Y.H.: Generative artificial intelligence in education, part one: The dynamic frontier. TechTrends **67**(4), 603–607 (2023)
22. Hwang, G.J., Chen, N.S.: Exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. Educational Technology & Society **26**(2) (2023)
23. Jovic, J., Milic, R.M., Cvetanovic, S., Gavrilovic, N., Domazet, D., Nikolic, R.: Intelligent Recommender System for Personalized Online Learning (IREP). In: Proceedings of the Eleventh Internacional Conference on E-Learning. pp. 22–29 (2020)
24. Jović, J., Milić, M.R., Cvetanović, S.: Multidimensional concept map representation of the learning objects ontology model for personalized learning. Journal of Internet Technology **24**(5), 1043–1054 (2023)
25. Leyking, K., Chikova, P., Loos, P.: Competency—and process-driven e-learning—a model-based approach. Electronic Journal of e-Learning **5**(3), pp183–194 (2007)
26. Li, D., Zhang, H.: Adaptive learning systems: Impact on student outcomes. Computers and education (2023)
27. Mao, J., Chen, B., Liu, J.C.: Generative artificial intelligence in education and its implications for assessment. TechTrends **68**(1), 58–66 (2024)
28. MarketsandMarkets: Artificial intelligence market in the education sector 2023-2030. Tech. rep. (2023), https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-education-market-200912291.html
29. McGehee, N.: Ai in education: Student usage in online learning. Michigan Virtual Learning Research Institute (2024)
30. OpenAI: Technical report: Performance and baseline evaluations of gpt-oss-safeguard-120b and gpt-oss-safeguard-20b. Tech. rep. (2025), https://cdn.openai.com/pdf/08b7dee4-8bc6-4955-a219-7793fb69090c/Technical_report__Research_Preview_of_gpt_oss_safeguard.pdf

31. Qadir, J.: Engineering education in the era of chatgpt: Promise and pitfalls of generative ai for education. In: 2023 IEEE global engineering education conference (EDUCON). pp. 1–9. IEEE (2023)
32. Ruiz-Rojas, L.I., Acosta-Vargas, P., De-Moreta-Llovet, J., Gonzalez-Rodriguez, M.: Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. Sustainability **15**(15), 11524 (2023)
33. Samala, A.D., Rawas, S., Wang, T., Reed, J.M., Kim, J., Howard, N.J., Ertz, M.: Unveiling the landscape of generative artificial intelligence in education: a comprehensive taxonomy of applications, challenges, and future prospects. Education and Information Technologies **30**(3), 3239–3278 (2025)
34. Thomson, S.R., Pickard-Jones, B.A., Baines, S., Otermans, P.C.: The impact of ai on education and careers: What do students think? Frontiers in Artificial Intelligence **7**, 1457299 (2024)
35. UNESCO: Global student survey on ai in education. Tech. rep. (2024), https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024
36. Vieriu, A.M., Petrea, G.: The impact of artificial intelligence (ai) on students' academic development. Education Sciences **15**(3), 343 (2025)
37. Walter, Y.: Embracing the future of artificial intelligence in the classroom: the relevance of ai literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education **21**(1), 15 (2024)
38. Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., Du, Z.: Artificial intelligence in education: A systematic literature review. Expert Systems with Applications **252**, 124167 (2024)
39. Yu, H., Guo, Y.: Generative artificial intelligence empowers educational reform: current status, issues, and prospects. In: Frontiers in Education. vol. 8, p. 1183162. Frontiers Media SA (2023)
40. Zhang, K., Aslan, A.B.: Ai technologies for education: Recent research & future directions. Computers and education: Artificial intelligence **2**, 100025 (2021)

Improving Student Engagement through Learning Analytics and Early Interventions with Learning Locker

Anđela Grujić[1][0009−0009−1098−4246], Jovana Jović[1][0000−0002−4204−0233], Mladen Opačić[1][0009−0000−6246−8002], Emilija Kisić[1][0000−0003−3059−2353], and Nemanja Zdravković[1][0000−0002−2631−6308]

Faculty of Information Technology, Belgrade Metropolitan University
Tadeuša Košćuška 63, 11000 Belgrade ,Serbia
andjela.grujic@metropolitan.ac.rs, jovana.jovic@metropolitan.ac.rs, mladen.opacic@metropolitan.ac.rs, emilija.kisic@metropolitan.ac.rs, nemanja.zdravkovic@metropolitan.ac.rs

Abstract. This paper presents the results of a pilot study conducted at Belgrade Metropolitan University within the framework of the Erasmus+ ISILA project which investigates how data-driven early interventions can enhance student engagement and academic performance in higher education courses. Three pilot implementations are carried out in three courses, where each course has integrated the University's Learning Management System, Learning Locker as a Learning Record Store, while Self-Regulated Learning (SRL) surveys are conducted to collect and analyze student activity data. Learning analytics dashboards are used to identify students at risk of disengagement or low achievement, prompting personalized and general interventions during key points in the semester. Results indicate that targeted communication, flexible deadlines, and additional learning sessions positively influenced engagement and submission rates. The study demonstrates how combining SRL data and learning analytics supports early identification of learning barriers and offers practical insights for improving academic outcomes through evidence-based decision-making .

Keywords: Learning Analytics · Early Intervention · Student Engagement · Self-Regulated Learning.

1 Introduction

In recent years, universities have increasingly relied on learning analytics (LA) to understand and improve how students engage with learning materials and digital environments. The collection and analysis of learner data, when aligned with pedagogical frameworks such as Self-Regulated Learning (SRL), enables educators to design timely and data-informed interventions that prevent dropout, increase motivation, and foster continuous engagement. Within the European Erasmus+ project ISILA, which aims to enhance the quality and sustainability of learning by developing and testing methodologies for early interventions based on

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.040G
Part of ISBN: 978-86-89755-37-4

learning, the authors implemented several pilot studies. For the study, data was used from various sources which comprise Belgrade Metropolitan University's learning ecosystem. Namely, we have use BMU Educational Information System (iMet), Learning Activity Management System (LAMS), and Learning Locker, an open-source Learning Record Store (LRS) that stores learner interactions in xAPI format. These pilots were conducted in three core Information Technology courses: Object-Oriented Programming 1, Fundamentals of Web Development, and Distributed Systems, which represent different levels of study and student experience. Together, the pilots provide a comprehensive overview of how learning analytics can be applied in both foundational and advanced programming contexts.

Student engagement is a persistent challenge in computer science education, especially in blended and online settings. Engagement is often multidimensional, encompassing behavioral participation, emotional involvement, and cognitive investment. When these dimensions decline, students risk academic underperformance or course withdrawal. The ISILA project's core hypothesis is that data-driven early intervention, based on real-time analytics and SRL feedback, can indeed mitigate these risks by identifying disengagement early and supporting students through personalized communication on one side, and utilizing adaptive instructional design as well.

Each of the three pilot courses follows a similar structure: data collection through the LMS and SRL surveys, creation of course dashboards in Learning Locker, and the design of both general interventions and personalized interventions. General interventions include announcements, extended deadlines, additional sessions, while personalized interventions include targeted emails or consultations. The timing and content of these interventions varied slightly across courses to reflect contextual needs and observed trends in engagement.

In the course CS101 Object-Oriented Programming 1, interventions are planned in weeks 6 and 13, focusing on students with missing activity or elevated anxiety levels. For the course CS105 Fundamentals of Web Development, three interventions are planned at weeks 9, 14, and 15, offering repeated opportunities for inactive students to rejoin the learning process. These two courses are for first-year students. Finally, the only course for students in the second year of studies, CS230 Distributed Systems, interventions are planned as for CS101. This multi-course implementation allows us to assess the effectiveness of early intervention strategies under diverse conditions. It also provides comparative insight into how different instructional approaches, such as in-person consultation versus email-based communication, impact re-engagement patterns. The goal of this paper is to synthesize findings across all three pilots to answer the following research questions:

1. How can learning analytics be used to identify early signs of student disengagement?
2. What forms of personalized and general intervention are most effective for improving student participation and performance?

3. How can SRL data complement quantitative learning analytics to guide more empathetic and context-aware academic support?

The paper is structured as follows. Section 2 outlines the research methodology, including the data sources, instruments, and analysis techniques. Section 3 describes the implementation process of the three pilots and summarizes the intervention designs. Section 4 presents the results and discussion, focusing on measurable engagement changes and qualitative feedback. Section 5 concludes the paper with implications for future applications of learning analytics in higher education.

2 Related Work and Methodology

2.1 Related Work

Learning analytics (LA) has emerged as a central component of educational innovation, enabling universities to transform raw digital traces of learning activity into actionable insights [1–3]. Defined by the Society for Learning Analytics Research (SoLAR) as "the measurement, collection, analysis and reporting of data about learners and their contexts", LA aims to improve both teaching practices and learning outcomes [4].

A growing body of research has explored the application of LA for early detection of at-risk students and targeted interventions [5, 6]. The authors of [7] demonstrated that the Course Signals system could predict students' academic risk and deliver proactive feedback that improved course completion rates. Similarly, in [8] showed that LMS interaction data could serve as a reliable indicator of engagement and academic success. More recent studies extend this approach by combining behavioral indicators with self-reported psychological data, linking analytics with motivational and affective constructs [9, 10].

Self-Regulated Learning (SRL) theory [11–16] provides a valuable framework for interpreting these analytics. SRL emphasizes learners' capacity to plan, monitor, and evaluate their own progress, a process that aligns well with data-driven feedback loops. In [17] and [18], the authors highlighted that dashboards supporting reflection on SRL processes increase students' metacognitive awareness and motivation. However, they also caution that analytics alone are insufficient without pedagogical scaffolding.

Recent European initiatives advocate for institution-wide frameworks that connect analytics to ethical governance, teacher training, and continuous improvement [19, 20]. Within this landscape, the ISILA project contributes by focusing not only on institutional adoption but also on practical intervention design, ultimately linking dashboard insights to individualized and empathetic student communication.

This paper builds on these foundations by presenting an integrated model that combines Learning Locker dashboards, SRL survey data, and context-sensitive interventions within three university-level IT courses from multiple sources at multiple universities, as shown in Fig. 1. Unlike prior studies that concentrate on a single course or technology, the ISILA pilots test the scalability of such interventions across multiple curricular contexts affected by external socio-political disruptions.

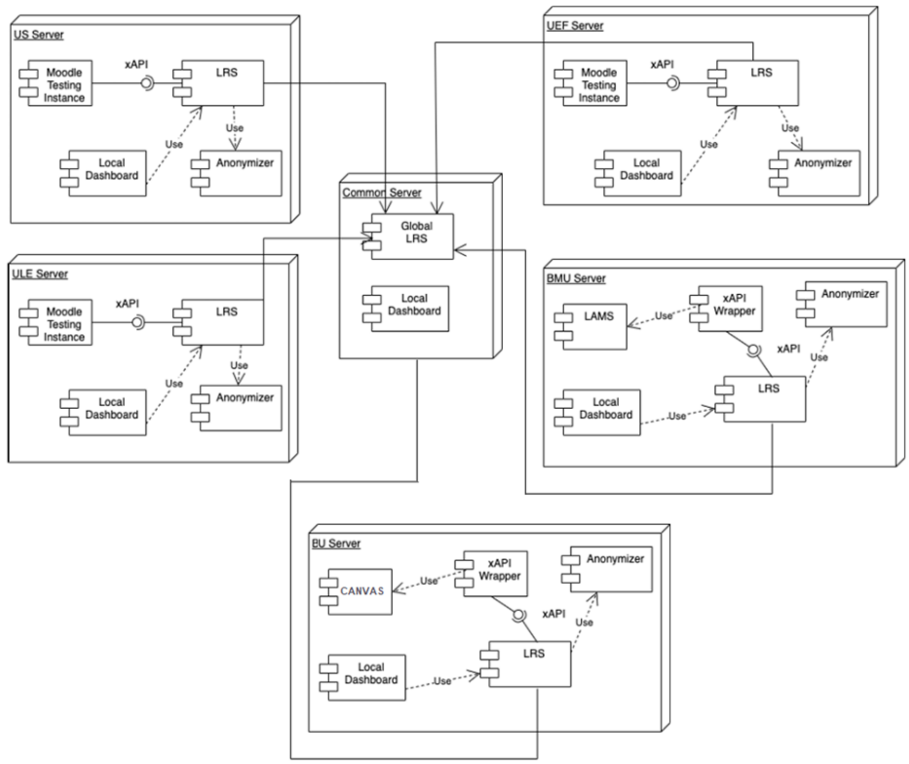


Fig. 1. The ISILA LRS model, comprising of multiple local LRSs connected to a global LRS running on a common server.

2.2 Research Design

The research followed a mixed-methods approach, combining quantitative analysis of learning analytics data with qualitative interpretation of student responses and instructor reflections. The pilot activities were implemented during the spring semester of 2025 at BMU as part of the Erasmus+ ISILA project. Each course served as a separate case study, but shared an aligned methodological framework. This design allowed both longitudinal tracking throughout the 15-week semester and cross-course comparison. We can summarize the phases as follows.

1. Data collection phase, in which extraction of student activity data from institutional platforms is performed.
2. Dashboard analysis phase, where Learning Locker gives the needed visualizations of learning patterns.
3. Intervention phase, performing the targeted actions which are emails, consultations, or extra sessions for students.
4. Evaluation phase, in which the analysis of behavioral changes and qualitative feedback after each intervention is performed.

3 Implementation and Interventions

3.1 Data Sources and Tools

Learning data were aggregated from several institutional systems, which include BMU's iMet Educational Management System (EMS), which is needed to provide enrollment records, assignment submissions, and grades. Next, we used the LAMS Learning Management System, where we captured weekly activity logs, exercise completion, and access time. For student communication, we used Discord. Finally, each week we gave students SRL surveys, focused on goal-setting, time management, anxiety, and perceived overload.

3.2 Participants and Context

The three courses involved 204 students in total, divided as follows.

- CS101 Object-Oriented Programming 1 with 90 students, first-year undergraduate, at two campuses.
- CS105 Fundamentals of Web Development with 61 students, first-year undergraduate at one campus.
- CS230 Distributed Systems with 53 students, second-year undergraduate students at one campus.

All participants were Computer Science or Information Technology majors. Participation in data collection and interventions was voluntary and conducted under explicit consent in accordance with institutional ethics guidelines. The Learning Locker is shown in Fig. 2.

3.3 Analysis Procedures

Data analysis combined descriptive statistics, dashboard interpretation, and thematic coding:

1. Quantitative metrics (number of active users, submission counts, average grades, SRL index values) were extracted from Learning Locker and analyzed in order to identify engagement trends before and after interventions.

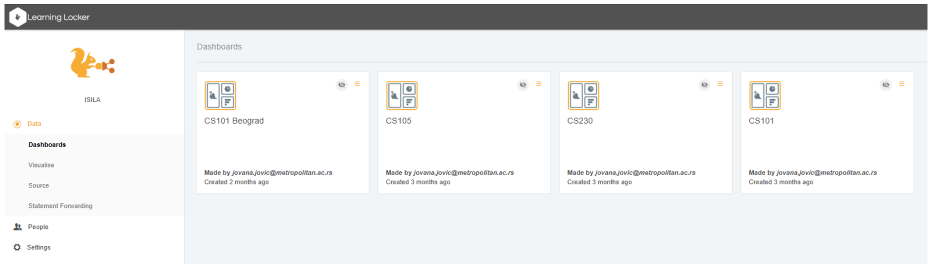


Fig. 2. Learning Locker dashboard.

2. Qualitative data, including open-ended survey responses and email correspondence, were coded using inductive thematic analysis [21] to identify recurring patterns related to motivation, stress, and perceptions of support.
3. Cross-case synthesis compared patterns among the three pilots to evaluate the consistency and scalability of early-intervention outcomes.

Reliability was ensured through double-checking dashboard exports and triangulating analytics data with SRL responses. Internal validity was strengthened by aligning interventions with clearly defined student categories (dropout risk, average activity, high achievers with anxiety, etc.).

Each course used a common set of institutional tools (iMet, LAMS, Discord, Learning Locker), but differed slightly in the frequency and type of interventions, as presented in Table 1.

Instructors were trained to interpret analytics dashboards and to act upon indicators of low engagement, such as missing assignment submissions, inactivity on the LMS, or negative SRL survey responses.

Table 1. Overview of pilot courses, number of students, and timing of interventions.

Course	**Students**	**Intervention Weeks**	**Tools Used**	**Key Focus**
Fundamentals of Web Development	61	Weeks 9, 14, 15	iMet, LAMS, Learning Locker	Flexibility and repeated support opportunities
Object-Oriented Programming 1	90	Weeks 6 and 13	iMet, LAMS, Learning Locker	Early detection and SRL-based communication
Distributed Systems	53	Weeks 6 and 13	iMet, LAMS, Learning Locker	Personalized feedback and stress management

Note that the implementation was heavily influenced by contextual factors: Serbia's ongoing student protests during the semester disrupted regular classes and reduced student motivation and attendance. Consequently, the pilots offered a unique opportunity to evaluate interventions under crisis conditions, testing whether timely communication and flexible scheduling could help mitigate external disruptions.

3.4 Data Collection and Dashboard Setup

Data were collected continuously from the institutional systems, first converted into the xAPI format, and then automatically exported to Learning Locker using system integrations and the `csv2xapi` converter developed under the ISILA project. Each student's activity, including login frequency, time spent on exercises, assignment submissions, and survey participation, was transformed into xAPI statements, allowing for cross-platform analysis. The Learning Locker dashboards were customized to visualize engagement trends at both course and individual levels. Key dashboard components included:

- Total time spent on the course per student
- Number of exercises and assignments completed
- SRL survey scores (motivation, anxiety, time management)
- Cumulative grade progress

Fig. 3. Learning Locker dashboards for CS230 after Week 6.

These dashboards were updated weekly, providing instructors with actionable insights into participation trends.

This early diagnostic phase was crucial for designing the first round of interventions, which aimed to re-engage inactive students before mid-semester deadlines.

3.5 General Intervention Strategy

Each pilot followed a similar three-step approach:

1. Identification, using dashboard data to detect low-engagement or high-anxiety students;
2. Classification, grouping students into categories according to performance and SRL data;
3. Intervention, sending personalized or general messages and providing additional support sessions.

The classification model distinguished five main student profiles:

1. No learning activity detected (dropout risk)
2. Average activity, missing SRL data
3. Below-average performance, high SRL
4. Average performance, high anxiety
5. High performance, high anxiety

This taxonomy allowed instructors to adapt both the tone and content of interventions. For example, students with no activity received empathetic outreach emphasizing re-engagement opportunities, while those with high anxiety but good performance received reassurance and stress management guidance.

In addition to individual emails, general interventions included announcements, deadline extensions, and group consultations designed to create opportunities for re-entry into the learning process.

Table 2. Example student classification and corresponding intervention types.

Student	Intervention 1: Week 6	Intervention 1 actions
S1	Average SRL. Submitted all exercises and high grades	NONE
S2	Average level of activity. No SRL data available.	CONTACT PERSONALLY
S3	No learning activity detected on the system – dropout.	CONTACT PERSONALLY
S4	No learning activity detected on the system – dropout.	CONTACT PERSONALLY
S5	No learning activity detected on the system – dropout.	CONTACT PERSONALLY

3.6 Course-Specific Implementation

CS101 Object-Oriented Programming 1
The CS101 course focused on core Java programming concepts such as classes, inheritance, and exception handling. Data analysis during the first six weeks revealed an alarming number of inactive students, with over 50% had no recorded activity in LAMS or iMet.

First Intervention (Week 6):
Personalized emails were sent to all students showing inactivity, missing SRL data, or high anxiety levels. The emails, written in a supportive tone, encouraged students to restart their coursework and offered additional help. Five different email templates were used, addressing dropout risk, low engagement, or anxiety

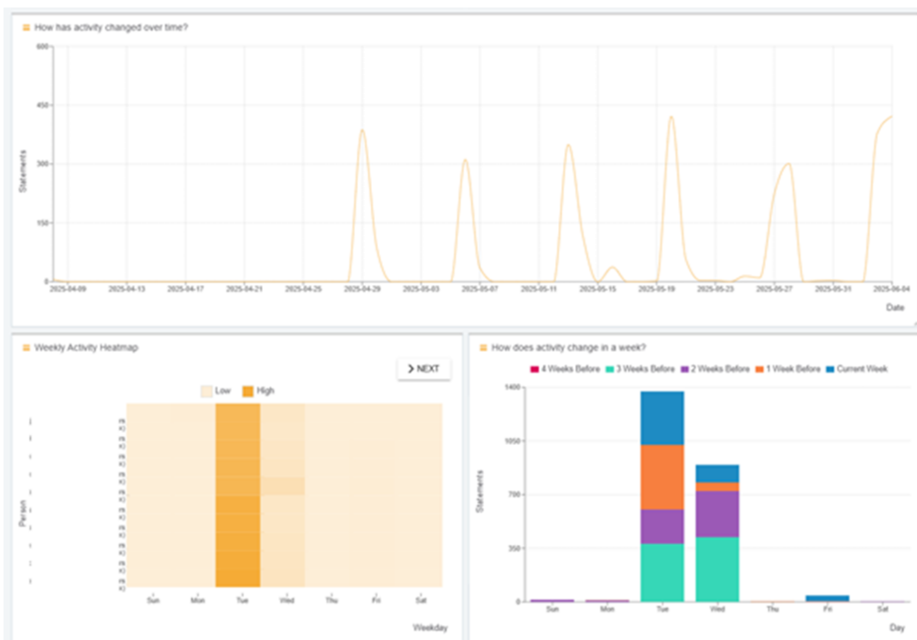


Fig. 4. Learning Locker dashboards for CS101 after Week 6.

Second Intervention (Week 13):
A follow-up round included reminders about final assessments and invitations to group consultations, organized on middle of July, 2025. Activity peaks were observed shortly afterward, showing a measurable increase in platform usage and assignment submissions. Several students explicitly stated in their replies that time management tips and clarity about remaining obligations motivated them to resume their studies.

Despite external disruptions, around 35% of previously inactive students re-engaged by Week 13, while high-performing students maintained consistent progress. However, approximately 20% of the cohort remained inactive throughout, often citing personal or work-related issues.

CS105 Fundamentals of Web Development

The CS105 course introduced students to front-end programming with HTML, CSS, and JavaScript. This pilot included four successive interventions, making it the most extensive implementation among the three.

First Intervention (Week 9):
Due to widespread inactivity, all students received an online questionnaire designed to identify barriers to engagement. Survey results revealed recurring challenges:

- Inability to attend classes due to student protests and transport blockades;
- Preference for online instruction and asynchronous completion of tasks;
- Personal or health-related obstacles;
- Lack of motivation due to non-mandatory attendance.

Second and Third Interventions (Weeks 14 and 15):
Based on feedback, instructors organized additional pre-exam sessions for students to catch up on missed work. Personalized emails invited low-performing students to these sessions.

Dashboard analytics in Fig. 5 indicated clear activity spikes in Weeks 14 and 15 with the number of submitted statements rising from 904 to 1200 and the average grade increasing from 2.3 to 3.0.

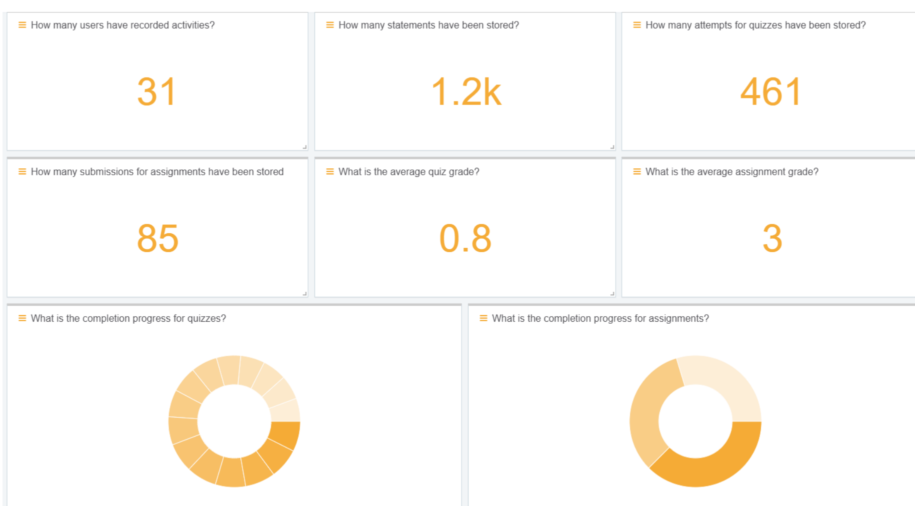


Fig. 5. Learning Locker dashboards for CS105 after Week 15.

Although some students remained inactive, most respondents expressed appreciation for the extra opportunities and flexibility. The iterative approach in CS105 showed that repeated, flexible interventions were most effective under crisis conditions.

CS230 Distributed Systems

The distributed systems course, which included topics on process synchronization, communication models, and blockchain security, served as an advanced-level pilot. Engagement analysis in Week 6 revealed that two-thirds of students had yet to begin any course activities.

First Intervention (Week 6):
Students were contacted individually according to the same five-profile model as in CS101. Particular attention was given to those who reported high anxiety or stress through the SRL survey. Some students expressed gratitude for the personalized support, noting that it reduced their sense of isolation.

Second Intervention (Week 13):
A similar follow-up was conducted with personalized feedback and suggestions for workload organization. Dashboard visualizations shown in Fig. 6 indicated modest improvements in engagement, especially among students who had previously responded positively to the first intervention. However, nearly half of the cohort remained unresponsive, highlighting the limitation of email-based outreach alone.

Fig. 6. Learning Locker dashboards for CS230 after Week 13.

Overall, the CS230 course underscored the importance of multimodal intervention strategies, including direct conversations and group consultations, rather than relying solely on asynchronous communication.

3.7 Cross-Course Summary

Across all three pilots, several common trends emerged:

- Early inactivity was widespread (40–60% in Weeks 1–6) but decreased after the first intervention.
- Response rates to personalized emails varied (10–30%), but those who replied often demonstrated consistent improvement.
- Flexible opportunities (extra sessions, extended deadlines) yielded the most significant engagement increases.
- SRL indicators (particularly self-reported anxiety) correlated strongly with performance fluctuations.

These findings collectively demonstrate that data-driven and empathetic intervention models can effectively re-engage students even in challenging academic environments. The pilots also validated the functionality of the Learning Locker dashboard as a tool for monitoring engagement and guiding instructors' decision-making

4 Results and Discussion

4.1 Overview of Results

The three pilot implementations at Belgrade Metropolitan University provided a broad dataset for examining how early interventions informed by learning analytics can affect student engagement and academic performance.

Across all courses, the data reveal consistent trends:

1. Student engagement was initially very low in the first half of the semester (Weeks 1–6 or 1–9).
2. Targeted interventions, whether through personalized emails, surveys, or additional sessions, corresponded to visible increases in LMS activity and higher assignment submission rates.
3. Students' self-reported SRL indicators (motivation, anxiety, and organization) were closely aligned with observed behavioral changes in Learning Locker dashboards.

The interventions were implemented at different points of the semester (Weeks 6, 9, 13, 14, and 15), but in all cases, dashboard visualizations displayed clear peaks in activity immediately following outreach.

In the CS101 course, approximately 35% of previously inactive students became active after the first intervention. In the CS105 course, the number of submitted learning statements increased from 904 to 1200, and the average assignment grade rose from 2.3 to 3.0 after two rounds of interventions.

In the CS230 course, engagement improved modestly, primarily among students who had responded to initial communication, while others remained inactive.

Table 3. Quantitative summary of key engagement metrics before and after interventions.

Course	Week of Initial Intervention	Average Submissions Before	Average Submissions After	Active Students Before	Active Students After	Average Grade Change
CS101	Week 6	45	82	43%	65%	+0.5
CS105	Week 9	904 (statements)	1200 (statements)	50%	70%	+0.7
CS230	Week 6	25	47	37%	53%	+0.4

The results summarized in Table 3 confirm the hypothesis that timely, data-informed interventions can reactivate disengaged students, even in semesters marked by significant external disruptions.

4.2 Impact on Student Engagement

Engagement was assessed through multiple dimensions, namely behavioral (activity logs), cognitive (task completion), and emotional (SRL self-reports).

Following the first intervention in each course, dashboards recorded increased frequency of logins, message exchanges, and assignment submissions. This trend aligns with prior findings that early communication and personal acknowledgment can trigger renewed motivation [7, 17].

However, engagement patterns differed by the type and timing of intervention.

In CS101 and CS230, where the first outreach occurred in Week 6, engagement improved temporarily but plateaued after several weeks, suggesting that a single intervention round is insufficient to maintain momentum.

By contrast, in the CS105 course, where three consecutive rounds were organized between Weeks 9 and 15, the engagement curve displayed sustained improvement, peaking during final pre-exam sessions.

The CS105 data therefore suggest that iterative interventions, even if similar in content, serve as recurring reminders that sustain engagement over time.

This finding supports [10], who observed that repeated prompts based on analytics encourage self-regulatory behaviors such as planning and time management.

4.3 Patterns in SRL Indicators and Student Feedback

The integration of SRL surveys provided deeper insight into the psychological and motivational factors underlying student behavior.

Across all pilots, high levels of reported anxiety were observed among both high- and low-performing students.

This aligns with findings from and [9, 12], who note that even high achievers may experience reduced self-efficacy when external disruptions or increased workload occur. In the BMU courses, SRL data informed the tone and focus of interventions:

- Students reporting high anxiety received messages emphasizing stress management and reassurance.
- Students with low self-regulation received suggestions for time planning and breaking tasks into smaller goals.
- Students lacking SRL data were reminded to complete the survey to enable tailored support.

Analysis of email correspondence and survey responses revealed several recurring themes, summarized in Table 4, and given as follows:

1. Appreciation of personalized contact: Many students expressed gratitude that instructors noticed their inactivity or stress, indicating that acknowledgment alone contributed to motivation.
2. Preference for flexibility: Students valued additional sessions and deadline extensions, especially when facing external challenges (e.g., protests, transport blockades).
3. Persistent barriers among some students: Despite support, roughly one-quarter of students remained inactive, suggesting deeper issues such as lack of intrinsic motivation or competing life obligations.

Table 4. Summary of thematic codes from student feedback.

Theme	Description	Example Student Comment
Recognition	Students valued being noticed and supported	"I didn't expect anyone to follow up on my work and it really helped me refocus."
Flexibility	Requests for online or additional sessions	"I could finally complete assignments thanks to the extra class."
Emotional support	Acknowledgment of anxiety and stress	"I feel less pressure knowing the teacher understands our situation."
Structural barriers	Transportation, protests, work obligations	"I simply couldn't attend because of the road blockades."

The inclusion of SRL-based personalization thus enhanced the human dimension of data-driven interventions, combining analytics precision with empathetic communication as an aspect emphasized in recent LA studies [17, 19].

4.4 Comparative Discussion Across Courses

While all three pilots shared common objectives and tools, they produced distinct insights depending on course structure and student maturity level. Regarding CS101 course, results shown that this was the most structured and technically demanding course. Early interventions had moderate impact, improving participation but not entirely eliminating inactivity. Advanced students responded best to individualized guidance. The most flexible and iterative implementation was conducted in CS105. The use of surveys and repeated sessions demonstrated the highest overall engagement increase. This supports the idea that multiple micro-interventions outperform single, large-scale actions. Finally, for CS230, despite clear data visualization and tailored communication, improvements remained modest. The findings for this course highlight the limitations of asynchronous email communication and the need for real-time support channels such as live consultations or mentoring.

These differences suggest that while the Learning Locker infrastructure effectively centralizes analytics, the pedagogical responsiveness of instructors ultimately determines impact. The best results occurred when interventions combined data insights with context-sensitive teaching actions, such as extending deadlines, organizing catch-up labs, and maintaining regular communication loops.

4.5 Interpretation in Light of Literature

The results from the BMU pilots support existing research emphasizing the value of continuously monitoring students' engagement through learning analytics dashboards, complemented by SRL survey data, in order to design timely and personalized interventions. Consistent with [7], early detection of disengagement led to measurable improvements in participation and grades. The integration of SRL surveys parallels findings in [11, 17], who argue that feedback rooted in metacognitive awareness enhances learners' capacity for self-regulation.

However, the pilots also reveal key practical limitations that echo concerns from [2, 18]:

– Data analytics alone cannot ensure engagement; they must be paired with pedagogical empathy.
– Sustained impact requires iterative, multi-channel communication.
– External contextual factors (sociopolitical disruptions, mental health issues) significantly shape students' responsiveness to interventions.

Moreover, these findings contribute to the emerging body of European research advocating for institution-wide frameworks that connect learning analytics with ethical use, teacher training, and early-warning systems. The BMU implementation shows that such systems can be scalable and adaptable across courses, provided that instructors receive proper support and interpret analytics as a diagnostic tool.

5 Conclusion

This study examined how learning analytics and early interventions can be applied to improve student engagement and learning outcomes in higher education. Within the framework of the ISILA project, three pilot courses at BMU served as testing grounds for the integration of Learning Locker dashboards, SRL surveys, and targeted communication strategies.

Across all pilots, the results confirmed the central premise that data-driven insights combined with timely pedagogical action can positively affect engagement, motivation, and academic performance. The use of learning analytics enabled early identification of disengaged or anxious students, while SRL data provided the necessary context to tailor interventions to individual needs. This dual approach demonstrated that learning analytics are most effective when interpreted through a human-centered, empathetic lens rather than as a purely technical monitoring tool. Several key findings emerged from the analysis. Firstly, early intervention matters significantly increased engagement after the first outreach in each course, confirming that students respond to acknowledgement and structured support. In addition, frequency and flexibility of interventions proved decisive. The course that implemented multiple interventions recorded the largest and most sustained improvement, showing that repeated prompts and opportunities for participation maintain engagement more effectively than single interventions. Third, the pilots revealed the importance of addressing affective factors, such as anxiety and stress, which directly influence academic persistence. Personalized messages and flexible deadlines mitigated these issues and encouraged students to take proactive control of their learning.

At the same time, several limitations must be acknowledged. A consistent group of students (roughly one quarter) remained inactive despite multiple interventions, suggesting that external or personal factors, such as employment, mental fatigue, or systemic disengagement, require alternative strategies beyond email communication. Furthermore, the socio-political context of the semester, characterized by nationwide protests and disrupted class schedules, limited the generalizability of the findings. These external stressors underscore how contextual factors can amplify disengagement, and why institutional resilience and adaptability are essential components of successful learning analytics initiatives.

Building on these findings, future work at BMU and within the ISILA consortium will focus on automation and predictive analytics by expanding the Learning Locker system with predictive models that can automatically flag at-risk students and trigger early notifications for instructors. Secondly, we will incorporate multimodal communication channels such as in-app notifications, learning platform chatbots, or short video messages to complement traditional email communication. Regarding expanding SRL, we plan to introduce longitudinal SRL instruments to measure how students' self-regulation evolves over the semester and how interventions affect metacognitive growth.

Acknowledgment

This paper was supported in part by the European Commission through the ISILA (2023-1-FI01-KA220-HED-000159757) project, and in part by the Blockchain Technology Laboratory at Belgrade Metropolitan University.

References

1. Long, P., Siemens, G.: Penetrating the fog: Analytics in learning and education. EDUCAUSE Review (Online) (2011)
2. Ferguson, R.: The state of learning analytics in 2012: A review and future challenges. Knowledge Media Institute, Technical Report KMI-2012-01 (2012)
3. Lang, C., Siemens, G., Wise, A., Gasevic, D.: Handbook of learning analytics (2017)
4. Siemens, G., Gasevic, D.: Guest editorial-learning and knowledge analytics. Journal of Educational Technology & Society **15**(3), 1–2 (2012)
5. Akçapınar, G., Altun, A., Aşkar, P.: Using learning analytics to develop early-warning system for at-risk students. International Journal of Educational Technology in Higher Education **16**(1), 1–20 (2019)
6. Kondo, N., Okubo, M., Hatanaka, T.: Early detection of at-risk students using machine learning based on lms log data. In: 2017 6th IIAI international congress on advanced applied informatics (IIAI-AAI). pp. 198–201. IEEE (2017)
7. Arnold, K.E., Pistilli, M.D.: Course signals at purdue: Using learning analytics to increase student success. In: Proceedings of the 2nd international conference on learning analytics and knowledge. pp. 267–270 (2012)
8. Macfadyen, L.P., Dawson, S.: Mining lms data to develop an "early warning system" for educators: A proof of concept. Computers & education **54**(2), 588–599 (2010)
9. Ifenthaler, D., Mah, D.K., Yau, J.Y.K.: Utilising learning analytics for study success: Reflections on current empirical findings. In: Utilizing learning analytics to support study success, pp. 27–36. Springer (2019)
10. Tempelaar, D.T., Niculescu, A., Rienties, B., Gijselaers, W.H., Giesbers, B.: How achievement emotions impact students' decisions for online learning, and what precedes those emotions. The Internet and Higher Education **15**(3), 161–169 (2012)
11. Zimmerman, B.J.: Attaining self-regulation: A social cognitive perspective. In: Handbook of self-regulation, pp. 13–39. Elsevier (2000)
12. Pintrich, P.R.: A conceptual framework for assessing motivation and self-regulated learning in college students. Educational psychology review **16**(4), 385–407 (2004)
13. Roll, I., Winne, P.H.: Understanding, evaluating, and supporting self-regulated learning using learning analytics. Journal of Learning Analytics **2**(1), 7–12 (2015)
14. Viberg, O., Khalil, M., Baars, M.: Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In: Proceedings of the tenth international conference on learning analytics & knowledge. pp. 524–533 (2020)
15. Heikkinen, S., Saqr, M., Malmberg, J., Tedre, M.: Supporting self-regulated learning with learning analytics interventions–a systematic literature review. Education and Information Technologies **28**(3), 3059–3088 (2023)
16. Matcha, W., Uzir, N.A., Gašević, D., Pardo, A.: A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. IEEE transactions on learning technologies **13**(2), 226–245 (2019)

17. Jivet, I., Scheffel, M., Drachsler, H., Specht, M.: Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In: European conference on technology enhanced learning. pp. 82–96. Springer (2017)
18. Verbert, K., Govaerts, S., Duval, E., Santos, J.L., Van Assche, F., Parra, G., Klerkx, J.: Learning dashboards: an overview and future research opportunities. Personal and Ubiquitous Computing **18**(6), 1499–1514 (2014)
19. Tsai, Y.S., Gasevic, D., Whitelock-Wainwright, A., Muñoz-Merino, P.J., Moreno-Marcos, P.M., Fernández, A.R., Kloos, C.D., Scheffel, M., Jivet, I., Drachsler, H., et al.: Sheila: Supporting higher education to intergrate learning analytics research report (2018)
20. Greller, W., Drachsler, H.: Translating learning into numbers: A generic framework for learning analytics. Journal of Educational Technology & Society **15**(3), 42–57 (2012)
21. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative research in psychology **3**(2), 77–101 (2006)

Sensitivity Analysis and Temporal Stability of Student Success Predictors based on Different Data Sources in Education

Marija Pokos Lukinec[1][0009-0006-1564-5970], Dijana Oreški[2] [0000-0002-3820-0126], and Dino Vlahek[3][0000-0002-3911-8685]

[1, 2, 3] University of Zagreb, Faculty of Organization and Informatics, Pavlinska 2, 42000, Varaždin, Croatia
mapokos@foi.hr, dijana.oreski@foi.hr, dvlahek@foi.hr

Abstract. Student success prediction is a central topic in educational data mining and learning analytics, as institutions increasingly rely on data-driven approaches to enhance learning outcomes. However, the dynamic nature of educational environments raises questions about the long-term reliability of predictive features used in these models**.** This study aims to investigate the temporal stability of features extracted by sensitivity analysis of predictive models developed by integrating data from various sources, including the e-learning system, student attendance records, teacher opinions, and meteorological data. In this study, the stability of success predictors is modeled using machine learning algorithms – Random Forest and Gradient Boosted Decision Tree. By applying regression metrics, the precision of the model is assessed to determine the reliability of predictive features over time. Identification of the relevant success predictors and their temporal stability provides insights into significant success predictors in the long term. The results support the development of robust predictive models and highlight key features that contribute to the reliable analysis of student success outcomes.

Keywords: Predictive Data Modeling, Machine Learning, Stability of Success Predictors, Learning Management System, Learning Analytics, Random Forest, Gradient Boosted Tree.

1 Introduction

Machine learning algorithms and the vast amounts of data generated daily have led to data-driven predictions aimed at advancement and development across various domains, including education. To provide high-quality feedback in the shortest possible time, new models and methods for tracking student progress are being developed. Data collected and analyzed from the Learning Management System (LMS) helps predict which features influence student dropout rates and which contribute to motivation, success, and continued engagement. The analysis of data from the e-learning system is defined as learning analytics. According to [1]*, learning analytics is the collection, analysis, interpretation and communication of data about learners and their learning that provides theoretically relevant and actionable insights to enhance learning and teaching*. Researchers aim to extract high-quality

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.058PL
Part of ISBN: 978-86-89755-37-4

data, specifically e-learning system features that would serve as a solid foundation for developing accurate and reliable student success models. Study [2] states that only LMS data is insufficient for prediction, highlighting the need for integration, making it one of the research objectives.

This study will analyze data extracted from the e-learning system, integrated with multiple sources, to determine the features that influence the student success model and which features demonstrate stability in prediction over time, thereby ensuring model robustness.

Objectives of the study are: (i) to integrate data from various sources (e-learning system, nastava.foi.hr, teachers' opinion, meteorological data), (ii) to determine the temporal stability of features over time, (iii) to compare the stability of predictive models obtained through different machine learning algorithms, (iv) to identify stable and relevant features from the formative assessment group.

Based on the established objectives, research questions are formulated: (i) Which group of integrated features contributes more to model stability?, (ii) Which success predictors can be extracted as stable within courses in the e-learning system for several years?, (iii) Which stable data features have greater predictive power in predictive models obtained using different machine learning algorithms, based on sensitivity analysis?, (iv) Which features from the formative assessment group have been identified as stable?

2 Course and Data Description

This section describes course included in the research as well as data used for predictive models development.

2.1 Course Description

The research is conducted within a blended course model. The course is delivered at the Faculty of Organization and Informatics, University of Zagreb, at the undergraduate level. It is a mandatory course that enrolls over 200 students annually. The course lasts 15 weeks and is organized so that the resources needed for lectures, exercises, and seminars are published in the e-learning system.

Materials are published in the system during both synchronous and asynchronous classes. Materials used during synchronous classes remain accessible for students to review and study later, while materials published during asynchronous classes additionally serve as preparation tools and support work in a live environment. Since this is a blended course model, in addition to data on student activity within the system and the number of points accumulated through various assignments during the course, data on student attendance records and teacher assessments are also available for analysis. This teaching format enables teachers to closely monitor student progress.

Assignments are given to students either synchronously or asynchronously, and as previously mentioned, all materials are published within the e-learning system. Assessment is conducted through formative (assignments) and summative (midterms, projects) tasks. Students must achieve at least 50% in each summative component to pass the course. The final grade is determined by summing the points earned in summative assessments. Assignments are submitted by students via the e-learning system, reviewed by teachers, and discussed through verbal feedback during live sessions. This process provides students with insights into their performance and areas for improvement.

For the 2023/2024 course, 241 students were enrolled, while in the 2024/2025 course, 246 students were enrolled.

The next subsection presents the data extracted from the e-learning system for a regression task, specifically predicting the number of points obtained in midterms.

2.2 Data Description

The data includes integrated information from: (i) the Moodle e-learning system; (ii) student attendance records; (iii) subjective teachers' opinions regarding students' performance during classes, and (iv) meteorological data retrieved from the Open-Meteo website [3].

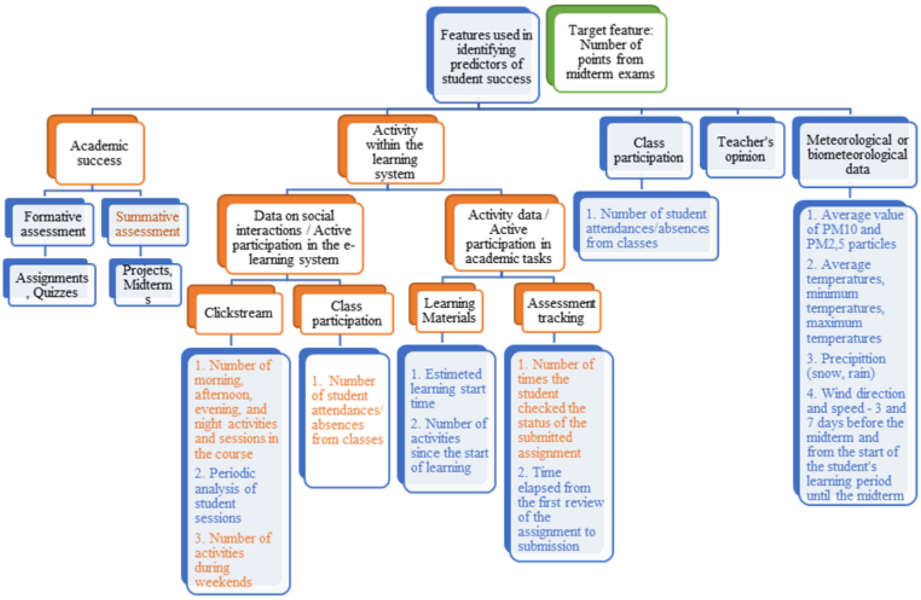


Fig. 1. Feature groups and examples of features used in the research

The first set of 46 features relates to LMS data. Since [4] states that the time spent within a course is not significant for the outcome unless dedicated to learning, the research analyzes those features that can track time spent on learning with the integration of new data. Some of these features, highlighted in orange, were also used in studies [2], [5], and [6].

The second group of features, meteorological features, includes data on average, minimum, and maximum temperature, precipitation, wind speed, and PM particles. PM10, and PM2.5 - airborne pollutant particles capable of penetrating the skin, bronchi, and bloodstream, thus posing a risk to human health [7]. Given their impact on the environment, climate, and visibility [7], it is hypothesized that they also influence human mood to some extent, potentially affecting motivation and academic success. The inclusion of meteorological data further contributes to this research, as there are no existing studies that have integrated such data into their analyses.

Meteorological data were incorporated to explore whether environmental conditions—such as temperature or daylight duration—affect student engagement and overall success. External factors of this kind can influence students both directly and indirectly. For example, severe weather conditions may impact attendance, limit students' mobility, or affect their concentration and motivation. Similarly, prolonged periods of poor weather could reduce opportunities for social interaction and extracurricular learning, while favorable conditions may foster more consistent participation. By including meteorological data, the study investigates whether academic performance is sensitive to non-academic variables, offering a broader view of factors shaping student outcomes.

Attendance record features are entered by teachers and can be retrieved from the nastava.foi.hr system or the e-learning system. The features provide data on the number of times a student has attended classes (lectures and/or seminars, and/or exercises).

The teacher's opinion feature represents the teacher's subjective assessment, based on the student's performance during classes or exercises, as well as their engagement in all aspects of academic work.

Teacher assessments of student engagement and performance, while valuable, introduce an inherently subjective dimension into the predictive model. Such judgments may capture behavioral student patterns that are not reflected in quantitative data from LMS logs or environmental records. However, reliance on teacher opinions must be interpreted cautiously, as individual perceptions can vary widely and may be influenced by implicit biases or limited classroom interactions. Including this predictor provides an opportunity to compare human-informed insights with objective metrics, helping to assess its relative stability and predictive power.

Integration of these data sources serves as the basis for development of predictive models for student success prediction.

3 Research Methods

The research follows the CRISP-DM standard for data analysis. This standard is applied through six phases: domain understanding, data understanding, data preparation, modelling, evaluation, and deployment [8]. In conducting the research, data is first collected from the e-learning system and other sources, then integrated to ensure a comprehensive analysis.

After collecting and explaining the data for understanding, the data is extracted, structured, and prepared for modelling using a machine learning algorithm.

Sensitivity analysis identifies stable features for making predictions. In the evaluation phase, models are assessed, and research questions are addressed.

Data sets are collected at the Faculty of Organization and Informatics, University of Zagreb, and approval from the Ethics Committee is obtained for conducting the research and using the data, with the necessary anonymization of personal indicators and proper storage. The data is divided into a training set and a test set, and multiple machine learning algorithms are applied to develop predictive models of student success. The modeling phase is central to the research, where the application of machine learning algorithms – Random Forest (RF) and Gradient Boosted Trees (GBT) – determines the role and significance of stable features, as well as their contribution to predicting model performance.

RF is an ensemble of decision trees where each tree relies on randomly selected features from the input data. This approach reduces dependency among data points and ensures robustness by minimizing sensitivity to noise in the data, such as missing values and outliers [5], [9].

GBT is a machine learning algorithm that combines multiple weak predictive models (decision trees) to create a strong predictive model [10]. In GBT, trees are built iteratively, with each new tree attempting to correct and minimize the errors of the previous one [10].

Many consulted studies apply these algorithms to similar types of tasks, which is why they are applied in this research. Study [11] states that the GBT algorithm provides the best prediction for student success and failure, while [5] reports similar findings for the RF algorithm. Decision tree-based algorithms have also been applied in studies [12] and [13].

The results obtained using the algorithms are evaluated using metrics commonly used in regression tasks, namely Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Normalized Root Mean Squared Error (NRMSE). All metrics measure prediction errors, where lower values indicate better model performance. Based on the sensitivity analysis, insights into key features for prediction are obtained, and the resulting model will then be applied to predict new cases.

4 Results and Discussion

This section presents the results of the pilot study based on an analysis of data from the academic years 2023/2024 and 2024/2025.

Table 1. Presentation of results obtained using RF and GBT algorithms

Algorithm	Metric	2023/2024	2024/2025
Random Forest	MAE	2.3474	2.38
	RMSE	2.8221	2.8303
	NRMSE	0.2352	0.2881
Gradient Boosted Tree	MAE	2.2011	2.3093
	RMSE	2.7050	2.9587
	NRMSE	0.2254	0.3012

On the prepared dataset, after modelling with the mentioned algorithms, an evaluation is conducted and presented in Table 1.

Table 2. Presentation of the top 10 features obtained through sensitivity analysis, sorted by importance

RF - 2023/2024	RF - 2024/2025	GBT - 2023/2024.	GBT - 2024/2025
topinion: 0.3839	**topinion:** 0.3457	**topinion:** 0.4530	**topinion:** 0.3382
Exercise: 0.0961	**Assignment: Submission of the task for Week 3 exercises (Exercise 2 and 3) time:** 0.0480	**Excercise:** 0.1740	**Assignment: Submission of the task for Week 3 exercises (Exercise 2 and 3) time:** 0.0433
Assignment: Submission of the task for lab exercises - ERA model time: 0.0481	**activity_sum*activity_start_learning_before_exam_sum:** 0.0303	**Research seminar:** 0.0641	**Assignment: Submission of the task for lab exercises - Relational schema time:** 0.0402
Research seminar: 0.0340	**AssignmentMean:** 0.0297	**weekend_10:** 0.0300	**activity_start_learning_before_exam_Afternoon**: 0.0399
start_learning_before_exam: 0.0287	**activities_during_learning:** 0.0287	**start_learning_before_exam:** 0.0271	**AssignmentMean**: 0.0388
weekend_10: 0.0260	**Assignment: Submission of the task for lab exercises - Relational schema time:** 0.0260	**Assignment: Submission of the task for lab exercises - ERA model time:** 0.0269	**activity_start_learning_before_exam_sum:** 0.0370
Assignment: Submission of the task for lab exercises - Relational schema time: 0.0248	**Assignment: Submission of the task for lab exercises 4 - Project management time:** 0.0185	**median_periodicity_h:** 0.0250	**mode_periodicity_h:** 0.0358
activity_start_learning_before_exam_Afternoon: 0.0212	**Assignment: Oracle Apex submission 1 time:** 0.0184	**activity_Afternoon:** 0.0189	**Assignment: Submission of the task for lab exercises 4 - Project management time:** 0.0242
Assignment: Submission of the task for lab exercises 5 - BPMN time: 0.0174	**Assignment: Submission of the task for lab exercises 6 - Use case diagrams time:** 0.0179	**Assignment: Submission of the task for lab exercises - Relational schema time:** 0.0166	**activity_sum*activity_start_learning_before_exam_sum:** 0.0227
Assignment: Submission of the task for Week 3 exercises (Exercise 2 and 3) time: 0.0171	**start_learning_before_exam:** 0.0178	**active ty_start_learning_before_exam_Afternoon:** 0.0142	**Assignment: Submission of the task for lab exer-cises - ERA model time:** 0.0223

Table 2 presents the top 10 features identified as the most important during the sensitivity analysis, classified by years and algorithms.
So, one measure of stability, mentioned in study [9], is sensitivity analysis, which tracks features that consistently appear over time. Another measure is the coefficient of variation. Study [16] states that the coefficient of variation is a measure of dispersion used to detect changes in accuracy and monitor process stability. In this case, it is applied to assess feature stability over two years.

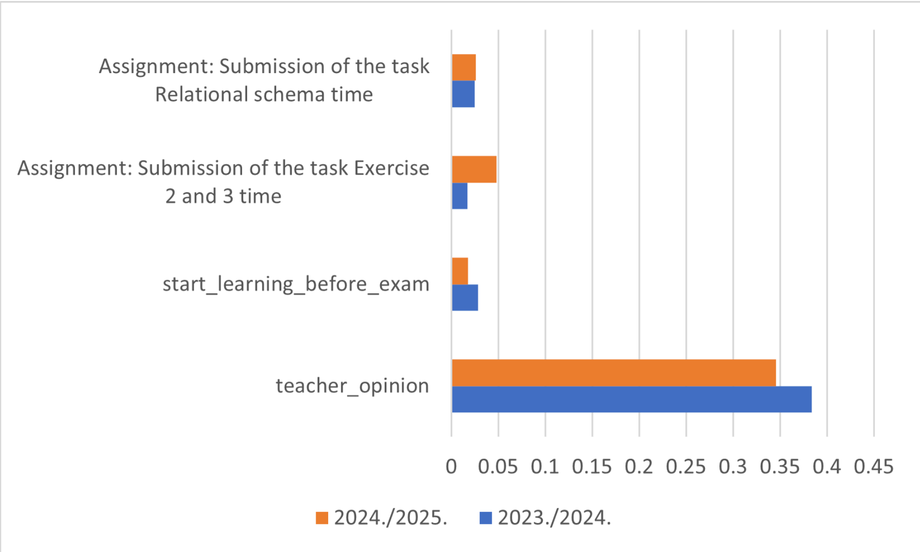


Fig. 2. Feature stability trends across years obtained using the RF algorithm

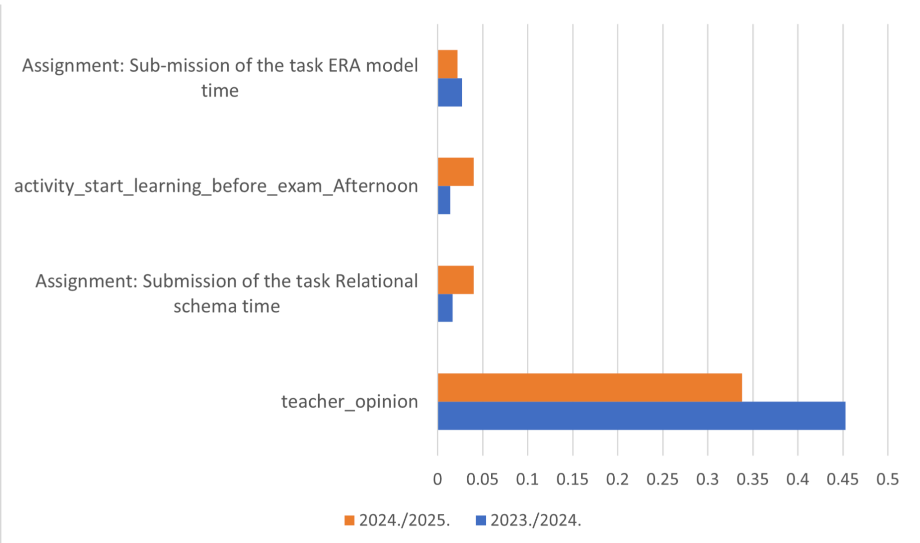


Fig. 3. Feature stability trends across years obtained using the GBT algorithm

The Figures 2 and 3 illustrate feature stability trends across years for four consistent features. The first image shows results obtained using the Random Forest (RF) algorithm, while the second reflects outcomes from Gradient Boosted Trees (GBT). The comparison highlights how feature importance and consistency vary between models, offering insights into their robustness and temporal reliability.

Table 3. Presentation of the coefficient of variation for features recurring in 2023/2024 and 2024/2025

Features in 2023/2024 and 2024/2025 - RF	Coefficient of variation	Features in 2023/2024 and 2024/2025 - GBT	Coefficient of variation
teacher_opinion	0.0524	teacher_opinion	0.1451
start_learning_before_exam	0.2344	Assignment: Submission of the task for lab exer-cises - Relational schema time	0.4155
Assignment: Submission of the task for Week 3 exercises (Exercise 2 and 3) time	0.4747	activity_start_learning_before_exam_Afternoon	0.4750
Assignment: Submission of the task for lab exercises - Relational schema time	0.0236	Assignment: Sub-mission of the task for lab exercises - ERA model time	0.0935

Table 3 presents the results of the pilot study for the coefficient of variation measure. The results show features that appear in both years for the RF and GBT algorithms, considering only the top 10 features identified through sensitivity analysis. Findings indicate that certain course assignments play a significant role in prediction, followed by the subjective teacher's opinion.

5 Conclusion

This study provides valuable insights into the temporal stability of student success predictors gained from multiple data sources. Reserach results highlight the importance of identifying robust features to ensure the long-term reliability of predictive models in educational settings.

LMS features: *Assignment: Submission of the task for lab exercises - Relational schema time* (using the RF algorithm) and - *Assignment: Submission of the task for lab exercises - ERA model time* (using the GBT algorithm), which belong to the formative assessment type, are the most stable predictors of success, demonstrating the highest predictive power. This conclusion is derived from answering the research question posed:

The first question examines which group of integrated features contributes most significantly to model stability. The pilot study found that features related to teacher opinions and academic success have a greater impact on model stability.

The second question aims to identify which success predictors remain stable within an e-learning course over several years, measured by the coefficient of variation. The results show that within the e-learning system, assignments, learning start time, and activities in the system during learning can be extracted as stable predictors over time.

The third question aims to identify stable data features with strong predictive power in models generated using various machine learning algorithms, based on sensitivity analysis. Some of the features that can be identified include teacher opinions, assignments, activities during learning, and weekend activity patterns.

Additionally, the fourth research question examines features from the formative assessment group that have been identified as stable. These features belong to the academic success category. Among the extracted LMS features, assignments stand out, as sensitivity analysis indicates, they have higher predictive power. Furthermore, the coefficient of variation shows that some assignments have remained stable over the years.

It can be concluded that the study's findings on predictor significance align with the consulted research, with further contributions focusing on verifying the stability of features. i.e., success predictors.

The study's limitations are reflected in the small dataset. However, future research will expand to include data covering more years. Additionally, in the continuation of the study, more machine learning algorithms will be applied, and evaluations will be conducted across multiple metrics.

The future goal is also to collect data over multiple years within a single course before expanding the research to multiple courses.

Acknowledgement

This paper is supported by Croatian Science Foundation under the project SIMON: Intelligent system for automatic selection and machine learning algorithms in social sciences, UIP-2020-02-6312.

Declaration on Generative AI

During the preparation of this work, the author(s) used Copilot and Grammarly in order to: Edit references and grammar and spelling check. After using tool(s)/service(s), reviewed and edited the content as needed, and take(s) full responsibility for the publication's content.

References

1. Society for Learning Analytics (SoLAR): What is Learning Analytics? Available: https://www.solaresearch.org/about/what-is-learning-analytics/, last accessed 2025/07/01.
2. Conijn, R., Snijders, C., Kleingeld, A., Matzat, U.: Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS. In: IEEE Transactions on Learning Technologies, vol. 10, no. 1, pp. 17–29 (2017).
3. Open-meteo: Free Weather API. Available: https://open-meteo.com/en/docs, last accessed 2025/03/21.
4. Saqr, M., Fors, U., Tedre, M.: How Learning Analytics Can Early Predict Under-Achieving Students in a Blended Medical Education Course. Medical Teacher 39(7), 757–767 (2017).
5. Saqr, M., López Pernas, S.: Why Explainable AI May Not Be Enough: Predictions and Mispredictions in Decision Making in Education. Smart Learning Environments 11(52), 1–16 (2024). Available: https://doi.org/10.1186/s40561-024-00343-4.
6. Abdulkareem Shafiq, D., Marjani, M., Ariyaluran Habeeb, R.A., Asirvatham, D.: Digital Footprints of Academic Success: An Empirical Analysis of Moodle Logs and Traditional Factors for Student Performance. Education Sciences 15(3), 304 (2025). Available: https://doi.org/10.3390/educsci15030304.
7. Klonkay, P.: Procjena indeksa kvalitete zraka za lebdeće čestice PM10 i PM2,5 grada Zagreba primjenom metoda strojnog učenja (Diplomski rad). Fakultet kemijskog inženjerstva i tehnologije (2024).
8. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.: CRISP-DM 1.0 - Step-by-Step Data Mining Guide (2000). Available: https://mineracaodedados.files.wordpress.com/2012/12/crisp-dm-1-0.pdf.
9. Breiman, L.: Random Forests. Machine Learning 45, 5–32 (2001).
10. Shah, J.: Gradient Boosting. Technical Report, no. 1–6, Computer Science Department, Nirma University, Ahmedabad, India (2020).
11. Blessed Deho, O., Joksimovic, S., et al.: Should Learning Analytics Models Include Sensitive Attributes? Explaining the Why. IEEE Transactions on Learning Technologies 16(4), 560–571 (2023).
12. Balabaeva, K., Kovalchuk, S.: Comparison of Temporal and Non-Temporal Features Effect on Machine Learning Models Quality and Interpretability for Chronic Heart Failure Patients. Procedia Computer Science 156, 87–96 (2019).
13. Teinemaa, I., Dumas, M., et al.: Temporal Stability in Predictive Process Monitoring. Data Mining and Knowledge Discovery 32, 1306–1338 (2018). Available: https://doi.org/10.1007/s10618-018-0575-9.
14. Chai, T., Draxler, R.R.: Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Arguments Against Avoiding RMSE in the Literature. Geoscientific Model Development 7, 1247–1250 (2014).
15. Lifesight: What is Normalized Root Mean Square Error (NRMSE)?. Available: https://lifesight.io/glossary/normalized-root-mean-square-error/, last accessed 2025/04/07.
16. Jalilibal, Z., Amiri, A., et al.: Monitoring the Coefficient of Variation: A Literature Review. Computers & Industrial Engineering 161, 107600 (2021).
17. Van Rossum, G., Drake, F.L.: Python 3 Reference Manual. Scotts Valley, CA: CreateSpace (2009).

Photogrammetry in 3D Game Development Education: A Case Study of Student Learning

Petar Pejic[1][0000-0003-4155-8038], Milos Nikolic [0009-0009-9023-9050], and Isidora Mitrovic[2][0009-0007-5734-2289]

[1] Faculty of Information Technology, University Metropolitan Belgrade, Serbia
[2] Faculty of Civil Engineering and Architecture, University of Nis, Serbia
petar.pejic@metropolitan.ac.rs

Abstract. This paper presents a case study on the integration of photogrammetry into undergraduate game development education. Conducted within the course *AD185 – Izrada 3D video igara* at Metropolitan University, the module engaged six students in capturing real-world objects and transforming them into optimized, textured 3D models for use in Unity Game Engine. A mixed-methods approach was employed, combining pre- and post-assignment questionnaires with technical evaluation of the resulting models. The results indicate significant improvements in student understanding, practical skill acquisition, and engagement. The assignment also encouraged students to critically evaluate different workflows, balancing automation and manual control. Despite a small cohort, the findings suggest photogrammetry is an effective, scalable addition to 3D modeling curricula. The paper concludes with recommendations for implementation and outlines future research directions in photogrammetry-based education.

Keywords: Photogrammetry, 3D Modeling, Game Development Education.

1 Introduction

Photogrammetry, the process of reconstructing three-dimensional (3D) digital models from photographs, has become an essential technique in the contemporary game development pipeline. By enabling the transformation of real-world objects and environments into textured 3D assets, photogrammetry offers a fast, scalable, and increasingly accurate alternative to manual modeling techniques. As gaming audiences continue to demand photorealistic content, studios have embraced photogrammetry for asset generation, especially in creating environments, props, and surface textures with fine detail and realism [1, 2].

Beyond industry applications, photogrammetry is gaining attention in higher education as a powerful pedagogical tool. In domains ranging from archaeology and cultural heritage to architecture and design, educators are adopting photogrammetric workflows to provide experiential learning opportunities that bridge the digital and physical worlds [3, 4]. These approaches support active learning through direct engagement with spatial data, while also fostering technical and creative skills [5]. The ability to scan real

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.068P
Part of ISBN: 978-86-89755-37-4

objects and process them into game-ready assets offers students immediate feedback and motivation, making photogrammetry a valuable addition to curricula focused on 3D modeling and game development.

Despite the growing relevance of photogrammetry in the digital content creation industry, its integration into game development education remains relatively underexplored in scholarly literature. While there are successful examples of its use in cultural preservation [6], landscape modeling [7], and historical reconstruction [8], few studies have examined how photogrammetry can enhance learning outcomes in game design programs specifically. Moreover, little is known about how students experience the technical and creative challenges of capturing, processing, and integrating photogrammetric models into real-time real-time game engines such as Unity or Unreal.

In this paper, we present a case study from the undergraduate course AD185 “Izrada 3D video igara” (3D Game Development) at Metropolitan University, where a dedicated photogrammetry module was introduced to teach students how to create their own 3D assets from real-world objects. Six students completed the module, including a practical assignment and pre/post questionnaires to evaluate expectations, experiences, and learning outcomes. The goal of this study is twofold: (1) to assess the pedagogical value of photogrammetry in a game development context, and (2) to document student-created models, tools used, and the overall effectiveness of this approach in developing practical skills relevant to digital game production.

2 Related Work

Photogrammetry has undergone significant development in recent decades, emerging as a foundational technique in both industrial and academic contexts. In digital content creation, particularly within the video game industry, photogrammetry allows the efficient and accurate reconstruction of real-world objects and environments for use as digital assets. It is valued for enabling rapid generation of realistic geometry and texture data, thereby shortening development cycles while increasing fidelity [1, 2].

The use of photogrammetry in education has also gained considerable attention. Chapinal-Heras et al. [3] demonstrated the pedagogical potential of photogrammetry in a history curriculum, where students used image-based 3D modeling techniques to digitize cultural artifacts. This hands-on approach helped students better engage with course material while developing transferable technical skills. Similarly, Sapirstein [4] reported on the value of photogrammetry in archaeological education, emphasizing how it enhances students’ spatial reasoning and data literacy through direct interaction with field data.

These findings align with broader efforts to integrate 3D modeling into immersive educational environments. Hughes et al. [5] compared photogrammetry and laser scanning in the context of real-world capture for virtual reality applications, concluding that photogrammetry provided a cost-effective and accessible solution for creating immersive learning assets. Their findings suggest that students benefit not only from acquiring technical competencies but also from applying those skills in multidisciplinary and experiential settings.

In the context of game development, photogrammetry offers an effective bridge between artistic and technical disciplines. Pejić et al. [6], in their study of a historical reconstruction project, compared manual and automatic photogrammetric approaches to model the Barutana building in Serbia. The semi-automatic approach yielded a simplified but manageable model suitable for visualization and presentation, whereas the fully automatic method produced a highly detailed and geometrically accurate model, albeit with a substantially larger file size. Their findings highlighted trade-offs between manual effort and computational complexity, and their conclusions remain relevant to game development workflows, where asset performance and visual fidelity must be balanced.

Further research by Vannini et al. [7] explored photogrammetry's role in landscape modeling, emphasizing its scalability and adaptability to various levels of detail. This is particularly relevant to open-world game environments, where accurate terrain and environmental models are required. Similarly, Koutsoudis et al. [8] conducted a performance evaluation of multi-image 3D reconstruction, concluding that photogrammetry techniques, if properly applied, can deliver high-quality models comparable to those obtained with more expensive methods.

Finally, Guidi et al. [9] and Dall'Asta and Roncella [10] contributed to the methodological literature by analyzing photogrammetric pipeline components, such as image alignment, dense matching, and mesh reconstruction. Their studies underscore the importance of algorithmic selection and parameter tuning, especially when photogrammetry is taught as part of an applied curriculum.

Prior literature illustrates photogrammetry's strong potential to support learning through practice, foster technical skill development, and promote interdisciplinary thinking. However, there remains a gap in empirical studies focusing specifically on photogrammetry in game development education. This paper aims to contribute to this emerging area by evaluating how photogrammetry was introduced in a game design course, what students created, and how their learning evolved through the process.

3 Methodology

3.1 Course Context and Objectives

This study was conducted within the undergraduate course AD185 "Izrada 3D video igara" (3D Game Development) at Metropolitan University. The course is part of the curriculum for students specializing in game design and interactive media and aims to equip them with practical skills in 3D modeling, animation, and game engine integration. The photogrammetry module was introduced mid-semester as an experiential learning intervention, designed to bridge real-world object acquisition with digital asset creation workflows used in the video game industry.

The learning objectives of the photogrammetry module were threefold:

1) introduce students to image-based 3D reconstruction methods and tools,
2) provide practical experience in converting real objects into game assets, and

3) enhance student understanding of the end-to-end digital content creation pipeline, including capture, modeling, optimization, and game engine integration.

3.2 Assignment Design and Workflow

Six students participated in the module. Each was assigned the task of independently selecting a real-world object, capturing a series of photographs from multiple angles, and processing the image set using photogrammetry software. Students were given a choice of both mobile and desktop-based applications, including:

- RealityScan (Epic Games) and Polycam – mobile applications offering guided photo acquisition and automated cloud-based processing,
- Meshroom (AliceVision) – an open-source desktop tool requiring an NVIDIA GPU or the OpenCL variant (MeshroomCL) for AMD systems,
- RealityCapture – a high-end desktop application available via student license.

Students were instructed to photograph their objects in consistent lighting conditions with sufficient overlap between images. The resulting image sets ranged from approximately 20 to 100 photos per student, depending on object size and complexity. Processing was performed either on students' personal devices or remotely via cloud services, yielding textured 3D mesh outputs.

Following reconstruction, students were required to:

- Import the model into Autodesk Maya for cleanup, including removal of background elements, and minor corrections,
- Export the optimized model in a game engine-compatible format,
- Test the model's integration into the Unity engine, including material assignment and scene placement.

Students also documented their workflows and outcomes in a structured report, including technical metadata such as:

- number of images used,
- time required for photography and processing,
- final polygon count,
- file size of the exported model,
- quality and clarity of generated textures,
- and whether further processing in Maya was required.

This documentation was used as the basis for comparative technical analysis.

3.3 Questionnaire Design and Data Collection

To evaluate the learning impact of the module, students completed two anonymous questionnaires: one prior to the assignment (pre-questionnaire) and one after completion (post-questionnaire). The pre-questionnaire assessed students' baseline knowledge and expectations, including questions such as:

- "Have you previously heard of or used photogrammetry?"
- "What challenges do you anticipate in capturing or processing real-world objects?"

The post-questionnaire measured shifts in perception and self-assessed learning outcomes, using both Likert-scale items and open-ended questions, such as:

- How would you rate understanding of photogrammetry after the assignment?
- What was the most challenging part of the process?
- Would you consider using photogrammetry in future? Why or why not?

Responses were coded and analyzed thematically to identify common experiences and emergent learning patterns. By combining quantitative data from model outputs and qualitative data from student reflections, this study employed a mixed-methods case study design, enabling a holistic assessment of both technical proficiency and perceived educational value.

4 Results

This section presents the outcomes of the photogrammetry module in terms of student feedback, learning progression, and technical characteristics of the generated 3D models. Results are derived from both the questionnaires and the technical evaluation of the submitted assignments.

4.1 Student Feedback and Learning Outcomes

Pre-assignment awareness and expectations. Prior to the module, only 2 out of 6 students had heard of photogrammetry, and none had previously used it in any form. Most expected the process to be either “highly technical” or “difficult to manage without professional equipment.” On a 5-point Likert scale (1 = no knowledge, 5 = expert knowledge), the average self-assessed understanding of photogrammetry was **1.8**.

Post-assignment reflections. After completing the assignment, students reported a significant increase in understanding, with the average self-assessment rising to **4.0**. All students rated the experience as “valuable” or “very valuable,” and five out of six stated that they would consider using photogrammetry in their future projects. Open-ended responses highlighted several recurring themes:

- **Realism and satisfaction**: “It was amazing to see how real the object looked in Unity with real textures.”
- **Technical challenge**: “Lighting was tricky; I had to redo the photo session twice.”
- **Learning motivation**: “I want to try scanning larger scenes next.”

The post-questionnaire also revealed an increased confidence in using new tools and integrating external workflows into the game development process.

Perceived challenges. Students identified three main challenges during the assignment:

1. Ensuring sufficient photo coverage to avoid holes in the mesh.
2. Managing lighting conditions and surface reflectivity.
3. Processing time and system requirements, especially when using desktop software.

Despite these difficulties, all students completed the assignment successfully and gained practical insight into both the benefits and limitations of photogrammetry.

4.2 Technical Evaluation of Models

Each student submitted a complete asset pipeline: image set, reconstructed mesh, cleaned model, and Unity-integrated prefab. A summary of the technical data is provided in Table 1.

Table 1. Technical Summary of Student Photogrammetry Projects

Student	Object	Photos	Software	Triangles	Size (MB)	Time (Min)
Veljko Kovacevic	Statue	58	Meshroom	176,574	99.4	85
Luka Kurtic	Statue	100	RealityScan	7.955	1.4	20
Luka Rankovic	Chair	45	RealityScan	99,993	15	17
Lazar Stanisavljevic	Bottle	20	RealityScan	1,000,008	32	30
Dimitrije Stojanovic	Slipper	30	Polycam	13,390	3	12
Milos Nikolic	Bike	100	RealityCapture	256,453	143	254

Across all projects, texture quality was generally strong due to the use of real surface photography. Polygon counts varied based on software and optimization (Figure 1). Students using mobile applications produced lower-poly models by default, while those using desktop tools (e.g., Meshroom) submitted higher-resolution meshes, which required manual decimation in Maya to be suitable for game engines.

Fig. 1. 3D models created by students using photogrammetry method

5 Discussion

The results of this study provide meaningful insights into the integration of photogrammetry into a game development curriculum. Both the qualitative and quantitative outcomes confirm that photogrammetry not only enhances technical proficiency but also promotes creativity, engagement, and interdisciplinary thinking among students.

5.1 Educational Value and Skill Development

One of the most notable outcomes was the increase in students' confidence and competence in working with unfamiliar digital tools. The shift in self-assessed knowledge from an average of 1.8 to 4.0 on a 5-point scale, demonstrates that even a short, well-structured module can significantly enhance student understanding. This aligns with findings from other studies [3, 4], which emphasize that experiential engagement with photogrammetry enhances spatial reasoning and procedural knowledge.

In this module, students did not merely follow a theoretical introduction but engaged directly in the complete asset pipeline, from data acquisition to deployment in a real-time game engine. This approach reinforced applied knowledge and mirrored real-world workflows in game studios [1, 2], validating the pedagogical decision to integrate photogrammetry as a hands-on, project-based learning experience.

5.2 Tool Selection and Workflow Comparison

Students used a variety of software tools depending on hardware availability and personal preference, resulting in a natural comparison between mobile-based and desktop-based workflows. Those using mobile apps like RealityScan or Polycam benefited from faster, more automated pipelines but often faced limitations in texture sharpness and mesh optimization. In contrast, desktop tools like Meshroom and RealityCapture offered higher accuracy and control at the cost of increased complexity and processing time.

This distinction parallels findings in professional and academic literature [6, 10], which show that automatic photogrammetry solutions prioritize accessibility and speed, whereas semi-automatic or manual workflows, though more labor-intensive—allow for better customization and asset fidelity. Within the classroom context, both approaches proved pedagogically valuable: mobile apps lowered the barrier to entry, while desktop solutions challenged students to refine their technical workflows and practice digital cleanup and optimization.

5.3 Trade-offs: Accuracy vs. Efficiency

The technical results revealed substantial variance in polygon counts and model sizes (from under 10,000 to over 1 million triangles), depending on the reconstruction method and user intervention. Students who spent more time optimizing their models, either by cropping extraneous geometry or decimating dense meshes, produced assets that were both game-engine ready and aesthetically compelling. This highlighted the importance of balancing visual realism with performance constraints, a fundamental concern in game development [5, 7].

The assignment fostered awareness of data integrity and artifact correction. For example, students who failed to capture the underside of an object or used inconsistent lighting learned firsthand how incomplete data affects model quality. These real-world challenges taught students critical lessons in planning and executing digital capture, mirroring the practical difficulties encountered by professionals in the field [9].

5.4 Student Engagement and Motivation

From a motivational perspective, the project format proved highly effective. Students frequently expressed enthusiasm for the tangible nature of the task, photographing real-world objects and seeing them appear, realistically textured, within a game engine. This tangible transformation from physical to digital fostered a sense of accomplishment and ownership, echoing observations from other pedagogical studies that highlight the power of hands-on digital fabrication to increase engagement [3, 8].

The open-ended nature of the assignment encouraged creativity. Students selected personally meaningful objects, from statues and bikes to bottles and furniture, resulting in a diverse and culturally resonant portfolio of models. This individualized approach promoted intrinsic motivation and made the technical learning process more relatable and enjoyable.

6 Conclusion

This study demonstrated the successful integration of photogrammetry into an undergraduate game development course, where students created game-ready 3D models from real-world objects. Through a structured assignment involving image capture, 3D reconstruction, and game engine integration, students acquired essential technical skills while gaining insight into realistic asset pipelines used in the industry.

The module proved effective in increasing both student engagement and understanding of key concepts such as mesh optimization, texture fidelity, and tool interoperability. Variations in workflows—between mobile and desktop tools—provided a practical lens through which students evaluated the trade-offs between automation and control. However, the study had limitations, including a small sample size and short duration. Technical barriers such as hardware requirements and software compatibility also posed challenges. Future implementations should incorporate peer feedback, structured performance evaluation, and more advanced topics like environment-scale scanning or error quantification.

Future research should explore longitudinal effects of photogrammetry training, its integration into full game development pipelines, and its value in multidisciplinary collaborations, particularly in design, heritage, and AR/VR applications.

References

1. Remondino, F., El-Hakim, S.: Image-based 3D modelling: A review. *The Photogrammetric Record* 21(115), 269–291 (2006). https://doi.org/10.1111/j.1477-9730.2006.00383.x
2. Chiabrando, F., Piatti, D., Rinaudo, F..: Multi-Scale Modeling of the Basilica of San Pietro in Tuscania (Italy). From 3D Data to 2D Representation. *Geoinformatics FCE CTU* 6, 291–298 (2011). https://doi.org/10.14311/gi.6.37
3. Chapinal-Heras, D., Díaz-Sánchez, C., España-Chamorro, S., Gómez-García, N., et al.: Photogrammetry and 3D modelling in university teaching: A case study applied to the history degree. *The Journal of Classics Teaching* 24(e14), 1–9 (2023).
4. Sapirstein, P.: Accurate measurement with photogrammetry at large sites. *Journal of Archaeological Science* 66, 137–145 (2016).
5. Guzman, S., Koshelev, A., Wagner, S.: From Reality to Virtual Worlds: The Role of Photogrammetry in Game Development, (2025). DOI:10.48550/arXiv.2505.16951
6. Pejić, P., Krasić, S., Krstić, H., Dragović, M., Akbiyik, Y.: 3D virtual modelling of existing objects by terrestrial photogrammetric methods – case study of Barutana. *Tehnički vjesnik – Technical Gazette* 24(Suppl. 1), 233–239 (2017). https://doi.org/10.17559/TV-20150901102802
7. Holt, K.: Photogrammetry: 3D Modeling Landscapes and Structures, Medium, (2018). https://medium.com/digital-heritage/photogrammetry-3d-modeling-landscapes-and-structures-fa97a4b96f73
8. Koutsoudis, A., Vidmar, B., Ioannakis, G., Arnaoutoglou, F., Pavlidis, G., Chamzas, C.: Multi-image 3D reconstruction data evaluation. *Journal of Cultural Heritage* 15(1), 73–79 (2014). https://doi.org/10.1016/j.culher.2012.12.003
9. Guidi, G., Russo, M., Angheleddu, D.: 3D Survey and Virtual Reconstruction of archeological sites. Digital Applications in Archaeology and Cultural Heritage 1(2), (2014). DOI:10.1016/j.daach.2014.01.001
10. Dall'Asta, E., Roncella, R.: A comparison of semi-global and local dense matching algorithms for surface reconstruction. The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences XL-5(5), 187-194, (2014). DOI:10.5194/isprsarchives-XL-5-187-2014 (2014).

Authoring tools for e-learning: critical analysis and pedagogical perspectives. The case of H5P in interactive digital teaching

Alberto Fornasari[1][0000-0003-0553-8945], Gaetano Monaco[2][0009-0008-7677-6936], and Rosa Minerva[3][0009-0008-9289-3024] [1]

[1] Associate Professor of Experimental Pedagogy at the University of Bari 'Aldo Moro'
[2] PhD Candidate in Experimental Pedagogy at the University of Bari 'Aldo Moro'
[3] PhD in Experimental Pedagogy at the University of Bari 'Aldo Moro'
alberto.fornasari@uniba.it
gaetano.monaco@lediel.it
rosa.minerva@uniba.it

Abstract. This paper critically analyses authoring tools for e-learning, with a focus on the H5P platform, highlighting its transformative potential in digital teaching. H5P is an open-source technology capable of promoting interactive, inclusive and cognitively stimulating learning environments. The article emphasises that such tools should not be seen merely as technical solutions, but as pedagogical levers for a profound methodological restructuring of teaching. Through the integration of dynamic multimedia content and immediate feedback, H5P promotes motivation, self-regulation and meaningful learning. However, the effectiveness of digital innovation depends on the quality of instructional design and teacher training. Only conscious and thoughtful use of these environments can lead to real educational transformation, capable of going beyond the simple digital transposition of traditional teaching.

Keywords: H5P, E-Learning, Pedagogical Innovation.

1 Introduction

In recent years, the education sector has undergone a profound structural transformation, leading to increasing interaction with digital, online and hybrid learning environments [1]. This process, further accelerated by the COVID-19 pandemic [2], has forced school systems and academic institutions to adapt quickly to distance learning and blended learning, combining the flexibility and accessibility offered by virtual spaces with the importance of face-to-face interaction. This momentous transformation has led to a paradigm shift in education, prompting those working in the education sector to reconsider teaching practices in light of the potential offered by digital technology. In this context, the design of educational content has necessarily had to reformulate delivery processes, considering that, in today's scenario, teaching cannot ignore

[1] Author of Section 4

Reveiw Paper
DOI: https://doi.org/10.46793/eLearning2025.077F
Part of ISBN: 978-86-89755-37-4

hybridisation with multimedia materials [3,4]. The integration of resources such as images, sounds, videos and animations is no longer a simple complement to traditional teaching, but rather an important pedagogical tool for motivating students and supporting highly complex cognitive processes, with the aim of promoting meaningful learning development [5,6,7,8]. The richness and interactive nature of these materials act as educational mediators capable of stimulating reflection, active participation and the constructivist construction of knowledge, essential elements for the educational success of new generations, who are increasingly called upon to interact with electronic devices in everyday and professional contexts.

Within this framework, e-learning authoring tools take on a strategic position, acting as digital devices that enable educators to design, create and distribute interactive content in virtual environments. These tools, among which H5P (HTML5 Package) stand out for its versatility and integration, allow the creation of dynamic and customisable materials, such as interactive videos, quizzes, enriched presentations, timelines, simulations and drag-and-drop activities, which can be easily integrated into the main Learning Management Systems (LMS) platforms. The uniqueness of these tools lies in their ability to overcome traditional transmissive logic, promoting constructivist and socio-cognitive teaching approaches that value active participation, reflection and self-regulation of learning [9]. Furthermore, as authoring environments do not require advanced programming skills, they enable the dissemination and customisation of educational content even among teachers who are less experienced in technology.

From a systemic perspective, the adoption of authoring tools represents a strategic lever for the design of inclusive, differentiated learning environments that can respond effectively to the challenges posed by distance and blended learning. In this sense, scientific literature shows that these technologies have a significant impact on student motivation, engagement and satisfaction, positively influencing overall learning outcomes [10]. In particular, H5P stands out for its ability to build non-linear learning paths, in which students actively interact with varied and dynamic content, developing a reflective and personalised learning process that not only activates higher cognitive processes but also increases students' autonomy and responsibility in their own learning path [11,12].

The digitisation of teaching, therefore, is not just a technological issue, but a profound pedagogical transformation that broadens accessibility, promotes personalised learning and responds more effectively to the multiple cognitive, emotional and social needs of students in the third millennium. In this perspective, the role of authoring tools becomes central, enabling the creation of multimedia and interactive teaching materials that overcome the static nature of traditional teaching, enriching educational communication and stimulating active, self-regulated and inclusive learning processes [13]. Parallel to the evolution and growing popularity of *Open and Distance Learning* (ODL), there has been an increasing need to develop *Self-Instructional Materials* (*SIMs) SIMs)* that are not only highly interactive and attractive, but also based on rigorous pedagogical structures capable of responding in a differentiated and targeted manner to the multiple cognitive, affective and motivational profiles of an increasingly diverse learning population [14]. ODL is a strategic training method for meeting the flexibility needs of university students, but it has also been extended to secondary school learners, shifting the educational model from a transmissive and linear logic to a dialogic training process focused on operational autonomy, metacognitive awareness and the co-responsibility of the learner. In this perspective, interactive tools for teaching mediation are technological and pedagogical resources capable of effectively responding to some of the recurring critical issues in learning pathways, helping to make the educational experience more engaging on a sensory level, denser from a cognitive point of view and more relevant in terms of personal development [13].

In this perspective, the H5P platform allows the construction of highly interactive multimedia content, including enriched videos, dynamic questionnaires, simulations and animated presentations, which promote active learning processes, supported by immediate feedback mechanisms and continuous opportunities for meaningful interaction with other users of the platform [9,10,15]. Recent empirical evidence confirms that the integration of H5P in ODL contexts can positively impact on the overall quality of the learning experience, helping to improve learning outcomes and student satisfaction and counteract forms of disengagement and dispersion through teaching strategies focused on cognitive activation and continuous interaction between content and users [11,12].

In light of the reflections developed so far, the introduction of digital authoring tools cannot be interpreted simply as an accessory technological integration to traditional teaching devices, but rather as an effective epistemological restructuring of educational practices, capable of profoundly affecting the quality of teaching and learning processes. These tools are vectors of methodological innovation, promoting the design of learning environments that are more responsive to the cognitive, metacognitive and socio-relational needs of students, with a particular focus on promoting regulatory autonomy, inclusiveness and active participation.

In this direction, this paper aims to analytically explore the educational potential offered by authoring tools, with a specific focus on the H5P platform, in order to highlight the transformative impact that these technologies can have on blended and distance learning. The advanced features of H5P and their pedagogical implications in terms of cognitive activation, educational accessibility and personalisation of learning paths will be critically examined.

The aim is to highlight how the conscious use of highly interactive digital environments can not only support but also reshape traditional teaching, promoting a transition towards multimodal, dialogic training models centred on the reflective action of the learner. The proposed reflection, therefore, is part of a pedagogy of innovation based on criteria of educational sustainability, design flexibility and the enhancement of digital resources as complex teaching mediators.

2 Authoring tools

In the context of digital education, *authoring* tools for eLearning are advanced technological environments that enable the design and development of structured educational content using a wide range of multimedia resources, such as text, images, audio, video and interactive presentations, without the need for specialist IT expertise [16].These environments, which have profoundly changed the paradigm of educational production since the dawn of online training, now meet the dual need to simplify creation processes and ensure the pedagogical quality and accessibility of training materials. The uniqueness of these tools lies in their intuitive interface, based on visual design logic, which allows even those without advanced technical skills to structure personalised training courses. This approach promotes a significant democratisation of educational design, as it allows for the direct and active involvement of teachers and education professionals in the process of designing and developing training materials [17]. By eliminating the need for specific technical skills and reducing dependence on external figures such as programmers or designers, these tools allow teachers to translate their pedagogical strategies into digital content with greater autonomy and timeliness, promoting a closer connection between real teaching needs and the technological solutions adopted, helping to make the training project more tailored to the specific needs of students and more consistent with the educational objectives pursued. In this way, accessibility to *authoring* tools is a key element in enhancing the professional skills of educators and promoting collaborative and participatory practices in the construction of learning experience [17].

The effectiveness of authoring tools is also evident in their ability to ensure high content usability, thanks to responsive interfaces that allow access from a variety of devices such as smartphones, tablets and computers, promoting continuity of the learning experience even when on the move. In this context, the integration of interactive features, such as simulations, multiple-choice exercises, branching paths and *gamification* elements [18], makes it possible to overcome the traditional linearity of transmissive teaching, promoting cognitive activation and personalisation of learning.

Particularly important is the focus on accessibility, which translates into the presence of technological devices capable of meeting the needs of a diverse student population: keyboard navigation aids, automatic subtitling, compliance with international guidelines for digital accessibility (WCAG), as well as automatic transcription of audiovisual content, which, in addition to ensuring inclusion, also facilitate translation and semantic search in training materials.

Another important aspect, according to [19], concerns the collaborative potential of these platforms, which facilitate co-design and content sharing between different actors involved in the training chain, facilitating the re , updating and integration of materials. The presence of centralised libraries for teaching resources, templates, lesson models and exercise databases is a further element supporting the standardisation and effectiveness of *instructional design* processes.

Authoring tools for creating eLearning content are mainly distinguished by how they are used and integrated into digital learning environments, thus forming different functional and operational categories. The first category includes desktop-based tools, such as Articulate and Adobe Captivate, which require local installation on dedicated devices. This software is known for offering a wide range of advanced features, allowing developers to design highly customised and complex content, integrating multimedia elements, simulations and sophisticated interactions. However, this type has limitations in terms of accessibility and collaboration, as its use is restricted to the machine on which the software is installed and requires specific technical skills to manage and update content.

In contrast, cloud-based tools, such as Genially, H5P and Canva, operate entirely on web platforms, offering greater flexibility and ease of access regardless of the device used. These tools facilitate collaborative work, allowing multiple users to work simultaneously on the same educational project and enabling immediate and centralised updates [20]. Furthermore, cloud-based nature significantly reduces the costs and time associated with software installation and maintenance, promoting an agile and inclusive development model. This approach effectively responds to contemporary needs for flexibility and scalability in educational production, especially in complex institutional and organisational contexts.

A further category of tools is represented by those integrated directly into Learning Management Systems (LMS), such as Moodle Book or Google CourseKit. These tools allow content to be created and managed within the same platform used for training delivery and tracking, promoting an integrated workflow that reduces discontinuities between production and distribution. This integration facilitates centralised course management, learning data collection and the customisation of training programmes, representing an effective solution especially for institutions that adopt LMS as their main training infrastructure.

To ensure that content produced with these different tools can be shared, reused and monitored effectively on heterogeneous platforms, the adoption of interoperability standards is essential. Among these, SCORM (Sharable Content Object Reference Model) is one of the most established standards for the packaging and distribution of reusable training modules, allowing the tracking of learning activities in a standardised way. More recently, the Experience API (xAPI) has expanded data collection capabilities, allowing learning experiences to be recorded in unstructured contexts or outside traditional LMSs, enabling a more detailed analysis of the training process. Finally, the Learning Tools Interoperability (LTI) standard facilitates the integration of external tools within LMSs, enabling smooth and secure communication and inter tion between different platforms, which is crucial for increasingly complex and interconnected training environments. Together, these standards form an essential pillar for ensuring the technical compatibility, accurate traceability and scalability of eLearning content, contributing substantially to the quality and effectiveness of digital training.

The pedagogical evaluation of eLearning authoring tools is a multidimensional process aimed at determining their ability to support effective educational practices consistent with contemporary learning principles. Firstly, the usability of the tool is a determining factor, as an intuitive, accessible and functionally efficient interface facilitates instructional design by teachers, minimising technical barriers and promoting more fluid content management and updating [21]. This cognitive and operational accessibility is essential to ensure that educators can focus on pedagogical aspects without being hindered by technological difficulties. At the same time, the inclusion of interactive elements plays a key role, as educational literature emphasises that active participation and the social construction of knowledge are fundamental to deep, lasting and meaningful learning processes [22,23,24]. Authoring tools that integrate interactive activities, such as dynamic quizzes, simulations and branching scenarios, not only facilitate students' cognitive engagement, but also promote reflection and critical thinking, which are essential for meaningful learning [25].

Another essential criterion concerns the adaptability of tools and their compliance with international accessibility standards, such as the Web Content Accessibility Guidelines (WCAG) developed by the World Wide Web Consortium (W3C). These standards are fundamental to ensuring that educational materials are accessible to users with different sensory, motor or cognitive abilities, helping to reduce educational inequalities and promote an inclusive paradigm that respects the right to education for all [26,27]. Finally, the ability of authoring tools to support the personalisation of the training path and continuous learning monitoring is a crucial aspect in a learner-centred teaching approach. The ability to adapt content and teaching strategies to individual needs, as well as to collect detailed performance data, enables targeted educational interventions and promotes metacognitive and motivational self-regulation processes [28,29]. This level of personalisation and traceability fits perfectly into the contemporary paradigm of dynamic; inclusive education geared towards the development of transversal skills. However, the widespread adoption of digital tools in education is not without limitations and critical issues that affect their pedagogical effectiveness and real innovativeness. Among the most common problems is the lack of adequate and systematic training for teachers on the potential and methods of integrating educational technologies. Numerous studies highlight how insufficient preparation and a lack of specific skills in the pedagogical use of digital tools and h s represent a significant obstacle to the implementation of truly transformative teaching practices [30,31] This training deficit often results in their use being limited to purely operational or technical functions, without any real reflection on the methodological reorganisation of teaching. Furthermore, there is a widespread tendency to adopt solutions that are predominantly 'content-based', i.e. focused on the simple digitisation of traditional teaching materials, rather than on innovative pedagogical design. This reductionist approach reduces digital tools to mere vehicles for the transmission of knowledge, neglecting the interactive, collaborative and adaptive potential offered by emerging technologies [32]. The side effect is digital teaching that mechanically replicates existing transmission models, without promoting active, critical and meaningful learning processes.

The most insidious risk, therefore, lies in the digital replication of traditional teaching practices, which continue to favour a unidirectional and passive mode of knowledge transfer. Such reproduction risks compromising the effectiveness of technological innovation, creating a dichotomy between the revolutionary potential of tools and their actual application in educational contexts [33, 34]. The failure to transform teaching and learning epistemologically and methodologically through digital technology results in an uninspiring and uninclusive educational experience that does not take full advantage of the opportunities offered by technological mediation.

Overcoming these critical issues requires a structural and ongoing investment in teacher training, aimed not only at acquiring technical skills, but above all at building an integrated and critical pedagogical vision. Only through a conscious and reflective approach will it be possible to avoid the simple digital transposition of traditional teaching, promoting instead innovative, participatory and learner-centred educational processes.

3 The H5P case: potential, experiences and pedagogical reflection

In the contemporary context of digital technologies applied to education, H5P stands out as one of the most relevant and versatile open-source solutions for the design, creation and dissemination of interactive educational content. This development environment, whose acronym stands for “HTML5 Package”, was launched in 2013 thanks to an initiative promoted and funded by the Norwegian Centre for ICT in Education. It is a technological artefact inspired by an inclusive and democratic logic, aimed at promoting the equal dissemination of digital educational resources [35].

The design philosophy underlying the platform is based on the adoption of open and interoperable paradigms that encourage systematic sharing, content reusability and the possibility of adaptation and reworking by users. In stark contrast to the proprietary and closed models of commercial platforms, H5P stands out for its open-source nature, which not only allows free use of the tools provided, but also access to and modification of the source code. This promotes the construction of a participatory pedagogical ecosystem, in which teachers, learners and developers collaborate in the co-creation of digital learning experiences.

From a structural point of view, the platform is based on an extremely flexible modular architecture that currently supports over forty types of educational content. These include, for example, interactive quizzes, enriched audiovisual content, timelines, memory games, branching paths, completion exercises and drag-and-drop activities. These resources are characterised by a high level of interactivity and extensive customisation options, making them ideal tools for supporting active, student-centred teaching practices.

Another strength of H5P is its native compatibility with major Learning Management Systems (LMS), such as Moodle, WordPress and Drupal. These platforms, being open source, are also widely used in education and professional settings for managing and publishing digital content, enabling the creation of virtual learning environments, academic blogs, institutional websites and complex content management systems. The seamless integration between H5P and these LMSs ensures not only the scalability of training content delivery, but also centralised and consistent governance of digital learning environments. This structural and functional interoperability is a competitive advantage for educational institutions, as it allows the implementation of dynamic and integrated educational pathways within established digital contexts, responding to a systemic and complex vision of technology-mediated teaching [36].

Motivation to learn can be seen as the main driving force that lets students see themselves as active and intentional agents in the educational process. It not only encourages genuine commitment but also leads to a process of cognitive and emotional restructuring through which students rework their self-perception and redefine their role in the educational relationship [37]. In this perspective, promoting a qualitative increase in school motivation is equivalent to designing educational environments characterised by a high level of cognitive and emotional stimulation, selecting content capable of arousing interest and meaning, and adopting digital tools that allow immediate and bidirectional interaction, providing timely, continuous and personalised feedback.

In this regard, the contributions of Susetyarini et al. [38] are particularly relevant, as they offer a comprehensive interpretative model aimed at assessing students' school motivation through a series of indicators observable in the classroom context. The authors' proposal is based on a threefold perspective, which considers, on the one hand, the level of active involvement of students and their ability to persevere with continuity and determination in carrying out the tasks assigned; on the other hand, the quality of the attitudes adopted and the emotions expressed during the educational experience, understood as signs of affective-cognitive disposition towards learning. Finally, the study also examines behaviours that occur in less structured moments of school life, such as breaks, transitions or non-directly educational phases, which, although not formally part of the teaching-learning process, are emblematic of the student's latent motivation and general attitude towards the educational environment.

To supplement and expand on the theoretical framework outlined above, [39] proposes a more detailed interpretative perspective, in which learner motivation is conceived as the dynamic result of a variety of interrelated factors, each with specific pedagogical and theoretical relevance. Among the elements identified, the first to emerge is the level of personal aspiration, understood as the student's ability to attribute meaning and purpose to their educational path, orienting their choices and efforts within a meaningful educational vision. A second element concerns confidence in one's abilities, i.e. the degree of perceived self-efficacy that directly influences commitment, perseverance and willingness to tackle complex cognitive challenges. Added to this are the psycho-physical conditions of the individual, which include emotional, physical and mental dimensions and act as facilitating or inhibiting factors with regard to active participation in learning processes. No less important is the influence of the living environment, including cultural, social and family variables, which, although external to the school setting, profoundly shape the motivational dispositions and educational expectations of the student. Finally, the teaching strategies implemented by the teacher play a crucial role: these include, in particular, clarity of presentation, the ability to create a positive and inclusive relational climate, and the adoption of creative methodologies capable of involving students in an authentic and meaningful way in the educational process.

It is in this theoretical context that the strategic usefulness of H5P can be found, as a digital platform capable of significantly increasing students' motivation levels while supporting teachers who, for various reasons, may be less inclined to experiment with teaching or creative design. Digital environments, due to their intrinsic media structure and the immediacy of communication that characterises them, are inherently more stimulating than transmissive teaching practices, which are rigidly frontal and based on the mere mnemonic reproduction of content [40].The use of a platform such as H5P stimulates students to take responsibility for their own learning, activating cognitive and metacognitive processes that shift the focus of educational interaction from a passive, teacher-centred paradigm to an active, participatory and *learner-centred* model. This model is not limited to the simple execution of tasks, but stimulates and supports the learner's awareness in decision-making, encouraging the development of willpower, critical thinking and autonomy, in line with their inclinations and potential.

Educational technologies such as H5P have a profound impact on the motivational and emotional aspects of learning: they simultaneously activate thought, emotions, interest and attention, promoting effective, interactive and multisensory pedagogical communication between teachers and students. As teaching mediation tools, these platforms are highly effective in transmitting content, as they combine standardisation with a high degree of customisation, helping to reduce interpretative ambiguity and increase the transparency of the educational message [41,42]. According to Kristanto [43], these environments promote clearer learning that is more consistent with the learner's cognitive style and facilitate effective educational interaction that respects the individual rhythms, modes and characteristics of the student.

It is worth emphasising the strategic role that teaching materials play in digital technology-mediated teaching, especially in online contexts. Thanks to the continuous evolution of information and communication technologies, these materials have undergone a profound transformation, evolving from predominantly transmissive tools to multimedia devices capable of generating immersive, dynamic and highly interactive learning experiences. Currently, teaching resources can be designed to synergistically integrate different communicative elements, including text, static images, audiovisual content and animations, in order to stimulate active participation by the learner, promoting continuous and meaningful interaction between the user and the system.

In support of this perspective, Chen et al. [3] highlighted the educational effectiveness of using interactive digital content developed through the H5P (HTML5 Package) environment, which is particularly advantageous in terms of greater visibility and accessibility of materials, as well as more careful consideration of students' cognitive, motivational and educational needs. In particular, H5P content is a suitable tool for promoting active and constructive learning, thanks to its ability to stimulate cognitive interaction and reflective processing: it requires learners to interact directly with the content, tackle complex problems and implement strategies for applying the knowledge acquired, thus promoting meaningful and situated learning.

One of the most important aspects of this tool is its immediate feedback function, which is a powerful self-reflection device, as it allows students to monitor their learning process in real time, correct any mistakes and reorient their cognitive strategies. This type of feedback, which is part of a self-assessment process, promotes the development of autonomy in the learning process and reinforces the sense of self-efficacy.

Being structurally based on the HTML5 standard, the H5P platform also guarantees high cross-device accessibility, making it usable on a wide variety of digital devices – from computers to tablets and smartphones – without compromising its usability or educational effectiveness. Its native integration with various virtual learning environments further amplifies its versatility, making it compatible with a variety of educational contexts and adaptable to a wide range of disciplinary and methodological objectives [44].

From a pedagogical point of view, the features offered by H5P align with student-centred learning models, supporting approaches such as *microlearning* and situated learning. Support for *microlearning* takes the form of ' ', the possibility of building modular, short-term teaching micro-units that promote continuous, just-in-time and highly personalised learning [45]. This approach is particularly effective in flexible educational contexts, such as continuing education and professional learning, but also in secondary schools, where short content episodes allow for greater attention and cognitive engagement. The promotion of interactivity is a further strength: H5P allows for the structuring of non-linear learning paths, encouraging active exploration and self-regulation of the learning process. The immediate feedback tools integrated into the content, as already mentioned, facilitate a continuous cycle of action-evaluation-reflection, in line with constructivist learning theories and the formative assessment model [46]. Furthermore, the ability to create easily reusable, clonable and adaptable learning objects makes H5P a highly efficient tool in terms of learning design sustainability: an activity designed for a university module can be easily reconfigured for a secondary school lesson, maintaining its structure and interactivity while updating the content. In terms of inclusivity, H5P complies with WCAG 2.1 accessibility standards and offers a responsive interface that can be used on mobile and desktop devices, thus contributing to reducing the digital divide and promoting equal access to education. The combination of interactivity, adaptability and accessibility makes H5P a tool capable of effectively responding to the diverse educational needs of today's world [47]. In school settings, studies conducted in Germany and Spain have shown that the use of H5P content in secondary schools is positively correlated with increased intrinsic motivation among students and greater participation in asynchronous activities, especially in integrated digital learning contexts [48]. The most recent systematic reviews [49], confirm that the use of interactive tools such as H5P has a positive impact on self-efficacy perceptions, sense of belonging to the learning community and course completion rates, contributing significantly to the construction of motivating and cognitively challenging learning environments. However, it should be emphasised that the effectiveness of H5P is not intrinsic to the tool but depends largely on the quality of the pedagogical design: technology can in fact become a facilitator or an obstacle, depending on its consistency with the educational objectives, the clarity of the instructions and its suitability for the students' profile. In this sense, pedagogical reflection on the use of H5P requires a rethinking of the roles of teachers, who must take responsibility for mediating the teaching- moving from simple content providers to designers of cognitive and metacognitive environments.

4 Conclusion and future prospectives

This study aimed to provide a critical and in-depth examination of authoring tools for e-learning, with a particular focus on the paradigmatic case represented by H5P. This tool is a highly versatile, interoperable, modular and accessible open-source platform that has demonstrated significant potential in terms of promoting interactivity, personalisation and cognitive activation in digital learning contexts [47]. H5P is therefore a virtuous example of educational technology capable of combining technical requirements and pedagogical needs, offering an environment conducive to the development of meaningful, constructivist and situated learning experiences. However, it is necessary to avoid any uncritical mythologisation of the tool. H5P, although highly functional and supported by a large international community of developers and educators, cannot be seen as a solution to the structural problems of digital teaching. Its effectiveness is not intrinsic to the technological device itself but is only realised to the extent that it is integrated into intentional, critical and conscious pedagogical design [33]. Converging empirical and theoretical evidence emphasises that the positive educational outcomes of using H5P are systematically mediated by the quality of the educational design, methodological consistency and the appropriateness of the educational setting in which these tools are used [50]. In outlining a critical perspective, it is essential to emphasise that educational technology is never neutral, but deeply conditioned by the underlying epistemologies and practices that substantiate it. The introduction of digital tools such as H5P, if not accompanied by adequate methodological transformation, risks remaining a purely aesthetic operation, generating digitisation that is devoid of pedagogical effectiveness in traditional teaching [34]. In other words, the digital replication of transmission models is one of the most insidious and regressive outcomes of technological innovation disconnected from genuine pedagogical reflection. From this perspective, the role of the teacher must be completely reconfigured: no longer a simple provider of content, but a cognitive architect, director of educational interaction and mediator of knowledge. This professional reconfiguration requires the solid acquisition of *instructional design* skills, i.e. the ability to design authentic, dialogic, multisensory and meaningful learning environments. Without specific training in this area, authoring tools risk being used in a mechanical or standardised way, losing their transformative potential.

Numerous international studies highlight a systemic deficiency in the initial and in-service training of teachers with regard to the effective integration of educational technologies [51,52]. The gap between technological availability and pedagogical competence (the so-called 'second-level digital divide') is a real obstacle to the construction of truly innovative and student-centred learning environments [53].

The experience with H5P highlights how the complexity of technological mediation requires a systemic vision that integrates devices, content and practices within a coherent pedagogical framework. In the absence of such a framework, there is a risk of confusing technical interactivity with cognitive interactivity, multimedia with the plurality of educational languages, and personalisation with the individualistic fragmentation of learning paths.

In light of these findings, a number of strategic lines of development are needed to consolidate the integration of authoring tools such as H5P in educational contexts, promoting their sustainable, informed and pedagogically sound use.

It is a priority to provide teachers and educational institutions with theoretical and operational frameworks to guide the educational use of authoring tools. These guidelines should be structured on three levels: the first concerns the pedagogical principles of reference (constructivism, situated learning, dialogicity); the second concerns the design criteria for the creation of effective multimedia objects (interactivity, accessibility, modularity); the third, on the other hand, concerns implementation and evaluation protocols.

Training courses aimed at developing integrated *instructional design* skills that go beyond mere technical ability should be promoted. In this sense, training should be situated, i.e. carried out in realistic and authentic contexts; collaborative, with peer co-design experiences; reflective, favouring processes of educational metacognition. Teaching workshops based on authoring environments such as H5P, if properly guided, can be fertile ground for methodological experimentation, the enhancement of creativity and the activation of continuous professional learning pathways [20]. In a context increasingly focused on the quality and effectiveness of teaching, it is essential to move beyond assessment based on purely quantitative indicators (frequency of use, connection time, number of clicks) and adopt multi-level assessment models capable of capturing the real impact of digital learning objects on learning. The integration of tracking tools such as xAPI allows the collection of granular data on student interactions, which, if correctly interpreted, can offer valuable insights into engagement, self-regulation, critical thinking and meaning construction [54].

Furthermore, the introduction of participatory assessment methodologies is recommended, directly involving students in reflection on learning processes and promoting forms of formative and metacognitive assessment.

The case of H5P provides a privileged observatory on the potential and limitations of the digitisation of teaching: on the one hand, it demonstrates the enormous opportunities offered by digital tools in the construction of interactive, adaptive and inclusive learning environments; on the other hand, it highlights the fragility of an education system that often struggles to integrate these tools into a mature and coherent pedagogical vision.

The challenge facing us today is not to use technologies without making any changes to the teaching-learning process, but rather to rethink education itself through and with technologies. This implies an epistemological, methodological and professional repositioning, which must start with teacher training and lead to the co-construction of a shared, critical and inclusive digital pedagogical culture.

Only through structural and multidimensional investment will it be possible to achieve a real transition towards an educational model capable of enhancing the potential of technology without sacrificing the centrality of relationships, reflection and educational planning.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Singh, Vandana, and Alexander Thurman. 2019. "How Many Ways Can We Define Online Learning? A Systematic Literature Review of Definitions of Online Learning (1988– 2018)." American Journal of Distance Education 33 (4): 289–306. https://doi.org/10.1080/08923647.2019.1663082.

[2] Mishra, S., Sahoo, S., & Pandey, S. (2021). Tendenze della ricerca nell'apprendimento a distanza online durante la pandemia di COVID-19. Distance Education, 42 (4), 494-519.

[3] Chen, L., Manwaring, P., Zakaria, G., Wilkie, S., & Loton, D. (2021). Implementing H5P Online Interactive Activities at Scale. Proceedings ASCILITE 2021, December, 81–92. https://doi.org/10.14742/ascilite2021.0112.

[4] Bozdag, Hüseyin C., and İsa Gokler. 2023. "Digital Content Design for the Flipped Classroom Model: Example of Biology Lesson." Journal of Computer and Education Research 11 (21): 335–355. https://doi.org/10.18009/jcer.1246524.

[5] Dinç, Emre, and April Millet. 2022. "Evaluating Different Assessment Types in an Online Geoscience Course." International Association for Development of the Information Society 53 (1): 226–232. https://doi.org/10.1177/00472395241254825.

[6] Chasani, Akrom, Muhammad Nasir, and Erviyenni. 2023. "Development of Interactive Learning Media for H5P-Based Elasticity Materials in the Mobilizing School Curriculum." Journal Penelitian Pendidikan IPA 9 (4): 2089–2096. https://doi.org/10.29303/jppipa.v9i4.3109.

[7] Wehling, Judith, Stefan Volkenstein, Stefan Dazert, et al. 2021. "Fast-Track Flipping: Flipped Classroom Framework Development with Open-Source H5P Interactive Tools." BMC Medical Education 21 (1): 1–10. https://doi.org/10.1186/s12909-021-02784-8.
[8] Lai, Jennifer, and Matt Bower. 2019. "Evaluation of Technology Use in Education: Findings from a Critical Analysis of Systematic Literature Reviews." Journal of Computer Assisted Learning 36 (3): 241–259. https://doi.org/10.1111/jcal.12412.
[9] Selvarasu, E., Mohammad, A. R., Farzana, S., Mohammed, N., Pillai, A., & Govindaraj, S.: Creating interactive teaching content using Moodle and H5P – Teachers' perspectives from the Colleges of Technology in Oman. Journal of Critical Reviews 7(13), 3976–3981 (2020).
[10] Mir, Kamran, Muhammad Iqbal, and Jahan Shams. 2022. "Investigation of Students' Satisfaction about H5P Interactive Video on MOODLE for Online Learning." International Journal of Distance Education and E-Learning 7 (1): 71–82. http://doi.org/10.36261/ijdeel.v7i1.2228.
[11] Siregar, Alfina Gustiany, and Friscilla Sembiring. 2022. "Interactive Learning Content Using H5P in Pronunciation Course." Journal of Education, Humaniora and Social Sciences (JEHSS) 5 (2): 1219–1225. http://doi.org/10.34007/jehss.v5i2.1474.
[12] Mutawa, A. M., Al Muttawa, J. A. K., & Sruthi, S. (2023). The effectiveness of using H5P for undergraduate students in the asynchronous distance learning environment. Applied Sciences (Switzerland), 13(8). https://doi.org/10.3390/app13084983.
[13] Marelli, M., & Castelli, L. (2021). Ripensare la scuola: gli strumenti tecnologici al tempo della Didattica Digitale Integrata. Dirigenti Scuola.
[14] Shawira Abu Bakar, Nazrai Ahmad Zabidi, Nooni Ezdiani Yasin, & Siti Aishah Hashim Ali. (2022a). Measuring the impact of COVID-19 pandemic for continuous development of learning materials. Proceedings of the 35th Asian Association Open University Annual Conference, Vol. 1, 54–60. https://2022aaou.org/file/Conference%20Proceedings%20-%20Vol.1.pdf.
[15] Palma, Carmen, and Mauro Garzón. 2023. "The Use of a Virtual Environment to Improve Students' Listening Skills: A Learning Analytics Approach." Ciencia Latina Revista Científica Multidisciplinar 7 (2): 2937–2953. https://doi.org/10.37811/cl_rcm.v7i2.5537.
[16] Naidu, VR, Najah, S., Saqib, M., Swathi, R. e Pandey, N. (2023). Trasformare l'e-learning attraverso soluzioni di authoring multimediale interattivo basate su cloud. In SHS Web of Conferences, EDP Sciences.
[17] Gratani F.,: Makers at school: L'apprendimento nell'era post-digitale. FrancoAngeli, Milano (2023).
[18] Feng-Jung L., Chia-Mei L., (2021), Design and implementation of a collaborative educational gamification authoring system, International Journal of Emerging Technologies in Learning, 16(17), 277.
[19] Nilson, L. B., Goodson, L. A.: Online teaching at its best: Merging instructional design with teaching and learning research. John Wiley & Sons, Hoboken (2021).
[20] Lo Presti F., Zizza S., (2024), L'uso delle nuove tecnologie nei percorsi didattici e di formazione. Possibilità di sviluppo e versanti critici, Nuova Secondaria, 346-355.
[21] Norman, D.: Le cose che ci fanno intelligenti. Feltrinelli, Milano (1995).
[22] Novak, J. (2002), L'apprendimento significativo, Trento, Erickson.
[23] Vygotsky L. S., (1978), Mind in Society: The Development of Higher Psychological Pro cesses, Harvard University Press.
[24] Mayer R. E., : Constructivism as a theory of learning versus constructivism as a prescription for instruction. In Constructivist instruction. Routledge, London (2009).

[25] Bransford J. D., Brown A. L., Cocking R. R., : How people learn. National Academy Press, Washington, DC (2000).
[26] Seale J.,: E-learning and disability in higher education: Accessibility research and practice. Routledge, London (2013)
[27] Al-Azawei, A., Serenelli, F., & Lundqvist, K., (2016), Universal Design for Learning (UDL): A content analysis of peer reviewed journals from 2012 to 2015, Journal of the Scholarship of Teaching and Learning, 16(3), 39-56.
[28] Zimmerman B. J., (2002), Becoming a self-regulated learner: An overview. Theory into practice, 41(2), 64-70.
[29] Panadero E., (2017), A review of self-regulated learning: Six models and four directions for research, Frontiers in Psychology, 8, 422.
[30]Ertmer P. A., Ottenbreit-Leftwich A. T., (2010), Teacher technology change: How knowledge, confidence, beliefs, and culture intersect, Journal of research on Technology in Education, 42(3), 255-284.
[31] Tondeur J., Van Braak J., Ertmer P. A., Ottenbreit-Leftwich A., (2017), Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence, educational technology research and development, 65, 555-575.
[32] Selwyn N.,: Is technology good for education? John Wiley & Sons, Hoboken (2016).
[33] Laurillard, D. (2014). Insegnamento come scienza della progettazione. Costruire modelli pedagogici per apprendere con le tecnologie. Milano: FrancoAngeli.
[34] Rivoltella P. C., Rossi P. G., (2024), Tecnologie per l'educazione, Torino:Pearson.
[35] Pagliara, S. M. (2025). Tecnologie educative e inclusione. Prospettive, metodologie e innovazione, Milano: FrancoAngeli.
[36] Panciroli C., (2021). Elementi di didattica post-digitale, Bononia University Press.
[37] Suhana, C. (2014). Konsep Strategi Pembelajaran. In Bandung: Refika Aditama (p. 78). PT. Refika Aditama.
[38] Susetyarini, R. E., Permana, T. I., Gunarta, G., Setyawan, D., Latifa, R., & Zaenab, S.: Motivasi dan tanggung jawab siswa dalam pembelajaran berbasis proyek, sebuah penelitian tindakan kelas. Jurnal Inovasi Pendidikan IPA 5(1), 1–9 (2019).
[39] Setiawan, M. A. (2017). Belajar Dan Pembelajaran Tujuan Belajar Dan Pembelajaran. In Uwais Inspirasi Indonesia (Issue August). Uwais Inspirasi Indonesia Redaksi.
[40] Ristika, R., & Lestari, H. (2024). Peningkatan Motivasi dan Tanggung Jawab Belajar Berbasis Project Based Learning dalam Memahami Karya Ilmiah Siswa Kelas XI SMA Negeri 3 Palembang. Socius: Jurnal Penelitian Ilmu-Ilmu Sosia, 1(March), 37–41.
[41] Prasong, F., Kalang, F., & Nurlailah.: Pengembangan Media Video Animasi Interaktif Untuk Meningkatkan Minat Membaca Dengan Tema Kegemaranku Pada Siswa Kelas 1 Mis Al-Fitrah Oesapa. Jurnal Pendidikan Dasar Flobamorata 3(1), 276–280 (2022).
[42] Pagarra, H., Syawaluddin, A., Krismanto, W., & Sayidiman. (2022). Media Pembelajaran. In Badan Penerbit UNM. Badan Penerbit UNM.
[43] Kristanto, A. (2016). Media Pembelajaran. In Bintang Sutabaya. Bintang.
[44] Sudarmawanto, S., Hartati, S. J., & Sucipto, S. (2025). Inductive Learning in The Digital Era: Assessing the Effectiveness of HTML 5 Package (H5P) Based Media on Learning Motivation and Responsibility. Jurnal Paedagogy, 12(3).
[45] Hug, T. (2021). Sound pedagogy practices for designing and implementing microlearning objects. In Microlearning in the Digital Age (pp. 33-57). Routledge.
[46] Black, P., & Wiliam, D.: Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability 21(1), 5–31 (2009).

[47] Reisoğlu, İ. : How does digital competence training affect teachers' professional development and activities? Technology, Knowledge and Learning 27(3), 721–748 (2022).
[48] Gómez-García, M., Ortega-Sánchez, D., & Dimitriadis, Y. (2022). Interactive videos and gamification in H5P: Impact on motivation and academic performance in secondary education. Computers & Education, 190, 104610.
[49] Tang, Y., Hew, K. F., & Chen, W. (2023). Interactivity in online learning: A meta-analysis and research synthesis of empirical studies. Computers in Human Behavior, 139, 107540.
[50] Molenaar, I., de Mooij, S., Azevedo, R., Bannert, M., Järvelä, S., & Gašević, D. (2023). Measuring self-regulated learning and the role of AI: Five years of research using multimodal multichannel data. Computers in Human Behavior, 139, 107540.
[51] De Sanctis G., (2010), TALIS. I docenti italiani tra bisogni di crescita professionale e resistenze, Fondazione Giovanni Agnelli Programma Education FGA Working Paper, 24(2), 1-34.
[52] Di Donato, D., De Santis, C. (2021), Il cambiamento delle pratiche didattiche dei docenti italiani durante il lockdown. Percezioni dell'efficacia nell'uso delle tecnologie didattiche digitali e collaborazione con i colleghi, RicercAzione, 13(1), 213-233.
[53] Ma, J. K. H., & Cheng, S. (2022). First-and Second-Level Digital Divides from 2009 to 2018. In Adolescent Well-Being and ICT Use: Social and Policy Implications (pp. 167-183). Cham: Springer International Publishing.
[54] Panciroli, C., & Rivoltella, P. C. (Eds.). (2025). IA in classe: Didattica con e sull'Intelligenza Artificiale. Edigita ebooks.

Monetization of a Multi-Sided Innovative Platform – A Case Study of the STEDDY Project

Sanja Dalton[1] [0009-0004-2163-232X] and Jefto Dzino[2] [0009-0007-8913-752X]

[1,2]Belgrade Metropolitan University, 63 Tadeuša Košćuška St, Belgrade, Serbia

sanja.dalton@metropolitan.ac.rs, jefto.dzino@metropolitan.ac.rs

Abstract: This paper analyzes the commercialization potential of an innovative multi-sided platform for personalized learning using a qualitative research approach based on both primary and secondary data. The data was collected through surveys and semi-structured interviews with high school students, university students, and employed individuals in need of personalized upskilling. Through case study methodology and the analysis and synthesis of the collected data, the results indicate that the monetization of the innovative platform is feasible and sustainable.

Keywords: Innovative platform; platform business model; personalized learning; monetization of platforms

1. Introduction

Modern educational technologies increasingly rely on personalized learning, adapting content, pace, and methods to meet the individual needs of users. Personalized learning platforms represent dynamic digital tools that utilize user data to provide optimal educational experiences for various target groups.

These platforms are an effective tool for educating all generations. Their primary value lies in the ability to respond to different learning styles, needs, and preferences, thereby improving the quality of education and contributing to better motivation and learning outcomes. In today's digital environment, personalized learning is becoming an increasingly important approach, as it allows for the customization of educational content, pace, and methods according to individual needs. Platforms that enable this type of education often function as multi-sided platforms, connecting various user groups – learners, students, instructors, training organizations, and even employers.

What contributes most to the profitability of platforms are cloud-based infrastructures (i.e., IaaS, PaaS, SaaS, and FaaS), among which IaaS should be highlighted for its scalability and flexibility.

Besides, in order to ensure sustainability and further development of such platforms, it is essential to develop an effective monetization model that enables revenue

[1] sanja.dalton@metropolitan.ac.rs
[2] jefto.dzino@metropolitan.ac.rs

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.092D
Part of ISBN: 978-86-89755-37-4

generation without compromising the user experience. Therefore, monetization of multi-sided educational platforms can be achieved through various mechanisms: subscriptions, course fees, instructor commissions, advertisements, as well as partnerships with institutions and companies investing in the skill development of their employees.

2. LITERATURE REVIEW

2.1. Platform-based business

In the literature, platform-based business is also known as a business model defined as "a system consisting of components, linkages among these components, and dynamics" [2]. On one hand, the business model identifies the transaction partners, defines the value proposition for each, and describes how the core firm connects with them [4]. On the other hand, it explains how value is delivered, monetized, and shared among the partners [20].

According to some authors, it is essential to define how value is delivered, monetized, and shared in platform-based business models [38], "emphasizing the interdependence between activities that helps to understand the structural relationships between the platform provider and its users" [5].

The literature also describes multi-sided platforms as hubs or intermediaries that facilitate the exchange of value between two or more user-producer groups [34, 17]. For instance, Cennamo and Santaló (2015, p. 12) define multi-sided platforms as "networks that connect two or more distinct types of users and enable transactions among them" [7], while McIntyre and Srinivasan (2017, p. 143) conceptualize them as "interfaces that can mediate transactions between two or more sides." [24] These definitions implicitly convey the idea that value creation through multi-sided platforms is where the supply and demand of multiple parties meet. Such platforms are characterized by strong network externalities [16].

The rise of digital platforms has transformed the way creative content is produced, shared, and consumed, opening up new opportunities for content creators to monetize their work. However, despite the rapid growth of digital content, challenges remain in understanding the most effective monetization strategies [21].

The rapid advancement of digital technologies has changed how creative content is produced, distributed, and consumed [25]. Platforms like social media have democratized content creation, enabling individuals to reach global audiences without relying on traditional intermediaries such as publishers or broadcasters [3]. This shift has opened up new possibilities for content creators to earn directly from their work instead of depending on intermediaries or conventional industry structures.

Monetization strategies have become essential for the success of digital content creators, and various models have been developed to meet the needs of both creators and their audiences [9]. Advertising revenue, brand partnerships, subscription models, and crowdfunding are among the most common monetization methods used by digital platforms to enable creators to generate income [6]. Many creators use multiple revenue streams simultaneously, leveraging a wide range of platform tools to maximize their earning potential.

There are various monetization models - the nature of monetization differs across platforms like YouTube, TikTok, Instagram, and Patreon. For example, YouTube relies heavily on ad revenue, whereas TikTok emphasizes brand partnerships and Instagram follows influencer-driven models. Likewise, Patreon's subscription-based model offers a more direct and sustainable form of income, contrasting with the fluctuating nature of ad-based revenue [21].

2.2 Conceptual framework of personalized educational platforms

Personalized learning methods can help meet individual needs and goals. Accordingly, personalized learning can be an effective approach that enhances motivation, engagement, and comprehension [30], maximizing student satisfaction, learning efficiency, and effectiveness [14]. Niknam and Thulasiraman (2020) argue that the education community is interested in developing personalized learning systems that adapt pedagogy, curricula, and learning environments to students' needs and preferences [27]. Schmid and Petko (2019) support this claim, stating that a clearly defined concept of personalized learning still does not exist; instead, it serves as an umbrella term for educational strategies aimed at addressing individual abilities, knowledge, and learner needs [35].

Personalized learning is an educational strategy that tailors instruction to learners' interests, abilities, or needs, and typically involves students having some degree of voice and choice (i.e., autonomy) in the customization process.

Today, schools, universities, and corporate environments possess the technological capabilities to personalize learning in line with students' unique needs. Technology provides numerous options for learners and educators to explore new approaches to personalized learning.

Horn and Staker (2014) proposed a framework for thinking about the dimensions of personalized learning in practice [19]. They suggested that personalization can be achieved by adapting the time, place, pace, and/or learning path. Graham et al. (2019) added a fifth dimension - learning goals [15]. Shewshack et al. (2021) proposed that a unique, evolving personalized learning approach should include four main components: learner profiles, prior knowledge, personalized learning pathways, and flexible self-paced learning environments formed based on dynamic learning analytics (Chatti & Muslim, 2019) [36, 8]. Learning environments that incorporate these various dimensions and components can empower learners to take responsibility for their learning and enhance their self-efficacy.

In his research, Short (2022) argues that the concept of personalized learning explains that although various definitions describe it as adjusting instruction based on a student's background, needs, abilities, or interests, a comprehensive description should include: firstly, what is being personalized - learning goals, assessments, or educational activities; secondly, how personalization is achieved - via goals, time, place, pace, and/or learning path; thirdly, who or what carries out the personalization - the teacher, the student, or an adaptive learning system; and last but not least, on what basis personalization is conducted - performance data, activity data, or learner profile data [37].

Other studies highlight the need to further explore the outcomes of personalized learning initiatives and the hope that technology will fulfill its transformative potential to enable tailored, individualized education [41, 45].

2.3 Monetization strategies for platform business

The continuous advancement of digital platforms has changed the way creative content is produced, shared, and consumed, offering new opportunities for content creators to monetize their work. However, despite the rapid expansion of digital content, there are still challenges in understanding the most effective monetization strategies [21].

Monetization strategies have become key to the success of digital content creators, with various models developed to meet the needs of both creators and their audiences [9]. Advertising revenue, brand partnerships, subscription models, and crowdfunding are among the most common methods used by digital platforms to help creators earn income [6]. Many creators use multiple revenue sources simultaneously, leveraging a variety of platform tools to maximize their income potential.

According to some authors, a combination of audience engagement, consistent content production, and platform-specific strategies (such as ads, subscriptions, and merchandise sales) has been key to successful monetization. Moreover, creators who diversify their income streams are more likely to achieve long-term financial sustainability [21].

Despite the increasing volume of content produced on these platforms, there is still limited understanding of the long-term sustainability of these monetization models [26]. While short-term earnings may be possible, it remains unclear how creators can build sustainable careers over time - especially in the context of constantly changing platform algorithms, audience preferences, and market trends [12].

Although a growing body of literature explores content creation and monetization, the specific factors contributing to long-term financial success remain underexplored [22]. While many studies have examined individual monetization methods, there is a gap in understanding how these strategies work synergistically across different platforms [10]. Creators often use multiple monetization tactics simultaneously, yet little is known about how these various income sources interact and affect financial stability over time.

Additionally, research on how different content types (e.g., educational, entertainment, lifestyle) influence monetization success remains limited [32]. While some content genres may be better suited for specific monetization models (e.g., tutorials for subscription-based services), this relationship has not been systematically studied. There is a need for a more sophisticated, genre-specific analysis to understand which types of content have the highest potential for success across different platforms [44].

The fast pace of technological change means that content creators must continuously adapt to new formats and consumer behaviors, which can significantly impact their earning potential [33].

Finally, the financial sustainability of small and medium-sized creators compared to large influencers has not been sufficiently addressed in academic research [13]. Many studies focus on top-tier creators, while little is known about the challenges faced by new and lesser-known creators who lack the resources or institutional support to succeed on these platforms.

Filling the gaps in understanding content monetization is essential for both creators and platforms [39]. By studying the synergistic effects of different monetization strategies, creators can better navigate the complex digital content ecosystem [31]. The insights gained from this research can help optimize revenue sources and achieve long-term financial sustainability. Moreover, a deeper understanding of how various types and genres of content function within different monetization models can lead to more precisely tailored strategies for creators at all levels [40].

The digital economy, largely driven by content consumption, has enabled creators to turn their passion into a profession, making monetization an essential aspect of content creation. Monetization strategies are now crucial for the sustainability and growth of content creation as a career. Content creators must explore revenue sources beyond ads, which, while accessible, are often unreliable and insufficient on their own [1].

3. RESEARCH DESIGN

3.1 Sample and Instruments

The personalized learning platform that is the subject of this paper can be implemented as IaaS (Infrastructure as a Service), providing scalable and flexible infrastructure on cloud platforms (e.g., AWS, Azure, or Google Cloud, or with other providers offering similar services). Naturally, cost plays a key role.

The approach allows dynamic resource management, security, and easy scalability, along with full functionality for analyzing qualitative data and personalizing educational content.

This solution was chosen primarily because it addresses the critical issue of infrastructure, which requires a significant initial investment. In addition to the aforementioned advantages, an IaaS solution allows for easy expansion of capacity and resources as needed—they can be increased or decreased, and thus costs adjusted, depending on demand. This approach also enables straightforward integration with other cloud services and tools.

For the purpose of conducting the monetization analysis, a survey was conducted. The study employs a qualitative research design, focusing on multi-sided platform-based business models and, accordingly, various monetization strategies available to content creators on digital platforms [42]. The research centers on the potential for successful monetization of work on the platform and examines the factors influencing financial success [43]. The methods of analysis and synthesis are used for the evaluation of projected financial values of the platform monetization.

The potential users are students from various universities, high school students from different schools, as well as employed individuals and professionals in need of personalized professional development.

Table 1: The Interview Respondents (Source: Author)

Respondents	Number of respondents	Percentage
High school students	66	30%
University students	115	52,27%
Employed individuals	39	17,73%
Total:	**220**	**100%**

Therefore, the research sample consisted of 220 respondents, purposefully selected to represent key target groups relevant to the study. The sample included 115 university students from various faculties, 66 high school students from different secondary schools, and 39 employed individuals from various organizations who are in need of professional upskilling (Table 1).

A semi-structured interview questionnaire was used as the primary data collection instrument. This method allowed for a flexible yet focused exploration of the participants perspectives on personalized learning platforms, their needs and expectations, as well as their willingness to pay for such services. The semi-structured format ensured that while a core set of questions was consistently asked across all interviews, participants were also encouraged to elaborate on their willingness experiences and opinions, thus providing rich qualitative insights.

A purposive sampling method was used in order to project monetization success, such as steady revenue from advertisements, sponsorships, or direct user support (Table 2). Purposive sampling is a non-probability sampling technique where the researcher selects participants based on specific characteristics or qualities, ensuring that the sample is relevant to the research objectives [29].

3.2. Results and Discussion

Data collection conducted using semi-structured interviews [18] was developed to cover the key area – the financial outcomes. In that respect, possible monetization strategies, which preceded financial outcomes, were also analyzed. Secondary data from various platforms and online articles were analyzed to complement the primary data collected through interviews (Table 1).

The data collected for this study include metrics related to content demand on the projected platform, such as the exchange of educational and teaching materials, organization of personalized courses, and consultations.

Table 2. The target Groups and Strategies of the Innovative Social Network (Source: Author's analysis based on research data)

Target group	Description of personalization	Strategies
High school students	- Individual lessons - Textbook exchange	- Microtransaction, Freemium[3] - Microtransaction[4]
University students	- Sharing of educational materials (summaries/outlines) - Exchange of educational materials (study notes) - Personalized consultations - Individual lessons / one-on-one sessions	- Freemium, Subscription - Subscription, Microtransaction - Freemium, Microtransaction - Freemium, Microtransaction
Employees and professionals	- Personalized courses based on profession - Technical and IT courses	- B2B model[5] - Subscription, Microtransaction, B2B model

The indicators were projected based on the collected primary data, i.e., semi-structured interviews with respondents from the target groups - as well as secondary data from platforms where advertisements for such services are posted. Based on the compiled dataset, the projected values of total investments, expenses, revenues, and net profit were calculated (Table 3).

The data indicate the projected economic values and the factors that need to be considered in calculating the necessary indicators. Creators on the platform primarily rely on revenue from advertising, subscriptions, and microtransactions, with a significant number of users and interactions contributing to their earnings.

[3] Derived from the word "free" and "premium", freemium is a type of business model where the basic features of a product or service are offered to users at no cost. However, a premium is charged for additional or advanced features.

[4] A microtransaction refers to the ability to spend real money on a platform in exchange for virtual currency that is used solely within that platform.

[5] The B2B model (Business-to-Business) refers to a business model in which a digital platform is used to connect two or more business entities. In this model, companies trade services, knowledge, and technologies with one another.

Table 3. Projected Investments, Revenues, Expenses, and Net Profit (Source: Author's analysis based on research data)

Indicators	Values
INITIAL INVESTMENT (CAPEX)	
- Software and Platform development	825.000 RSD
- Infrastructures + initial licenses	275.000 RSD
- Content development	825.000 RSD
- (UI/UX)	412.500 RSD
- Initial Legal and Administrative Costs	137.500 RSD
- Launch Marketing Campaign	275.000 RSD
Initial Investment Total:	**2.750.000 RSD**
EXPENCES	
Operational Expenditure (OPEX):	
- Infrastructure (IaaS hosting)	810.000 RSD
- Maintenance and IT Administration	810.000 RSD
- Marketing	1.620.000 RSD
- Content creation and Development & outsourcing	1.080.000 RSD
- Customer support and tutoring	378.000 RSD
- Operational Financial Cost	270.000 RSD
- Administrative Costs (G&A) + Tax	432.000 RSD
Operational Expenditure Total:	**5.400.000 RSD**
REVENUE	**6.472.500 RSD**
ESTIMATED NET PROFIT	**1.072.000 RSD**
PROJECTED PROFIT RETENTION	**50.000 RSD**

The dataset from the ratio analysis reveals an efficiency ratio calculated as the relationship between achieved effects (results) and resources spent, which is greater than 1. This means the platform is economical, generating 11.09 RSD in revenue for every RSD spent. Also, the analysis of the accumulation rate of 4.7% shows that a portion of the profit that is not spent but reinvested into the further development of the platform amounts to 4.7% as the percentage of profit retained for accumulation.

By analyzing the projected values, the PBP (Payback Period) can be observed as 2.57 years, which is the time needed for the initial investment to be covered through generated profit or savings, meaning that monetization of the platform will cover the entire investment within 2.57 years. The ROI (Return on Investment) calculation indicates how profitable the investment was, i.e., the profit achieved relative to the invested capital. Since the ROI is positive at 39%, the investment yields profit (Table 4). The justification for the investment is reflected in the ROI. The investment is justified and sustainable.

Table 4. Ratio Analysis (Source: Author's analysis based on research data)

Indicator	Calculation	Value
Efficiency ratio	Total revenue/Total expenses	**11.9%**
Accumulation rate	Retention/Net profit * 100	**4.7%**
Payback Period (PBP)	Total investments/Net income	**~ 2.57 years**
Return on Investment (ROI)	Net income /Investments * 100	**39%**

Monetizing this platform for personalized learning, which is implemented as IaaS, represents both a key challenge and an opportunity for sustainability and further development of digital education. Since personalized learning is based on tailoring content, pace, and methods to learners, the revenue model must be carefully designed to simultaneously support inclusiveness, quality, and commercial interests.

Launching an innovative platform enables (Table 2):

- Personalized lessons
- Exchange of teaching and educational materials (scripts, textbooks, etc.)
- Sharing of educational materials (summaries/outlines)
- Personalized consultations
- Personalized courses based on professions
- Technical and IT courses
- B2B model.

One of the strategies is the freemium model, where basic functionalities remain free for all users, while advanced features - such as access to mentors or specialized courses - are charged through subscriptions. This model ensures wide accessibility while creating stable revenue through loyal users. The freemium concept involves a combined approach where users can choose between using a free version of the product with basic features or a premium version for a fee [28] (Table 2).

Another approach is the B2B model, where the platform collaborates with educational institutions or companies, offering integration of the system into their internal training and educational processes. In this context, monetization is achieved through contracts, training, or content tailored to specific user needs. The literature extensively examines how the B2B model affects content revenue on platforms [23] (Table 2).

Additional revenue sources include advertising and microtransactions, where users can purchase individual lessons, tests, consultations, and classes. This model offers greater flexibility and accessibility, especially for groups with limited purchasing power, allowing users to pay selectively.

Subscription services - models more directly focused on users - also provide creators with more reliable and diversified income sources that are less vulnerable to platform changes. Many creators today capitalize on their loyal audience by offering exclusive content through subscription services [11] (Table 2).

According to the research data, the projections indicate that the number of users will increase by 50% after the first year, and by 100% each subsequent year thereafter. This means that ROI will increase proportionally with the growth of the number of users. The reason for this is that IaaS stems from scalability as the main characteristics of this type of platforms – as the number of users increases, the total revenue grows as well.

4. CONCLUSION

The analysis of the personalized learning platform demonstrates that the investment is financially viable. Both the ROI (39%) and the PBP (~2.57 years) indicate that the investment is sustainable, with great potential for success. In that respect, the projected analysis of investments, revenues, and profits (Tables 3 and 4), indicate that monetization of an innovative multi-sided platform for personalized learning is feasible, but requires a carefully designed implementation strategy in practice.

Primarily, the application of a freemium business model is recommended, which would allow the platform's basic functionalities - such as the exchange of teaching materials and access to general courses - to be available to all users free of charge, while advanced services (individual consultations, mentoring programs, specialized courses) would be monetized through subscriptions and microtransactions. This approach enables both inclusivity and revenue generation, which is especially important during the initial development phase when attracting a wide user base is critical.

To stabilize revenue and expand the market, the development of B2B collaborations with educational institutions and corporations is recommended, through which the platform could be integrated into internal education and training systems. This approach may result in contractual revenues that ensure long-term business sustainability and strengthen institutional support.

As the key factor contributing to this profitability is the use of Infrastructure as a Service (IaaS), which ensures scalability, cost optimization, and flexibility in adapting to user demand. Consequently, the combination of solid financial indicators and scalable infrastructure confirms the sustainability and long-term potential of this platform.

Considering that the purpose of this paper is to demonstrate that a personalized learning platform can be successfully monetized, a detailed presentation of the IT solution will be the subject of further discussion in one of the upcoming scientific papers.

References

[1] Adewumni, O. (2024). Monetization strategies for content creators. *IOSR Journal of Economics and Finance (IOSR-JEF)*, *15*(6, Ser. 5), 57–66. https://www.iosrjournals.org

[2] Afuah, A., & Tucci, C.I. (2000). Internet Business Models and Strategies: Tet and Cases. McGraw-Hill Higher Education, New York.

[3] Ardolino, M., Saccani, N., Adrodegari, F., & Perona, M. (2020). A business model framework to characterize digital multisided platforms. *Journal of Open Innovation: Technology, Market, and Complexity, 6*(1), 10. https://doi.org/10.3390/joitmc6010010

[4] Baden-Fuller, C., & Heafliger, S. (2013). Business models and technological innovations. *Long Range Planning,* 46(6), 419–426.
[5] Baden-Fuller, C., & Mangematin, V. (2013). Business models: A challenging agenda. *Strategic Organization, 11*(4), 418–427.
[6] Baltas, H., Sirin, M., Gökbayrak, E., & Ozcelik, A. E. (2020). A case study on pollution and a human health risk assessment of heavy metals in agricultural soils around Sinop Province, Turkey. *Chemosphere, 241*, 125015.
https://doi.org/10.1016/j.chemosphere.2019.125015
[7] Cennamo, C., & Santaló, J. (2015). How to avoid platform traps. *MIT Sloan Management Review, 57*(1), 12–15.
[8] Chatti, M. A., & Muslim, A. (2019). The PERLA framework: Blending personalization and learning analytics. *International Review of Research in Open and Distance Learning, 20*(1), 244–261. https://doi.org/10.19173/irrodl.v20i1.3936
[9] Cheng, K.-H., & Tsai, C.-C. (2019). A case study of immersive virtual field trips in an elementary classroom: Students' learning experience and teacher-student interaction behaviors. *Computers & Education, 140*, 103600.
https://doi.org/10.1016/j.compedu.2019.103600
[10] Da Silva, E. K. N., Dos Santos, V. B., Resque, I. S., Neves, C. A., Moreira, S. G. C., Franco, M. D. O. K., & Suarez, W. T. (2020). A fluorescence digital image-based method using a 3D-printed platform and a UV-LED chamber made of polyacid lactic for quinine quantification in beverages. *Microchemical Journal, 157*, 104986. https://doi.org/10.1016/j.microc.2020.104986
[11] Debutify. (2021). *How content creators are making a living in the creator economy*. https://debutify.com/blog/what-is-creator-economy
[12] Feng, W., Zhang, Q., Ji, H., Wang, R., Zhou, N., Ye, Q., Hao, B., Li, Y., Luo, D., & Lau, S. S. Y. (2019). A review of net zero energy buildings in hot and humid climates: Experience learned from 34 case study buildings. *Renewable and Sustainable Energy Reviews, 114*, 109303. https://doi.org/10.1016/j.rser.2019.109303
[13] W. Gao, P. Veeresha, H.M. Baskonus, D.G. Prakasha, P. Kumar, A new study of unreported cases of 2019-nCOV epidemic outbreaks, Chaos, Solitons & Fractals 138 (2020) 109929, https://doi.org/10.1016/j.chaos.2020.109929
[14] Gómez, S., Zerva, S. P., Sampson, D. G., & Fabregat, R. (2014). Context-aware adaptive and personalized mobile learning delivery supported by UoLmP. *Journal of King Saud University – Computer and Information Sciences, 26*(1), 47–61. https://doi.org/10.1016/j.jksuci.2013.10.008
[15] Graham, C. R., Borup, J., Short, C. R., & Archambault, L. (2019). *K-12 blended teaching: A guide to personalized learning and online integration*. Provo, UT: EdTech Books. http://edtechbooks.org/K12blended
[16] Hagiu, A. (2007). Merchant or two-sided platforms? *Review of Network Economics, 6*(2), 115–133.
[17] Hagiu, A., & Wright, J. (2015). Marketplace or reseller? *Management Science, 61*(1), 184–203.
[18] Harvey Arce, N. P., & Cuadros Valdivia, A. M. (2020). Adapting competitiveness and gamification to a digital platform for foreign language learning. *International Journal of Emerging Technologies in Learning (iJET), 15*(20), 194–209. https://doi.org/10.3991/ijet.v15i20.16135

[19] Horn, M. B., & Staker, H. (2014). *Blended: Using disruptive innovation to improve schools*. Jossey-Bass.
[20] Johnson, M. W., Christensen, C. M., & Kagermann, H. (2008). Reinventing your business ecosystems: Evidence from application software developers in the iOS and Android smartphone ecosystems. *Organization Science, 28*(3), 531–551.
[21] Kadeni, K., Santoso, E., & Jing, W. (2025). Creative content monetization: Case studies on digital platforms. *Journal of Social Entrepreneurship and Creative Technology, 2*(2), 81–91. https://doi.org/10.70177/jseact.v2i2.2059
[22] Kannan, K., & Arunachalam, N. (2019). A digital twin for grinding wheel: An information sharing platform for sustainable grinding process. *Journal of Manufacturing Science and Engineering, 141*(2), 021015. https://doi.org/10.1115/1.4042076
[23] Kohtamäki, M., Parida, V., Patel, P. C., & Gabauer, H. (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. *Technological Forecasting and Social Change, 151*, 119804.
[24] McIntyre, D. P., & Srinivasan, A. (2017). Networks, platforms and strategy: Emerging views and next steps. *Strategic Management Journal, 38*(1), 141–160.
[25] Mourtzis, D., Angelopoulos, J., & Panopoulos, N. (2021). A survey of digital B2B platforms and marketplace for purchasing industrial product service systems: A conceptual framework. *Procedia CIRP, 97*, 331–336. https://doi.org/10.1016/j.procir.2020.05.246
[26] Nam, K., Hwangbo, S., & Yoo, C. (2020). A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renewable and Sustainable Energy Reviews, 122*, 109725. https://doi.org/10.1016/j.rser.2020.109725
[27] Niknam, M., & Thulasiraman, P. (2020). LPR: A bio-inspired intelligent learning path recommendation system on meaningful learning theory. *Education and Information Technologies, 25*, 3797–3919. https://doi.org/10.1007/s10639-020-10133-3
[28] Osipov, I. V., Nikulchev, E., Plokhou, D., & Volinsky, A. A. (2015). Study of monetization as a way of motivating freemium service users. *Contemporary Engineering Sciences, 8(20)*, 911–918. https://doi.org/10.12988/ces2015.57212
[29] Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y
[30] Pontual Falcão, T., Peres, F. M. A. e., Sales de Morais, D. C., & da Silva Oliveira, G. (2018). Participatory methodologies to promote student engagement in the development of educational digital games. *Computers & Education, 116*, 161–175. https://doi.org/10.1016/j.compedu.2017.09.006
[31] Porter, B., & Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability, 12*(24), 10243. https://doi.org/10.3390/su122410243
[32] Sakamiya, M., Fang, Y., Mo, X., Shen, J., & Zhang, T. (2020). A heart-on-a-chip platform for online monitoring of contractile behavior via digital image processing and piezoelectric sensing technique. *Medical Engineering & Physics, 75*, 36–44. https://doi.org/10.1016/j.medengphy.2019.10.001

[33] Samui, P., Mondal, J., & Khajanchi, S. (2020). A mathematical model for COVID-19 transmission dynamics with a case study of India. *Chaos, Solitons & Fractals, 140*, 110173. https://doi.org/10.1016/j.chaos.2020.110173
[34] Schamshack, A., & Spector, J. M. (2020). A systematic literature review of personalized learning terms. *Smart Learning Environments, 7*(33).
[35] Schmid, R., & Petko, D. (2019). Does the use of educational technology in personalized learning environments correlate with self-reported digital skills and beliefs of secondary-school students? *Computers & Education, 136*, 75–86. https://doi.org/10.1016/j.compedu.2019.03.006
[36] Shewshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education, 1*(19). https://doi.org/10.1007/s40692-021-00188-7
[37] Short, L. C., & Shemshack, A. (n.d.). Personalized learning. *EDTECHNICA*. https://doi.org/10.59668/371.11067
[38] Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning, 43*(2–3), 172–194.
[39] Tien Bui, D., Hoang, N.-D., Martínez-Álvarez, F., Ngo, P.-T. T., Hoa, P. V., Pham, T. D., Samui, P., & Costache, R. (2020). A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *Science of The Total Environment, 701*, 134413.
https://doi.org/10.1016/j.scitotenv.2019.134413
[40] Wang, Q., & Su, M. (2020). A preliminary assessment of the impact of COVID-19 on environment – A case study of China. *Science of The Total Environment, 728*, 138915. https://doi.org/10.1016/j.scitotenv.2020.138915
[41] Watters, A. (2023). *Teaching machines: The history of personalized learning*. The MIT Press.
[42] Wu, F., Wang, Y., Leung, J. Y. S., Huang, W., Zeng, J., Tang, Y., Chen, J., Shi, A., Yu, X., Xu, X., Zhang, H., & Cao, L. (2020). Accumulation of microplastics in typical commercial aquatic species: A case study at a productive aquaculture site in China. *Science of The Total Environment, 708*, 135432. https://doi.org/10.1016/j.scitotenv.2019.135432
[43] Xiao, R., Guo, D., Ali, A., Mi, S., Liu, T., Ren, C., Li, R., & Zhang, Z. (2019). Accumulation, ecological-health risks assessment, and source apportionment of heavy metals in paddy soils: A case study in Hanzhong, Shaanxi, China. *Environmental Pollution, 248*, 349–357. https://doi.org/10.1016/j.envpol.2019.02.045
[44] Xu, X., Zhang, Q., Song, J., Ruan, Q., Ruan, W., Chen, Y., Yang, J., Zhang, X., Song, Y., Zhu, Z., & Yang, C. (2020). A highly sensitive, accurate, and automated single-cell RNA sequencing platform with digital microfluidics. *Analytical Chemistry, 92*(12), 8599–8606. https://doi.org/10.1021/acs.analchem.0c01613
[45] Zhang, L., Basham, J. D., & Yng, S. (2020). Understanding the implementation of personalized learning: A research synthesis. *Educational Research Review*, 31, 100339. https://doi.org/10.1016/j.edurev.2020.100339

A Scalable Microservices-Based Architecture for E-learning Platforms

[1]Pedro Guedes Monteiro and [2]Paulo Tomé

[1] Polytechnic of Viseu, Portugal
[2] CISeD - Research Centre in Digital Services, Polytechnic of Viseu, Portugal

[1]pv20272@alunos.estgv.ipv.pt, [2]ptome@estgv.ipv.pt

Abstract. E-learning platforms have become critical infrastructure for knowledge dissemination, flexible scheduling, and lifelong upskilling. Their relevance was underscored during the COVID-19 pandemic, which exposed recurring gaps in widely used systems: brittle scalability under surges, uneven customization across institutional contexts, fragmented support workflows, and limited, siloed analytics. These constraints hindered timely interventions, raised operational costs, and complicated governance and maintenance as deployments grew and heterogeneity.
This paper proposes a scalable architecture for e-learning that combines a microservices-based LMS with an integrated knowledge hub and gamification. The design supports institutional customization, observability, and elastic scaling. A comparative analysis with six leading platforms highlights the architectural advantages and addresses scalability and support gaps exposed during the pandemic.
To evaluate the relevance and advantages of this architecture, an integrated Artificial Intelligence layer provides conversational support and data-driven recommendations grounded in usage and performance signals. We complement the architectural description with a comparative analysis against six established platforms, using a set of evaluation variables defined in the methodology (architecture, licensing, implementation type, scalability, customization, ecosystem, integration/extensibility, and gamification), showing how the proposed approach addresses pandemic-revealed gaps while delivering SaaS-like elasticity without forfeiting institutional control.

Keywords: E-learning; Knowledge Hub; Gamification; Artificial Intelligence; Microservices; Digital Platform.

1 Introduction

The COVID-19 pandemic exposed critical limitations in existing e-learning platforms. Critical issues, as pointed in [1], include technological limitations, lack of adequate technical support, ineffective change management and cultural barriers. These gaps highlighted an urgent need for more effective and adaptable solutions.

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.105GM
Part of ISBN: 978-86-89755-37-4

In response to these challenges, this led to the conception of a platform designed to combine e-learning features with an emerging knowledge center concept. This knowledge center refers to a collaborative digital space that integrates functionalities for creating, sharing and managing educational content. In contrast to traditional platforms, this implementation promotes an interactive and dynamic environment where users can consume information but also actively contribute to its enrichment.

To overcome challenges such as lack of interaction and inadequate support, the knowledge center includes features such as opening technical and pedagogical support tickets for quick resolution of problems, continuous interaction between users through comments and collaborative forums and direct publication of content by users, an approach aligned with evidence showing that user-generated content enhances collaborative knowledge building in online education [24].

The adoption of modern architectures, such as microservices supported by Docker and Fastify, has proved crucial to ensuring scalability and efficiency on large-scale platforms, as advocated by the literature [28]. According to these authors, microservice-based architectures, when combined with lightweight, high-performance tools such as Fastify, promote greater modularity and autonomy between services, allowing different modules to operate independently, which facilitates maintenance, incremental updates and the expansion of the platform as new needs arise.

This article is structured into sections that follow a logical sequence to ensure a clear and structured understanding of the platform being proposed. In the **Introduction section**, the context of the project is outlined, along with the motivation for its development.This section also defines the problem to be addressed and presents the main objectives established for the work. The **Methodology section** details the methods and procedures adopted for the platform's development. It covers the technical approaches, tools to be used and methodological strategies that support the implementation process. In the **Comparative Analysis section**, a range of current e-learning solutions available on the market is explored and analyzed. The goal is to understand their architectures, functionalities and and underlying technological choices. The **Proposed Architecture section** explains how the selected methodologies are to be implemented in an integrated and efficient manner. It describes the roles of each component within the platform and how they interact with one another. Within the **Comparison with Existing Solutions section**, the proposed platform is contrasted with the previously analyzed systems. This comparison highlights both the strengths and the potential limitations of the proposed solution. Finally, the **Conclusions and Future Work section** summarizes the key contributions and findings of the project. It highlights the strengths of the proposed architecture and suggests future directions that could enhance the platform's impact and capabilities.

2 Problem Statement

Despite the progress that has been made in the area of digital education, [1] highlight that there are still persistent shortcomings, such as the lack of interaction between users, the absence of adequate technical support and the lack of adapted pedagogical approaches, problems which, when analyzed as a whole, represent a major challenge that was highlighted in several of the platforms analyzed, which is the lack of integration of robust technologies and collaborative approaches. This obstacle directly compromises the effectiveness of e-learning platforms, since there is a restriction on offering personalized educational experiences aligned with users' needs.

Given the scenario described in the previous paragraph, the following central question arises: How can we develop an e-learning platform that is scalable, comprehensive and integrates a robust knowledge center in order to promote personalized and collaborative learning?

To meet the challenges of the issue raised, the proposed platform will rely on technologies capable of enabling user-generated content, but will also provide reactive technical support systems to help students face the challenges presented to them. This need has become especially evident during the COVID-19 pandemic, when many institutions found it difficult to effectively support students due to the lack of integrated technical assistance, as noted by the authors [15]. In addition, [26], in their article, demonstrate that the use of gamified elements has proven to be effective in improving student engagement and motivation in digital environments. Finally, to provide personalized learning experiences, the integration of Artificial Intelligence is increasingly seen as essential, allowing platforms to adapt content and resources to the individual needs of each student.

3 Methodology

In order to understand the current panorama of e-learning platforms, we carried out a comparative analysis of six solutions on the market. The aim of this analysis is to identify the architectural approaches adopted, the licensing models used and the functionalities that each platform offers. The following platforms were chosen for analysis: Moodle, Canvas, Open edX , Coursera, Udemy and Sakai.

As part of the comparative analysis of the Moodle, Canvas, Open edX, Coursera, Udemy and Sakai e-learning platforms, a set of variables were defined that were considered fundamental for a comprehensive and rigorous assessment of the solutions on the market. The purpose of selecting these variables was to ensure a balanced analysis of the technical, functional and strategic aspects of each platform. The variables analyzed can be seen in the table 1.

The selection of these variables was based on the need for a critical and complete analysis, which is not limited to the functionalities offered, but also considers technical sustainability, operational flexibility and user experience.

The comparison between the platforms was based on various criteria considered fundamental for evaluating e-learning solutions. In terms of *architecture*, the

Table 1: Description of Comparative Evaluation Variables

Variable	Description
Architecture	Refers to the technical foundation of the platform, including client-server models, layer separation, and development technologies. This variable assesses the platform's internal organization, robustness, and long-term sustainability.
Licensing Model	Distinguishes between open-source and proprietary solutions, evaluating implications for customization, cost, and community support, which directly affect adoption feasibility.
Deployment Type	Analyzes whether the platform uses a monolithic or microservices-based architecture, influencing scalability, maintainability, and modular updates.
Scalability	Evaluates the platform's capacity to support increasing user volumes and activities without performance degradation—crucial for institutional growth.
Customization	Assesses flexibility in adapting the interface, workflows, and features to institutional or pedagogical needs, ensuring contextual alignment.
Community and Ecosystem	Considers project vitality, including forums, documentation, third-party extensions, and development activity—indicators of maturity and sustainability.
Integration and Extensibility	Measures the ability to connect with external tools and systems via Application Programming Interfaces (APIs), LTI support, and interoperability standards.
Gamification	Analyzes the presence of motivational elements such as badges, levels, and challenges, reflecting the platform's potential for user engagement.

technical foundations of each platform were analyzed, such as the client-server model, the separation of layers and the technologies used in development.

With regard to *licensing*, an attempt was made to distinguish between open source solutions and proprietary platforms, reflecting on the implications of this distinction in terms of flexibility, costs and community. The *deployment dimension* was addressed by identifying monolithic or microservice-based architectures, considering their impact on scalability, maintenance and ease of updating.

Scalability, in turn, was assessed based on the platforms' ability to support a growing number of users, courses and interactions without compromising performance. With regard to *customization*, the aim was to understand the extent to which each platform allows the interface, functionalities and flows to be adapted to the specific needs of each organization or educational context.

The *community* and *ecosystem* analysis focused on the vitality of the project around the platform, including the existence of support forums, regular contributions, comprehensive documentation and extensions developed by third parties. With regard to *integration* and *extensibility*, compatibility with other tools and services, support for standards such as LTI (Learning Tools Interoperability) and the existence of APIs to expand functionality were considered.

Finally, the degree of *gamification* offered by each platform was also assessed, i.e. the presence of elements that promote user involvement through mechanisms such as badges, progression levels, rankings and motivational challenges.

4 Comparative Analysis

To establish a comprehensive understanding of the current landscape, we analyzed six prominent e-learning platforms widely adopted in both academic and corporate contexts. Each of these platforms demonstrates distinct architectural patterns, integration capabilities, and educational approaches. The following sections detail the characteristics of each platform, highlighting their strengths, limitations, and relevance to the development of a modern, scalable learning environment.

4.1 Moodle

Moodle is a widely adopted Learning Management System (LMS) globally and, according to its official documentation, [17], is licensed under the GNU General Public License v3 (GPLv3). This license allows the use, modification and redistribution of the source code, promoting a strong international development community .

On a technological level, the platform is mostly developed in PHP, with additional support for web technologies such as JavaScript, HTML, CSS and SQL, as described in [17]. This base ensures compatibility with web services (Apache HTTP Server, Nginx, IIS) and with database management systems such as MySQL, PostgreSQL, MariaDB, Oracle Database and Microsoft SQL Server (MSSQL).

Its modular architecture, as detailed in the official repository [18], allows extensive customization through plugins and themes. In addition, offers RESTful APIs and web services that enable integrations with external systems, such as academic information systems, digital libraries and videoconferencing tools [17].

With regard to the interface, Moodle uses the Boost theme, based on Bootstrap, guaranteeing responsiveness and accessibility on various devices [17]. In terms of security and authentication, it supports protocols such as LDAP, OAuth2, SAML and Shibboleth, allowing integration with federated identity infrastructures [17].

4.2 Canvas LMS

Canvas is an LMS developed by the US company Instructure and, according to its official documentation, [12], offers an open-source version under the Affero GNU General Public License v3 (AGPLv3). This license allows the use, modification and redistribution of the source code, requiring the sharing of modifications even in Software as a Service (SaaS) environments, promoting collaboration and transparency.

On a technical level, the platform is mostly developed in Ruby on Rails, complemented with JavaScript, HTML, CSS and SQL, as described in [12]. It supports web servers such as Apache HTTP Server and Nginx, and uses PostgreSQL as its database management system [13].

Its modular architecture allows extensions and integrations via RESTful APIs and support for the LTI standard, making it easy to connect to various external tools, [12]. It also provides webhooks and endpoints for real-time data synchronization [13].

In terms of the interface, Canvas adopts a responsive and accessible design, aligned with the Web Content Accessibility Guidelines (WCAG) 2.1, ensuring compatibility with multiple devices and promoting inclusion [12]. In terms of security and authentication, it supports protocols such as SAML, LDAP and OAuth2, ensuring integration with complex identity infrastructures [12].

4.3 Open EdX

Open edX is an open source LMS initially developed by the Massachusetts Institute of Technology (MIT) and Harvard University, and currently maintained by the Open edX community and the company tCRIL [19]. It is licensed under the Apache License 2.0, allowing wide freedom of use, modification and redistribution of the source code.

From a technological point of view, it is developed mostly in Python, using the Django framework, as well as JavaScript, HTML, CSS and SQL [19]. It supports web servers such as Nginx and Apache HTTP Server and uses databases such as MySQL and MongoDB [20].

The architecture is based on microservices and independent modules, allowing for higher scalability and better customization. It offers RESTful APIs, LTI support and integration with external tools via Open edX plugins [19]. It also offers webhooks for synchronizing data with external systems [20].

In terms of interface, Open edX adopts a responsive design, in line with the WCAG 2.1 guidelines, ensuring accessibility and compatibility with various devices [19]. In terms of security, it supports protocols such as OAuth2 and SAML, guaranteeing integration with institutional systems [19].

4.4 Coursera

Coursera is a SaaS-based commercial LMS, founded in 2012 by professors from Stanford University with the aim of making Massive Open Online Courses (MOOCs) available globally [5]. As a proprietary platform, it does not make its source code available, but operates under institutional and corporate licensing.

In terms of technology, it is developed with a distributed architecture based on microservices [4]. It uses languages and technologies such as Java, Python, JavaScript, HTML, CSS and databases such as MySQL, MongoDB and Cassandra, operating essentially on public cloud computing infrastructures (Amazon Web Services (AWS)).

Its integration with external systems is ensured through RESTful APIs and support for LTI, in addition to providing Software Development Kit (SDK)s and webhooks for educational partners and institutional platforms [6].

In terms of interface, Coursera adopts a responsive and accessible design, in line with WCAG 2.1 guidelines [5]. Security and identity management are supported by protocols such as OAuth2 and SAML, ensuring integration with federated authentication infrastructures [5].

4.5 Udemy

Founded in 2010, Udemy, [34], is an online course platform that operates under a closed commercial model. Unlike open-source platforms, it does not make its source code available and access is regulated by individual subscription or via Udemy Business, [32], aimed at organizations.

In technical terms, its infrastructure is based on a distributed microservices architecture [33]. The platform uses Python, React, JavaScript, HTML, CSS and databases such as MySQL and Redis, operating in Cloud Computing environments supported by AWS.

Although it does not offer an extensive public API, enterprise integrations are possible through authentication solutions such as Single Sign-on (SSO) and partial support for the LTI standard [32]. The platform favours stability and scalability, but with less scope for institutional customization.

Udemy's interface is responsive and user experience-oriented, meeting the criteria defined by WCAG 2.1 and security is ensured by protocols such as SAML and OAuth2, facilitating integration with corporate directories and external LMS platforms [34].

4.6 Sakai

Sakai is an open-source LMS, primarily oriented towards higher education, and is developed and maintained by the Apereo Foundation [2]. The platform is made available under the Educational Community License 2.0 (ECL 2.0), a variant of the Apache License 2.0, allowing wide reuse and modification of the code, in line with free software principles.

From a technical point of view, it is built in Java and follows a modular architecture based on independent components [3]. It uses application servers such as Tomcat and database systems such as MySQL, Oracle Database and PostgreSQL. The user interface is developed using web technologies such as

JavaScript, HTML and CSS, and the platform supports integration with LDAP and external services via RESTful APIs.

Sakai's extensibility is one of its main features: it supports the LTI standard for integrating external tools and has a system of built-in tools that can be activated, customized or removed according to the institution's profile. Additional web services allow interoperability with digital libraries and academic management systems [2].

The interface design is responsive and in line with the good accessibility practices established by WCAG 2.1, allowing fluid use on different devices [2]. Federated authentication mechanisms such as SAML, LDAP and OAuth2 are also supported, ensuring security and compatibility with complex institutional infrastructures.

5 Proposed Architecture

The proposed architecture aims to provide a robust, scalable, and modular e-learning platform capable of adapting to growing user demands without compromising performance or maintainability.

The platform is optimized for responsiveness and stability, even under peak load. The technology stack was selected based on maturity, community support, and alignment with agile and secure development practices. In parallel, the architecture supports extensive customization—both visual and functional—to meet the diverse needs of target user groups and institutional contexts.

Designed to address gaps in existing e-learning platforms, the solution combines scalability, resilience, and personalization with a collaborative knowledge center. The platform enables individualized experiences, supports real-time pedagogical intervention, and sustains engagement through gamification, all without compromising technical robustness. A modular approach ensures long-term adaptability, allowing for the system's evolution without full rewrites [16].

At its core, the platform adopts a microservices architecture, isolating responsibilities into independent services packaged in Docker containers [7] and orchestrated by Kubernetes [31]. This enables elastic scaling, fault tolerance, and controlled deployment cycles. Horizontal Pod Autoscaler [30] and dynamic load balancing ensure resource alignment with demand, while container isolation improves environmental consistency. This setup allows intensive components like analytics or content delivery to scale independently.

For the user interface, Next.js [35] was chosen, combining Server-Side Rendering (SSR) and Static Site Generation (SSG) [36], enhancing response times, indexing, and usability on constrained devices. Its automatic routing and reusable components facilitate rapid development while preserving visual and semantic coherence. Documentation confirms that SSR and SSG contribute both to performance and SEO, reinforcing its relevance in educational platforms.

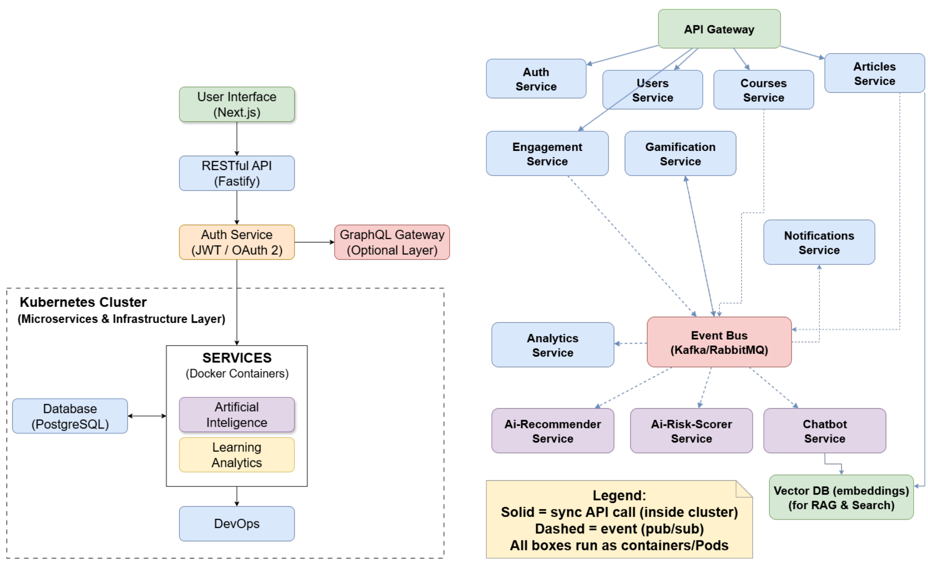
The services layer is implemented in Fastify [9], a lightweight Node.js framework with native JSON Schema validation. It enables the creation of secure, performant RESTful APIs in TypeScript. Benchmarks [29] highlight its performance advantage over other Node.js frameworks, making it ideal for large-scale platforms. Its plugin system also supports modular additions (e.g., authentication, rate limiting) without bloating the core.

PostgreSQL [21] was selected as the database for its ACID compliance, performance optimizations, and backup features, ensuring integrity and availability in high-throughput scenarios. For security, JWT [14] and OAuth2 [11] enable distributed authentication and authorization across services. GraphQL is optionally used for efficient data aggregation when REST endpoints are insufficient [10]. General Data Protection Regulation (GDPR) principles [8] guide personal data handling, emphasizing auditable logging and data minimization.

Artificial Intelligence and Learning Analytics modules analyze behavioral patterns to deliver personalized recommendations, identify at-risk learners, and support chatbot interactions [23]. Gamification reinforces engagement and motivation, with design guided by ongoing evaluation [25].

The platform's operation is sustained by DevOps practices, including CI/CD pipelines, automated testing, monitoring, and distributed tracing [27]. Observability is a key pillar in microservice ecosystems, supported through central metrics, logs, and correlation mechanisms [22]. Complexity is mitigated through automation and standardization, in line with microservices best practices [16].

Figure 1a presents an overview of the adopted architecture.



(a) Adopted Architecture (b) Microservices Communication

Fig. 1: Overview of the proposed platform: (a) adopted architecture and (b) microservices communication.

5.1 Customization, Ecosystem, Integration/Extensibility, and Gamification

Ecosystem maturity varies markedly. Moodle and Canvas provide rich plug-in/LTI catalogs and public APIs; Sakai offers tool activation and LTI; Open-EdX supports modular extensions and webhooks; Coursera and Udemy expose narrower partner or enterprise interfaces. These differences shape how institutions tailor assessment, content, and analytics pipelines.

The proposed architecture presented in this paper, emphasizes contract-first composition (versioned REST, optional GraphQL) rather than in-process plug-ins, trading breadth-at-day-one for safer upgrades and clearer failure boundaries. Integration with SSO (OAuth2/SAML) and support for standards such as LTI sustain interoperability while keeping authorization and data minimization consistent across services.

Finally, the knowledge-center layer and an event-driven analytics/gamification stack speak to learner engagement and timely support: user-generated resources, ticketed assistance, and telemetry-based recommendations are integrated as first-class services; gamified elements motivate participation while remaining decoupled from core grading flows, in line with evidence that such mechanics can enhance engagement when carefully designed.

Taken together, the microservices approach delivers elasticity and operational visibility comparable to commercial SaaS, yet preserves institutional control over cadence, data, and customization. It complements—rather than merely replicates—the strengths of open-source incumbents by localizing change, aligning resource usage with real demand, and foregrounding support workflows, analytics, and motivation as integral, horizontally implemented capabilities. These outcomes map directly to the comparison variables defined in the methodology and to the shortcomings surfaced during the pandemic period.

The communication between the microservices that work together in Kubernetes can be seen in the figure 1b. There, synchronous interactions are terminated at the API Gateway, which applies JWT/RBAC and routes requests to core microservices (Auth, Users, Courses, Articles, Engagement, Gamification, Notifications, Analytics). Business events—such as "Enrollment Created", "Progress Updated", "Course Completed", and "Article Published" are published to the Event Bus (e.g., Kafka/RabbitMQ), where downstream subscribers consume them without tight coupling.

Gamification, Notifications, and Analytics process these events to award badges, trigger messages, and maintain metrics/warehouse views. AI services subscribe to the same streams: the Recommender updates personalized suggestions, the Risk Scorer flags at-risk learners, and the Chatbot performs RAG over a Vector DB (embeddings) for semantic retrieval. Solid arrows represent synchronous API calls; dashed arrows represent asynchronous pub/sub flows between containers/Pods.

Table 2: Comparative summary of the analyzed e-learning platforms

Criterion	**Moodle**	**Canvas**	**Open edX**	**Coursera**	**Udemy**	**Sakai**
Licensing	GPLv3	AGPLv3	Apache 2.0	Proprietary	Proprietary	ECL 2.0
Open source	✓	✓	✓	×	×	✓
Commercial model	Community (third-party SaaS available)	Hybrid (open-source & SaaS)	Community (with providers)	Closed SaaS	Marketplace / SaaS	Community
Base technology	PHP (+ JS)	Ruby on Rails (+ PostgreSQL)	Python/Django (+ MySQL/MongoDB)	Cloud microservices	Cloud microservices	Java (Tomcat)
Implementation type	Modular monolith	Cohesive web app with extensions	Service-oriented modules	SaaS (provider-managed)	SaaS (provider-managed)	Modular (cohesive deployment)
Scalability	Medium (tuning/cluster dependent)	Good (deployment-dependent)	Very good / Excellent	Excellent (SaaS)	Excellent (SaaS)	Good
Customization	High (plugins)	Good (LTI/APIs)	Good (XBlocks/APIs)	Limited (vendor-controlled)	Very limited (course/enterprise)	High (built-in tools/LTI)
LTI/API integration	✓	✓	✓	✓	Partial (enterprise)	✓
Primary focus	Formal education (multiple levels)	Higher education	University-scale MOOCs	Global MOOCs	Open marketplace / enterprise	Higher education
Installation complexity	Medium	Medium–High	High	SaaS (ready to use)	SaaS (ready to use)	High
Accessibility support	✓	✓	✓	✓	✓	✓
Active community	Very active	Active	Active	—	—	Smaller

6 Conclusions and future work

The proposed architecture offers a modular and scalable foundation for modern e-learning platforms. By adopting a microservices model with independent deployment, containerization, and orchestration, it ensures fault isolation, resource efficiency, and flexibility in evolving components. The use of a contract-first, CI/CD-enabled stack with observability and performance monitoring tools supports agile development and reliable operations.

Security, interoperability, and governance are treated as first-class concerns, with support for federated authentication, GDPR-aligned data practices, and integration via standards such as LTI, REST, and optionally GraphQL. Beyond infrastructure, pedagogical features such as a collaborative knowledge center, educational analytics, and gamification are embedded as core services to enhance learner engagement and enable timely interventions.

6.1 Future Work

Future development will focus on delivering a mobile client using cross-platform frameworks like React Native or Flutter. This tier will adopt an offline-first model and integrate with the existing identity and notification systems to ensure session continuity and user engagement under varying connectivity.

To maintain parity with the web experience, the mobile application will replicate accessibility guarantees, instrument telemetry, and leverage automated release strategies for progressive deployment.

Artificial Intelligence enhancements will span adaptive learning guidance, conversational assistance powered by retrieval-augmented generation from the knowledge base, and the refinement of gamification via A/B testing and behavioral analytics. All AI-driven features will adhere to privacy-preserving data practices, include human-in-the-loop mechanisms where appropriate, and undergo continuous efficacy evaluation to ensure responsible impact and trustworthiness.

Acknowledgements

"This work is funded by National Funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project Ref UIDB/05583/2020. Furthermore, we would like to thank the Research Centre in Digital Services (CISeD), the Polytechnic of Viseu, for their support."

Bibliography

[1] M. A. Almaiah, A. Al-Khasawneh, and A. Althunibat. Exploring the critical challenges and factors influencing the e-learning system usage during covid-19 pandemic. *EDUCATION AND INFORMATION TECHNOLOGIES*, 25(6):5261–5280, NOV 2020.

[2] Apereo Foundation. Sakai documentation. https://sakai.readthedocs.io/, 2025. Accessed on: Jun 11, 2025.

[3] Apereo Foundation. Sakai lms source code repository. https://github.com/sakaiproject/sakai, 2025. Accessed on: Jun 11, 2025.

[4] Coursera Inc. Coursera engineering. https://engineering.coursera.org/, 2025. Accessed on: Jun 11, 2025.

[5] Coursera Inc. Coursera for campus documentation. https://www.coursera.org/campus, 2025. Accessed on: Jun 11, 2025.

[6] Coursera Inc. Partners api documentation. https://partners.coursera.help/hc/en-us/categories/360002478312-Partners-API, 2025. Accessed on: Jun 11, 2025.

[7] Docker, Inc. Docker overview, 2025. Accessed on: Jul 23, 2025.

[8] European Parliament and Council. Regulation (eu) 2016/679 (general data protection regulation). Official Journal of the European Union, 2016. Accessed on: Jul 23, 2025.

[9] Fastify Core Team. Fastify — fast and low overhead web framework, 2025. Accessed on: Jul 23, 2025.

[10] GraphQL Foundation. Graphql specification, 2021. Accessed on: Jul 23, 2025.

[11] D. Hardt. The oauth 2.0 authorization framework. Technical Report RFC 6749, IETF, 2012.

[12] Instructure Inc. Canvas lms documentation. https://community.canvaslms.com/t5/Canvas-Basics-Guide/tkb-p/canvas-basics, 2025. Accessed on: Jun 11, 2025.

[13] Instructure Inc. Canvas lms source code repository. https://github.com/instructure/canvas-lms, 2025. Accessed on: Jun 11, 2025.

[14] M. B. Jones, J. Bradley, and N. Sakimura. Json web token (jwt). Technical Report RFC 7519, IETF, 2015.

[15] A. M. Maatuk, E. K. Elberkawi, S. Aljawarneh, H. Rashaideh, and H. Alharbi. The covid-19 pandemic and e-learning: challenges and opportunities from the perspective of students and instructors. *JOURNAL OF COMPUTING IN HIGHER EDUCATION*, 34(1):21–38, APR 2022.

[16] Martin Fowler. Microservices guide. https://martinfowler.com/microservices, 2019. Accessed on: Jul 23, 2025.

[17] Moodle HQ. Moodle developer documentation. https://moodledev.io/, 2025. Accessed on: Jun 11, 2025.

[18] Moodle HQ. Moodle open source lms source code. https://github.com/moodle/moodle, 2025. Accessed on: Jun 11, 2025.

[19] Open edX Consortium. Open edx documentation. https://docs.openedx.org/en/latest/, 2025. Accessed on: Jun 11, 2025.

[20] Open edX Consortium. Open edx source code repository. https://github.com/openedx/edx-platform, 2025. Accessed on: Jun 11, 2025.

[21] PostgreSQL Global Development Group. *PostgreSQL 16 Documentation*, 2024. Accessed on: Jul 23, 2025.

[22] Prometheus Authors. *Prometheus Documentation*, 2025. Accessed on: Jul 23, 2025.

[23] C. Romero and S. Ventura. Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3):e1355, 2020.

[24] M. S. Roodposhti, E. Ezami, M. H. Hedayati, and A. Karimi. User-generated content effectiveness in co-creation of online higher educational services. *JOURNAL OF MARKETING FOR HIGHER EDUCATION*, 2024 MAY 8 2024.

[25] M. Sailer and L. Homner. The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32:77–112, 2020.

[26] A. N. Saleem, N. M. Noori, and F. Ozdamli. Gamification applications in e-learning: A literature review. *TECHNOLOGY KNOWLEDGE AND LEARNING*, 27(1):139–159, MAR 2022.

[27] B. H. Sigelman, L. A. Barroso, M. Burrows, and P. Stephenson. Dapper, a large-scale distributed systems tracing infrastructure. Technical report, Google, 2010.

[28] F. Tapia, M. A. Mora, W. Fuertes, H. Aules, E. Flores, and T. Toulkeridis. From monolithic systems to microservices: A comparative study of performance. *APPLIED SCIENCES-BASEL*, 10(17), SEP 2020.

[29] TechEmpower. Web framework benchmarks (round 22), 2023. Accessed on: Jul 23, 2025.

[30] The Kubernetes Authors. Horizontal pod autoscaler, 2025. Accessed on: Jul 23, 2025.

[31] The Kubernetes Authors. Kubernetes documentation, 2025. Accessed on: Jul 23, 2025.

[32] Udemy Inc. Udemy business documentation. https://business.udemy.com/, 2025. Accessed on: Jun 11, 2025.

[33] Udemy Inc. Udemy engineering blog. https://tech.udemy.com/, 2025. Accessed on: Jun 11, 2025.

[34] Udemy Inc. Udemy terms of use. https://www.udemy.com/terms/, 2025. Accessed on: Jun 11, 2025.

[35] Vercel. Next.js documentation, 2025. Accessed on: Jul 23, 2025.

[36] Vercel. Rendering: Ssr vs ssg, 2025. Accessed on: Jul 23, 2025.

Applying Artificial Intelligence in the Promotion of Art Faculties

Milica Slijepčević[1] [0000-0002-0431-2998], Nevenka Popović Šević[2] [0000-0002-7435-2979], and Tanja Devetak[3] [/]

[1] Belgrade Metropolitan University – Faculty of Management and Faculty of Digital Arts, Serbia,

[2] University Business Academy in Novi Sad, Faculty of Contemporary Arts, Belgrade, Serbia,
[3]Faculty of Design, Independent Higher Education Institution, Ljubljana, Slovenia
milica.slijepcevic@metropolitan.ac.rs
nevenka.popovic.sevic@fsu.edu.rs
tanja.devetak@fd.si

Abstract. Over the past decade, artificial intelligence (AI) has become a pivotal driver of transformation across multiple sectors, including higher education institutions. Art faculties, as integral components of the academic system, face the challenge of modernizing their promotional strategies to attract prospective students and effectively present their educational programs. In this context, AI presents new opportunities for developing personalized and interactive marketing campaigns tailored to the interests and online behavior of prospective students. This analytical approach allows for the generation of customized promotional content and advertisements that resonate with the specific interests of the target audience. Additionally, AI can automate processes such as responding to inquiries via chatbots, thereby enhancing user experience and improving communication efficiency. By leveraging AI technologies, art faculties can design interactive virtual environments where prospective students can navigate campus spaces, access detailed program information, and engage in real-time workshops. Moreover, by analyzing users' emotional responses to various forms of promotional content, AI enables the optimization of content appeal, including the selection of appropriate communication channels and the creation of effective, AI-driven materials.

Keywords: media promotion, art faculties, artificial intelligence.

1 Introduction

In the contemporary educational system, media promotion constitutes a key element in attracting the attention of prospective students and the broader public. As specialized institutions within the educational system, art faculties demand customized promotional approach to effectively communicate their programs, values, and accomplishments. Recognizing the inspirational and motivating nature of art, it is imperative that media promotion strategies be precisely designed to effectively convey the essence and distinctive qualities of art education.

Reveiw Paper
DOI: https://doi.org/10.46793/eLearning2025.119S
Part of ISBN: 978-86-89755-37-4

Media promotion in art faculties serves not only to recruit prospective students but also to develop and sustain the faculty's reputation, and to cultivate strong relationships with alumni and cultural stakeholders. In addition, media promotion strengthens the presence of art faculties across local, national, and international arenas, which is increasingly vital in the context of global competition within the education landscape [1]. Choosing the right promotional strategy involves selecting appropriate media channels and generating effective content with the support of artificial intelligence. This contributes to the development of institutional reputation and reinforces connections with the cultural community. Achieving success in this field requires thoughtful planning, creativity, and adaptability to clearly convey the distinct value of artistic education.

2 Media promotion in higher education institutions

Effective media promotion of educational institutions can be instrumental in student recruitment, increasing awareness of academic programs, and enhancing institutional reputation and brand identity [2]. This paper reviews and analyzes the ways in which educational institutions—particularly art faculties—employ media promotion, supported by artificial intelligence, to attract students, promote their programs and achievements, and build a recognizable institutional brand. It also explores how media promotion shapes public perception of educational institutions and how it can be employed to enhance institutional reputation.

Media promotion refers to a set of strategies and activities aimed at promoting products, services, or—within this context—educational institutions through various media channels. The primary objective of media promotion is to attract the attention of the target audience, provide them with relevant information about available offerings, and encourage them to take specific action—such as enrolling at an academic institution [3].

The success of media promotion in educational institutions frequently results from a combination of media channels and techniques that are specifically designed for the characteristics of the target audience and the academic institutions' objectives. Consequently, the successful attainment of the intended results is contingent upon the proper planning, implementation, and evaluation of media campaigns [4]. Media channels can be employed by educational institutions to communicate information regarding their programs, services, and activities, while simultaneously emphasizing their fundamental values and competitive advantages. Institutions can also monitor trends and interests within their target audience through media, which enables them to adjust their promotional activities accordingly [2].

Media plays a critical role in the promotion of educational institutions by assisting in the recruitment of prospective students. Media can help attract new generations of students by publishing pertinent information regarding enrollment periods, admission requirements, and student experiences [5], [6]. Artificial intelligence's contribution to the advancement of educational institutions is multifaceted and challenging. The institution's success, the recruitment of prospective students, and the establishment of a positive public image can all be significantly influenced by the appropriate utilization of media channels.

3 Media promotion strategy for art faculties

Due to their specific nature, art faculties necessitate a promotional strategy that is meticulously crafted. The primary component is the identification of the target audience and the selection of the most effective method to capture their attention.

3.1 Identifying the target audience of art faculties

Target audience analysis and market segmentation are essential components of an effective media promotion strategy for art faculties. These institutions have a specific target audience that must be carefully identified and understood in order to establish effective communication. The target audience of art faculties may include prospective students, their parents or guardians, faculty members and other academic staff, as well as the general public interested in the faculty's activities and achievements. [3].

By applying market segmentation to tailor their marketing strategies and messaging to distinct audience segments, art faculties can enhance engagement and achieve more effective outcomes. For instance, a faculty may provide targeted information about available programs, faculty members, and career opportunities within a specific artistic discipline to a segment of high school graduates who have already expressed interest in the field. [5]. In addition, market segmentation can help faculties develop a more comprehensive understanding of the needs and trends within each audience segment, enabling them to adapt their programs and activities accordingly. For example, a faculty may observe a growing interest in digital art among younger audiences and broaden its academic programs to cover these trends [7].

The target audience and market segmentation represent critical success factors in the media promotion strategies of art faculties. By identifying their audience and understanding its interests, the faculties can enhance the effectiveness of their communication, attract qualified audiences, and achieve both their promotional and developmental objectives.

3.2 Selection of media channels for promotion

For art faculties, it is essential to select media channels that align with the preferences of their target audience and facilitate outreach to individuals with similar interests. For example, faculties offering visual arts programs may opt for promotional channels—such as art magazines, galleries, specialized art websites, or popular social media—that align with the preferences of the artistic community. [6]. A well-balanced mix of media channels, tailored to the specific needs and characteristics of the target audience, can serve as a critical success factor in the effective promotion of art faculties. [5] It is important for faculties to follow trends and innovations in the media landscape in order to stay up-to-date with the most effective ways of communicating with their target audience. For example, digital marketing tools—such as Google Ads, Facebook and Instagram advertising, and influencer marketing—can be particularly effective in capturing the attention of younger audiences active on social media.
Creating engaging content is a vital component of a media promotion strategy for art faculties. Well-crafted content can capture the attention of the target audience, foster

engagement, raise institutional visibility, and ultimately contribute to student enrollment or enduring stakeholder support.

For art faculties, content may encompass a wide array of creative formats—including photographs and videos featuring student artworks, interviews with faculty members and students, event coverage, and lecture and exhibition highlights, as well as value-driven educational content that communicates the faculty's identity. [6]. By producing relevant and engaging content, art faculties can foster meaningful audience engagement, attract the attention of prospective students and advocates, and achieve their promotional and institutional goals [7].

These initiatives support real-time audience engagement and contribute to increased brand recognition in the digital era. In addition, websites and search engine optimization (SEO) are indispensable for enhancing the faculty's visibility in the digital space. Conversely, email marketing and digital campaigns facilitate personalized communication with prospective students and support throughout the conversion process. By combining these three segments of digital promotion, art faculties establish an effective mechanism for attracting and engaging prospective students [8]. Through the integration of these strategies, they can efficiently utilize digital channels to promote their offerings and enhance student recruitment. This is particularly critical in the current digital era, where an online presence is a prerequisite for success in a highly competitive educational environment [9].

Social media has become an indispensable tool for promoting art faculties in today's digital era. Moreover, their interactive social media functionalities—such as Q&A sessions, polls, and audience-engaged live streaming events—enable faculties to build stronger connections with their target audiences and gather real-time feedback [10]. As a strategic promotional channel, social media enables art faculties to expand their reach, build brand identity, foster community engagement, and connect with prospective students in a way that is relevant, engaging, and interactive [11].

Establishing communities of engaged followers and alumni can provide art faculties with sustained benefits. Creating dedicated groups or pages for students, alumni, faculty members, and other stakeholders can further reinforce a sense of belonging and support. Social media has become an indispensable channel for communication and promotion in art faculties. Through the creative and strategic use of these platforms, faculties can engage effectively with their audiences, build a distinctive brand identity, and attract new talent—thereby enhancing their reputation and strengthening their position within the educational landscape.

Through continuous monitoring of analytics across all media promotion channels, universities can gain insights into website traffic, source attribution, and visitor behavior. Such an approach allows them to assess the impact of their digital strategies and adapt them to achieve the best outcomes possible [12], [13].

3.3 **Applying artificial intelligence in the promotion of art faculties**

Art faculties can leverage artificial intelligence (AI) in various ways to promote their activities, academic programs, and student achievements more effectively and creatively. AI enables precise segmentation of target audiences by analyzing social media data, thereby facilitating the development of promotional campaigns tailored to specific interests—for example, students interested in visual arts, music, design, and related fields. With the support of generative AI tools —such as ChatGPT and DALL·E—it is possible to automatically generate text, graphic, and video advertisements adapted for various media platforms [14]. AI can also be employed for the automated production of posters, artistic catalogs, and audiovisual materials, as well as for generating video and animation content using a range of AI-based tools.

As part of its efforts to develop a strong digital identity, Metropolitan University Belgrade was among the first universities in Serbia to implement a chatbot on its official website. The chatbot serves as a digital assistant, offering information on academic programs, enrollment periods, and other relevant topics. By employing this tool, the university enhances information accessibility while simultaneously optimizing the workload of administrative staff—particularly during peak promotional periods.

To evaluate the effectiveness of promotional efforts, AI tools are capable of conducting detailed analyses to identify which campaigns yield the most favorable outcomes for the faculty. For example, emotional analysis tools can assess prospective audience responses, enabling the adaptation of each piece of content. AI technologies also enable the creation of virtual tours of art faculties, presenting various departments—such as visual arts, fashion design, painting studios, multimedia laboratories, and similar facilities. Additionally, AI can help implement interactive exhibitions and various performances by creating digital installations that allow the audience to participate in the artistic experience using mobile devices or augmented reality (AR) glasses.

These tools are particularly effective in engaging prospective students, as they can deliver information on academic programs, enrollment periods, exhibitions, and other relevant activities. A major advantage of these tools is the information delivery outside standard working hours, which helps lower administrative expenses for academic institutions. With the help of AI, art faculties can launch virtual ambassadors—digital characters that promote the faculty and student life on social media platforms such as TikTok and Instagram.

3.4 Case studies in best practice and prospects for future development

In recent years, art faculties across the globe have increasingly embraced targeted models of artificial intelligence applications to improve communication with prospective students. Whereas the preceding sections focused on overarching applications of AI in the promotion of educational institutions, case studies from international practice further illustrate the efficiency and range of such methods.

An illustrative example of innovative practice involves the use of gamified virtual tours enhanced by generative AI technologies. [15] describe the development of three-dimensional virtual walkthroughs of an educational institution that integrate user interaction, spatial simulation, and narrative flows, offering prospective students a comprehensive digital experience. Such tools not only present content but also contribute to building an emotional bond between candidates and the faculty's spatial context and identity.

An additional significant advancement lies in the application of predictive analytics for selecting specific audience segments and tailoring promotional strategies accordingly. According to [16], AI models are used by universities to analyze the digital footprints of prospective students, supporting timely and more precise campaign interventions. Such strategies contribute to greater relevance in communication and lead to improved conversion rates.

In addition, [17] explores the impact of immersive virtual reality (VR) environments in art-oriented educational institutions, emphasizing their potential to engage students' creative capacities. These findings suggest the potential for applying VR technologies not only within the classroom but also in the promotional and motivational phases—by offering prospective students a "creative experience" prior to enrollment.

Significant use cases of generative AI in visual communication include the application of Generative Adversarial Networks (GANs) for the production of tailored visual content, multimedia materials, and interactive presentations [18]. Within this framework, [19] developed algorithms that predict audience preferences in the arts, providing a foundation for the design of promotional messages and the development of faculties' visual identity.

Prospects for future development include the integration of emotional analytics and multimodal content personalization, wherein AI is employed to simultaneously process visual, verbal, and behavioral data. Such an approach facilitates a comprehensive understanding of prospective students' emotional responses to promotional content, allowing for advanced segmentation and real-time dynamic campaign adaptation.

4 Conclusion

Effective media promotion of art faculties requires ongoing engagement and responsiveness to evolving technologies and communication trends. By systematically monitoring outcomes and collecting feedback, faculties can identify best practices and adjust their future media strategies. Successful promotional efforts enable faculties to attract prospective students while also contributing to the broader dissemination of art and culture within society. Investing in the development and implementation of effective media strategies is therefore a necessary step toward achieving the long-term objectives and ensuring the sustained success of art education institutions.

Developing an effective promotional strategy of art faculties entails a holistic approach that integrates strategic planning, innovative content creation, the use of cutting-edge digital tools, and cooperation with relevant stakeholders. Such an approach supports not only the faculty's success but also promotes art and culture as fundamental societal values.

Disclosure of Interests.

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bilovodska, O., Kravchuk, T., Ponomarenko, I., Danko-Bliumska, K., Kononenko, A.: Artificial intelligence for marketing product strategy in the online education market. Economics of Development, 23(3), 18–31 (2024)
2. Vuletić, S., Todorović, V.: Internet marketing and media promotion of educational institutions. Singidunum Journal of Applied Sciences, Belgrade (2014)

3. Zečević, M.: Media promotion and media destruction. Research Media Center, Podgorica (2019)
4. Mevorah, V.: Internet art in the institutional context: The future prospects and challenges of digital art in Serbia. Proceedings of the Faculty of Dramatic Arts, Vol. 29, Belgrade (2016)
5. Brakus, A.: Creative marketing in the media. In Medias Res, 5(8), Zagreb (2016)
6. Šiđanin, I.: Social media management and the media landscape. School of Business, Vol. 4, Novi Sad (2012)
7. Pavlović, D.: New media technologies in the education of journalism students at state faculties in Serbia. PhD thesis, Faculty of Philosophy, University of Novi Sad (2023)
8. Ferenčić, M.: Marketing Communication in the Digital World. Praktični menadžment, 3(2), Virovitica (2012)
9. Starčević, M.: Strategic marketing planning in cultural institutions. PhD thesis, Faculty of Organizational Sciences, University of Belgrade (2016)
10. Raduški, D.: Marketing management in culture and the arts. Marketing 48(4), Belgrade (2017)
11. Prnjak, D.: The competitiveness of art faculties in the higher education market: Insights from students and faculty members. Marketing 53(1), Belgrade (2022)
12. Milić, K.: Cultural creativity in the digital environment. CM: Communication and Media 17(52), Belgrade (2022)
13. Ćetojević, H.: The impact of digital literacy on the development of higher education. Pedagoška stvarnost, 69(2), Novi Sad (2023)
14. Kumar, V., Ashraf, A.R., Nadeem, W.: AI-powered marketing: What, where, and how? International Journal of Information Management 77, 102783 (2024)
15. 7. 8. Garcia, M.B., Perez, M.R.L., Pempiña, E.B., Mansul, D.M.C., Adao, R.T.: A playable 3D virtual tour for an interactive campus visit experience: Showcasing school facilities to attract potential enrollees. In: 2023 9th International Conference on Virtual Reality. IEEE. https://doi.org/10.1109/ICVR57957.2023.10169768 (2023)
16. Bird, K.: Predictive analytics in higher education: The promises and challenges of using machine learning to improve student success. AIR Professional File No. 161. Association for Institutional Research. https://doi.org/10.34315/apf1612023 (2023)
17. Guerra-Tamez, C.R.: The impact of immersion through virtual reality in the learning experiences of art and design students: The mediating effect of the flow experience. Education Sciences 13(2), 185 (2023)
18. Anantrasirichai, N., Bull, D.: Artificial intelligence in the creative industries: A review. arXiv preprint arXiv:2007.12391 (2020)
19. Abernethy, J., Anderson, C., Chojnacki, A., Dai, C., Dryden, J., Schwartz, E., Shen, W., Stroud, J., Wendlandt, L., Yang, S., Zhang, D.: Data science in service of performing arts: Applying machine learning to predicting audience preferences. arXiv preprint arXiv:1611.05788 (2016)

Affective Generative AI for Adaptive and Inclusive eLearning: Prompt Engineering, Ethics and Pedagogical Innovation

Spyridon Kontis[1[0009-0009-7892-6966]] and Sofia Anastasiadou[1[0000-0001-6404-5003]]

[1] University of Western Macedonia, Greece
dmw00034@uowm.gr, sanastasiadou@uowm.gr

Abstract. Over the past few years, artificial intelligence (AI) has moved from being a background tool to becoming a central actor in education [49]. Generative AI is transforming the way learners interact with knowledge, offering personalized pathways, adaptive content, and new forms of digital collaboration. Yet most existing systems remain focused on efficiency and performance, while overlooking the emotional side of learning, factors such as motivation, frustration, or anxiety that often determine whether a student succeeds or disengages [1], [2].

This paper proposes the idea of Affective Generative AI in eLearning, combining large language models (LLMs), prompt engineering, and emotion-aware computing to design learning environments that are not only intelligent but also empathetic and inclusive. We argue that digital tutors capable of recognizing affective cues can adapt their responses in real time, providing encouragement, reframing explanations, or reducing cognitive load, thereby supporting both well-being and achievement [3], [4].

At the same time, handling emotional data [47] raises critical ethical and legal [12] questions. Issues of privacy, bias [51], and transparency must be addressed if such systems are to be trusted and responsibly deployed [5]– [7]. Our conceptual framework seeks to balance pedagogical innovation with these concerns, highlighting a path towards human-centered [53] AI in education that values inclusion, equity, and emotional resilience alongside cognitive performance.

Keywords: Generative AI, Affective Computing, Prompt Engineering, Adaptive Learning, LLMs, AI Ethics, Inclusive Education.

1 Introduction

Education in the 21st century is increasingly mediated by digital technologies that promise to expand access, enhance personalization, and support lifelong learning. eLearning platforms have become essential, yet their design has traditionally focused on delivering content and assessing performance, while neglecting the affective dimension of learning [8]. Emotions such as engagement, frustration, and motivation are now recognized as central to the learning process, influencing persistence, memory, and overall achievement [1], [2]. When these affective states are ignored, learners,

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.126K
Part of ISBN: 978-86-89755-37-4

especially those from vulnerable or marginalized groups, face higher risks of disengagement and exclusion [9].

Generative artificial intelligence (AI) has introduced new possibilities for adaptive and scalable digital education. Large language models (LLMs) can generate personalized explanations, tailor exercises to a learner's level, and provide immediate feedback. However, these systems remain cognitively adaptive but emotionally indifferent [10]. Recent scholarship argues that the next step for AI in education must be its integration with affective computing, enabling digital tutors not only to adapt content but also to respond to learners' emotions in real time [3], [4]. This shift aligns with broader movements in inclusive education, which emphasize equity, empathy, and responsiveness to learner diversity [11].

At the same time, the adoption of affective AI raises significant ethical and legal challenges. Emotional data is inherently sensitive, requiring careful consideration of privacy, fairness, and transparency [5], [6], [7]. Scholars in AI ethics stress that without robust governance [46] [14] frameworks, the same technologies that promise inclusion and wellbeing may instead amplify bias and erode trust [12], [13]. This paper therefore positions affective generative AI at the intersection of pedagogy, technology, and ethics, aiming to contribute to a more human-centered vision of eLearning.

The aim of this study is threefold:

1. To conceptualize an affective generative AI framework for adaptive and inclusive eLearning.
2. To examine how prompt engineering can integrate affective signals into LLM-driven educational interactions.
3. To discuss the ethical, legal, and psychosocial implications of deploying such systems in real-world contexts.

By pursuing these aims, the paper addresses a critical gap in current eLearning systems and argues for a paradigm shift: from AI that merely delivers knowledge to AI that empathizes, supports, and includes.

Unlike previous adaptive eLearning approaches that primarily emphasize cognitive personalization, our study contributes by explicitly integrating affective signals into generative AI systems through prompt engineering. This integration allows learning environments to be not only knowledge-centered but also emotion-sensitive, addressing the overlooked psychosocial dimensions of digital education. By situating our framework at the intersection of pedagogy, computational modeling, and ethics, the paper provides a novel contribution that bridges theoretical perspectives with actionable design principles for inclusive eLearning.

2 Affective Computing and the Role of Emotions in Learning

The study of emotions in human–computer interaction was first introduced through the field of affective computing, which highlighted the potential for machines to recognize and respond to emotional states [1]. In education, this perspective has become increasingly important, as re- search shows that emotions such as frustration, curiosity, or anxiety directly shape attention, memory, and motivation [2]. Learners who feel emotionally supported are more likely to persist, collaborate, and achieve meaningful outcomes, whereas neglecting the affective dimension can lead to disengagement and inequality [3].

Affective learning environments move beyond traditional instruction by embedding empathy into digital systems. Such environments aim to provide scaffolding that adapts not only to cognitive performance but also to emotional needs. For instance, when a student experiences cognitive overload, an affective-aware system may slow the pace, simplify explanations, or offer encouragement, helping to sustain resilience and wellbeing [4], [5]. This human-centered approach underscores the fundamental link between emotions and inclusion, ensuring that every learner's experience is acknowledged.

Generative artificial intelligence (AI) has rapidly advanced from ex- perimental tools to mainstream applications in education. Large language models (LLMs) can generate explanations, create adaptive exercises, and provide personalized feedback, thereby extending the possibilities of digital learning environments [6]. These systems represent a significant step towards more flexible and individualized learning, as they can dynamically tailor content to a learner's level of knowledge and preferred style of interaction [7].

Despite these advantages, current implementations of generative AI in education remain primarily cognitively adaptive. They adjust what and how of learning materials, but they do not yet address how learners feel during the process [8]. Without accounting for affective states, adaptive learning [48]-[56] risks becoming mechanistic, capable of personalizing tasks but not of sustaining engagement when learners struggle with motivation, stress, or anxiety [9].

Integrating affective signals into generative systems can bridge this gap. For example, a digital tutor powered by an LLM could detect signs of disengagement or cognitive overload and adjust its responses accordingly: simplifying explanations, providing motivational scaffolding, or reframing challenges to reduce frustration [10], [11]. Such responsiveness- ness moves adaptive learning closer to true inclusivity, where technology does not simply optimize performance but also recognizes and supports the learner as a whole person

3 Ethical and Legal Dimensions of Affective AI

While the pedagogical potential of affective generative AI is significant, its adoption raises equally important ethical and legal questions. Emotional data is among the most sensitive categories of personal information, and their collection and analysis in digital learning environments must be governed by strict principles [54] of privacy, transparency, and accountability [12], [13]. The risk is that technologies designed to foster inclusion could unintentionally reinforce bias, manipulate behavior, or exacerbate inequalities if deployed without adequate safeguards [14].

Scholars in AI ethics have repeatedly stressed the dangers of opacity, the so-called "black box" problem, where even system designers may struggle to explain why an algorithm produced a certain output [15]. In education, such opacity undermines trust, as learners and educators need clarity about how affective data are being interpreted and used [16]. Calls for responsible [50] artificial intelligence emphasize principles such as fairness, non-discrimination, and respect for human dignity [17], [18].

Embedding these principles into affective systems is essential if they are to become trustworthy allies in inclusive education.

Legal perspectives also contribute to this discussion. The rapid pace of AI development has outstripped existing regulatory frameworks, leaving gaps around liability, accountability, and cross-border governance [19], [20]. Recent debates around the European Union's AI Act and related directives illustrate the complexity of balancing innovation with protection of fundamental rights [21]. Scholars argue that frameworks must not only regulate technical standards but also anticipate new risks, including the commodification of emotions and the potential misuse of biometric affective data [22], [23].

Finally, affective AI must be situated within the broader socio-political- cal critique of digital technologies. Zuboff's analysis of surveillance [55] capitalism warns of the dangers of extracting personal experiences for profit [7]. When applied to education, this critique raises uncomfortable questions: should learners' emotions be treated as data points for optimization, or as integral parts of their human identity that deserve protection? Addressing such tensions requires an interdisciplinary dialogue between pedagogy, law, technology, and ethics [24].

4 Methodology

The methodological approach of this study combines a conceptual framework with computational underpinnings that support adaptive and inclusive e-learning. The design draws on socio-technical perspectives that emphasize the interplay between governance structures and algorithmic- mic processes [44,46].

From a technical standpoint, the methodological grounding is influenced by prior research in distributed systems and big data management. For example, dynamic scheduling of data streams has been shown to improve system responsiveness and efficiency in heterogeneous environments [48,52]. Similarly, pipeline-based approaches and linear scheduling in the cloud demonstrate how complex learning workloads can be optimized for scalability and fairness [51].

Advanced modeling tools, such as Colored Petri Nets, enable the representation of hierarchical and dynamic system behaviors, allowing researchers to simulate and validate resource allocation strategies in e-learning infrastructures [53,55]. Probabilistic detection methods applied to social and learning networks provide additional insights into the collective behaviors that shape adaptive learning environments [56]. This methodological synthesis ensures that the research framework is not only ethically and legally grounded [45,47], but also supported by rigorous computational models capable of sustaining adaptive, scalable, and inclusive AI-driven learning platforms.

5 Results

The synthesis of the literature reveals several recurring patterns that highlight both the opportunities and the challenges of embedding affective generative AI into eLearning. Three clusters of findings are particularly noteworthy: (a) the pedagogical benefits of integrating affective signals into adaptive systems, (b) the risks of ethical and legal misalignment, and (c) the broader socio-psychological implications for inclusion and learner wellbeing.

Pedagogical benefits. Research consistently demonstrates that emotions act as critical drivers of learning outcomes. Learners who feel motivated, supported, and emotionally engaged are more likely to persist and achieve higher cognitive gains [1], [2]. Digital environments that incorporate affective feedback can transform the learning experience from passive content reception to active, emotionally enriched participation [8], [9]. Generative AI offers new potential in this regard: through large language models (LLMs) and prompt engineering, systems can dynamically adapt instructional strategies, reframe explanations, or provide empathetic encouragement when signs of frustration or disengagement are detected [10], [11]. Such responsiveness aligns with broader educational goals of personalization and equity, supporting learners not only as knowledge receivers but as whole people with emotional needs.

Risks and challenges. Alongside these pedagogical promises, the literature reveals deep concerns about the ethical deployment of affective technologies. The interpretation of emotional data is fraught with risks of misclassification, cultural bias, and oversimplification of complex human experiences [12]. Furthermore, emotional data are highly sensitive: their collection, storage, and use raise acute concerns of privacy, surveillance, and manipulation [13], [14]. Scholars warn that without robust safeguards, affective AI could amplify systemic inequities rather than reduce them, for example by reinforcing stereotypes about certain groups of learners [15], [16]. The “black box” nature of advanced generative systems further complicates this picture, as even system designers may struggle to explain how outputs are derived [17]. In educational contexts, where trust is fundamental, such opacity undermines confidence among both learners and educators [18].

Psychosocial and inclusion dimensions. A third cluster of insights highlights the psychosocial importance of embedding affective aware- ness into educational technology. Inclusive education is not only about access but also about recognition, ensuring that learners’ diverse emotional experiences are acknowledged and valued [19]. Studies on inclusive pedagogy emphasize that when students feel their emotions are validated, their sense of belonging and participation improves markedly [20]. Conversely, neglecting affective dimensions risks marginalizing those who already face barriers, such as learners with disabilities, neurodiverse students, or those experiencing anxiety and stress in digital environments [21]. By combining generative AI with affective computing, systems have the potential to reduce these barriers, offering timely support and fostering resilience [22], [23]. However, this promise will only materialize if ethical principles such as fairness, accountability, and transparency are embedded into design and governance processes from the outset [24], [25].

Taken together, these findings suggest that affective generative AI occupies a paradoxical position. On the one hand, it has the capacity to humanize digital learning environments, bringing empathy, motivation, and inclusion into spaces that often feel abstract and isolating. On the other hand, it risks intensifying surveillance, commodification, and inequality if developed without adequate ethical foresight. Addressing this tension requires moving beyond isolated technical solutions towards a comprehensive framework that integrates pedagogy, ethics, and psycho-social well-being.

This discussion points to the need for a structured model of adoption that can guide institutions in balancing innovation with responsibility. Such a model must ensure that generative AI systems in education are not only technically effective but also aligned with human values—supporting equity, protecting rights, and nurturing the emotional lives of learners. It is on this foundation that the proposed framework for this paper is built.

6 Framework Proposal: The Emotion-Aware Generative eLearning Model

Building on the insights from the literature, this paper proposes the Emotion-Aware Generative eLearning Model (EAGeL), a three-layered framework designed to integrate affective computing with generative AI for adaptive and inclusive digital education. The framework emphasizes both pedagogical innovation and ethical responsibility, ensuring that learners' cognitive and emotional needs are recognized while safeguarding- ing their rights and well-being.

Layer 1 – Emotional Data Recognition and Interpretation.

The first layer focuses on capturing and interpreting affective signals from learners. These may include behavioral indicators (e.g., hesitation, error patterns), self-reports (e.g., quick surveys), or biometric cues where ethically permissible (e.g., facial expression, heart-rate variability) [26]. Unlike traditional affective computing approaches that often reduce emotions to simplistic categories, this layer emphasizes contextualized interpretation. The system must recognize frustration differently in a child with learning disabilities than in an adult learner in higher education [27].

Layer 2 – Generative Adaptation through Prompt Engineering.

The second layer integrates affective inputs into large language models (LLMs) through carefully designed prompt engineering strategies. Emotional cues dynamically shape the AI's responses: for example, when frustration is detected, the model may reframe the explanation, simplify instructions, or introduce encouraging feedback [28]. Similarly, signs of boredom could trigger more interactive or gamified tasks, while indications of anxiety may lead to shorter, step-by-step guidance [29]. This adaptive loop ensures that learners receive not only personalized content but also emotionally attuned support.

Layer 3 – Ethical and Inclusive Governance.

The third layer provides the ethical "scaffolding" of the framework. Handling affective data requires strong safeguards to ensure privacy, fair- ness, and transparency [5], [13], [24]. This includes anonymization protocols, explainable AI mechanisms, and clear policies about data use. In addition, the governance layer aligns the framework with principles of Universal Design for Learning (UDL) and the European DigCompEdu framework, ensuring that emotion-aware generative systems support inclusion and accessibility across diverse contexts [30], [31]. Interdisciplinary- nary collaboration, among educators, technologists, ethicists, and policymakers, is essential to maintain alignment with human values.

7 Scalability and Sustainability

The framework is designed to be implemented in stages, beginning with low-risk affective inputs such as learner self-reports and progressing to more advanced multimodal signals as ethical and technical maturity increases [32]. This phased approach allows institutions to test, refine, and scale emotion-aware generative systems responsibly, avoiding premature adoption that could undermine trust or equity [33].

In essence, the EAGeL model seeks to reimagine digital education not only as a space for knowledge delivery but as an environment where learners feel recognized, supported, and included. By integrating emotion-sensitive prompts into generative AI, the framework offers a roadmap towards empathetic and equitable eLearning ecosystems.

8 Practical Scenarios

8.1 Scenario 1 – Managing math anxiety in online tutoring.

A high-school learner repeatedly struggles with algebra tasks on an online platform. Delayed responses, high error rates, and a quick self- report about feeling “stressed” are interpreted as signals of math anxiety [34]. Instead of re-explaining the same formula, the AI tutor adapts its response style: it breaks the problem into smaller steps, praises incre- mental progress, and introduces motivational scaffolding. Over time, the system also suggests short reflection prompts to help the learner regulate stress, aligning with evidence that affective support can improve resili- ence in mathematics learning [35]. This illustrates how generative AI can become not only a tutor but also a coach for emotional self-efficacy, pre- venting dropout in a subject often associated with student anxiety.

8.2 Scenario 2 – Rekindling motivation in disengaged learners.

In a digital history course, several students display low engagement: minimal contributions to forums, rapid skimming of digital resources, and skipping of optional assignments. The platform flags this as disengagement and possible boredom. The AI responds by reframing the learning path: introducing short quizzes framed as challenges, weaving narratives that connect historical events to current issues, and suggesting group debates through interactive prompts [36]. These interventions draw on gamification strategies, which research shows can reignite curiosity and enhance persistence in online courses. Importantly, the system does not punish disengaged learners but seeks to re-motivate them, reflecting inclusive pedagogy principles where all students are given opportunities to re-enter the learning process on their own terms [9].

8.3 Scenario 3 – Supporting neurodiverse learners in MOOCs.

A programming MOOC includes a student on the autism spectrum who struggles with vague instructions and unpredictable tasks. Over time, the system notices repeated requests for clarification and signs of withdrawal. The AI responds by offering structured task lists, step-by-step explanations, and optional peer-support groups. Although technically simple, these adjustments reduce frustration and create a sense of pre- predictability. This practice resonates strongly with the Universal Design for Learning (UDL) principles, which call for multiple means of representation and engagement, and with the DigCompEdu framework, which emphasizes teacher capacity to personalize learning through technology [37], [38]. Such scenarios illustrate how affective generative AI can extend inclusion policies into practical, everyday learning contexts, supporting not only academic performance but also social belonging.

8.4 Scenario 4 – Promoting wellbeing in higher education.

During exam season, university students in an online course show sign of overload: high dropout from optional tasks, shorter attention spans, and careless mistakes. The system interprets this as stress [39]. In re- response, it recommends pacing strategies, encourages short breaks, and even reframes assessment with alternative formats. These well-being-focused interventions are modest, but they remind learners that the digital environment is designed with care for their mental health. Such approaches are consistent with international recommendations, such as those of UNESCO and the WHO, which emphasize the integration of mental well-being into digital education systems [40].

These scenarios illustrate how the EAGeL framework can translate abstract concepts into real-world impact. By addressing anxiety, boredom, neurodiverse needs, and well-being, affective generative AI demonstrates its potential to create empathetic, adaptive, and inclusive learning ecosystems.

9 Discussion

The case scenarios and methodological approach demonstrate that affective generative AI has the potential to foster inclusion and personalization in eLearning while also raising critical governance and ethical questions. The results highlight not only the feasibility of adaptive models but also the tension between technical optimization and societal expectations. Previous works confirm that the co-evolution of AI and law must be considered as a dynamic, autopoietic process, where educational practices and legal norms evolve together [44]. At the same time, digital learning environments are deeply intertwined with the "digital DNA" of the modern workforce, linking AI competences with employability and organizational culture [45].

From a technical standpoint, advanced computational approaches such as pipelined dynamic scheduling and Markov process modeling provide scalable solutions for real-time educational platforms [48], [50], [52]. These methods support adaptive allocation of resources and ensure efficient personalization of learning flows, especially in heterogeneous cloud environments [54]. Furthermore, community detection in social networks contributes to identifying clusters of learners with similar needs, thereby reinforcing the inclusivity of affective eLearning systems [56].

On the governance side, the concept of "legal entropy" offers a useful lens to analyze the uncertainty and fragmentation in AI regulation [46]. Education cannot be isolated from these dynamics, since the integration of affective AI implies new responsibilities for teachers, institutions, and policymakers. The use of ESG-aligned data governance frameworks further emphasizes the link between AI adoption in eLearning and broader sustainability agendas [47]. This dual perspective, technical and legal, hows that building trustworthy, human-centered AI in education requires bridging algorithmic design with socio-ethical constraints.

Overall, the discussion illustrates that while the presented scenarios show promise, they must be interpreted within broader debates on responsible AI, digital ethics, and the resilience of educational institutions to disruptive technologies. In light of these insights, it becomes evident that affective generative AI in e-learning cannot be evaluated solely on the basis of its technical performance. Its real value lies in how effectively it aligns with ethical imperatives, legal frameworks, and the human-centered principles of inclusion and mental well-being. The in- interplay between technological innovation and responsible governance highlights both the opportunities and the unresolved challenges that ed- educators, policymakers, and developers must jointly address. These con- considerations naturally lead to the concluding reflections of this study, where the implications and forward-looking directions are drawn together.

The contribution of this study lies in extending existing models of adaptive eLearning towards an affective generative paradigm. While earlier research has demonstrated the potential of large language models and adaptive scheduling techniques for personalization, few works have systematically integrated emotional data into these processes. By linking computational models such as Petri Nets and Markov-based scheduling with real-time affective cues, our framework provides a distinctive pathway to connect technical scalability with pedagogical empathy. This dual focus positions the paper as a bridge between conceptual discussions of AI in education and practical, implementable architectures for inclusive learning.

10 Conclusions and Recommendations

This paper has explored the emerging field of Affective Generative AI in eLearning, emphasizing how the integration of emotion-aware computing with large language models can reshape digital education. Through the proposed EAGeL framework, we outlined how affective signals can be recognized, embedded into generative responses, and governed through ethical safeguards. Literature synthesis and practical scenarios demonstrated that such systems could promote motivation, resilience, and inclusion, while also highlighting the risks of bias, surveillance, and opacity.

Three key recommendations arise from this study. First, educational institutions should approach affective AI through phased adoption, beginning with low-risk applications such as self-report data, before expanding to multimodal affective sensing [41]. This gradual pathway allows for testing and adaptation without overwhelming learners or educators. Second, teacher professional development must be prioritized: without adequate training, even the most advanced tools risk underuse or misuse [42].

Teachers need not only technical skills but also critical awareness of ethical issues, so that they can act as mediators between learners and AI systems. Third, ethical and legal framework [44]s must evolve in parallel with technological development. Regulatory instruments such as the EU AI Act, alongside guidelines from UNESCO and WHO, should inform national and institutional policies to safeguard privacy, ensure fairness, and prevent misuse of emotional data [21], [40], [43].

Future research should examine how affective generative AI operates across diverse cultural and educational contexts, and whether its benefits can be sustained over time. Comparative case studies, longitudinal evaluations, and participatory research with learners and educators will be essential in validating the promises of the EAGeL framework. Ultimately, the goal is not to replace human educators but to design empathetic digital ecosystems where technology amplifies care, inclusion, and well-being.

By embedding empathy into generative AI, education can move closer to a vision where every learner, regardless of background, ability, or emotional state, finds recognition, support, and opportunity in the digital classroom.

As a next step, the proposed framework could be tested in pilot studies across different learning environments, such as open online courses (MOOCs), blended university classes, or special education programs. A first stage could focus on simple, low-risk methods like learner self-reports or analysis of interaction logs, and then gradually expand to more advanced indicators such as optional biometric data, always under strict ethical safeguards. The aim of these pilots would be to explore whether affect-sensitive prompt engineering helps learners stay engaged, reduce anxiety, and feel more included. Although still hypothetical at this stage, describing such a pathway shows that the framework is not only theoretical but also has real potential for future application

Disclosure of Interests.

he authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chennupati, A. (2024). The evolution of AI: What does the future hold in the next two years. World Journal of Advanced Engineering Technology and Sciences, 12(1), 022–028. https://doi.org/10.30574/wjaets.2024.12.1.0176
2. Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision–making. Frontiers in Human Dynamics, 6, 1421273. https://doi.org/10.3389/fhumd.2024.1421273
3. Chipiga, I. V. (2023). Legal entropy in International Law. Russian Journal of Legal Studies, 10(2), 113–120. https://doi.org/10.17816/RJLS322823
4. Fomin, V. V. , & Astromskis, P. (2023). The black box problem. In J. S. Gordon (Ed.), Philosophy and Human Rights (pp. 112–125). Brill. https://doi.org/10.1163/9789004682900_012

5. Frosio, G. (2024). Algorithmic enforcement tools: Governing opacity with due process. In S. Francese & R. King (Eds.), Driving Forensic In- novation in the 21st Century (pp. 195–218). Springer. https://doi.org/10.1007/978–3–031–56556–4_9
6. Huang, K. , Joshi, A. , DunAnon., & Hamilton, N. (2024). AI regulations. In K. Huang, Y. Wang, B. Goertzel, Y. LiAnon. Wright, & J. Ponnapalli (Eds.), Generative AI Security (pp. 61–98). Springer. https://doi.org/10.1007/978–3–031–54252–7_3
7. Koridena, B. K. , Malaiyappan, J. N. A., & Tadimarri, A. (2024). Ethical considerations in the development and deployment of AI systems. Euro- pean Journal of Technology, 8(2), 41–53. https://doi.org/10.47672/ejt.1890
8. Kostic, M. M. (2014). The elusive nature of entropy and its physical meaning. Entropy, 16(2), 953–967. https://doi.org/10.3390/e16020953
9. Miazi, M. A. N. (2023). Interplay of legal frameworks and artificial in- telligence (AI): A global perspective. Law and Policy Review, 2(2), 01–25. https://doi.org/10.32350/lrp.22.01
10. Mylrea, M. , & Robinson, N. (2023). Artificial Intelligence (AI) trust framework and maturity model: Applying an entropy lens to improve security, privacy, and ethical AI. Entropy, 25(10), 1429. https://doi.org/10.3390/e2510142937895550
11. Nove, I. C. , Taddeo, M., & FloridiAnon. (2023). Accountability in ar- tificial intelligence: What it is and how it works. AI & Society, 39(4), 1871–1882. https://doi.org/10.1007/s00146–023–01635–y
12. Papademetriou, C. (2012). To what extent is the Turing Test still im- portant? Pliroforiki, 1(22), 28–32.
13. Papademetriou, C. , Ragazou, k. , & Garefalakis, A. (2023). Adapting Artificial Intelligence in Cypriot Hotel Industry: The views of hotel man- agers. In Eurasia Business and Economics Society Conference Proceed- ings (pp. 143–155).
14. Papademetriou, C. , Ragazou, K. , Garefalakis, A., & Anon. (2024). The role of Artificial Intelligence in employee wellness and mental health. In Global I. G. I. (Ed.), Human Resource Strategies in the Era of Artificial Intelligence (pp. 29–54). IGI Global. https://doi.org/10.4018/979–8– 3693–6412–3.ch002
15. Poli, P. K. R., PamidiAnon., & PoliAnon. K. R. (2025). Unraveling the ethical conundrum of artificial intelligence: A synthesis of literature and case studies. Augmented Human Research, 10(1), 2. https://doi.org/10.1007/s41133–024–00077–5
16. Rashid, A. B. , & Kausik, M. A. (2024). AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications. Hy- brid Advances, 7, 100277. https://doi.org/10.1016/j.hy- badv.2024.100277
17. RawasAnon. (2024). AI: The future of humanity. Discover Artificial In- telligence, 4(25), 25. https://doi.org/10.1007/s44163–024–00118–3
18. Robles, P. , & Mallinson, D. J. (2023). Catching up with AI: Pushing toward a cohesive governance framework. Politics & Policy, 51(3), 355–372. https://doi.org/10.1111/polp.12529
19. Scholes, M. S. (2025). Artificial intelligence and uncertainty. Risk Sci- ences, 1, 100004. https://doi.org/10.1016/j.risk.2024.100004
20. Anon. , & Kougias, I. (2019). Legal issues within ambient intelligence environments. In 2019 10th International Conference on Information, In- telligence, Systems and Applications (IISA) (pp. 1–2). IEEE. https://doi.org/10.1109/IISA.2019.8900748
21. Anon. , & Kougias, I. (2024). The AI spectrum under the doctrine of necessity: Towards the flexibility of the future legal praxis. Computers and Artificial Intelligence, 2(1), 1258. https://doi.org/10.59400/cai.v2i1.1258
22. Sichelman, T. (2021). Quantifying legal entropy. Frontiers in Physics, 9, 665054. https://doi.org/10.3389/fphy.2021.665054

23. SouravlasAnon. , & Anon. (2022). On implementing social community clouds based on Markov models. IEEE Transactions on Computational Social Systems. https://doi.org/10.1109/TCSS.2022.3213273
24. SouravlasAnon. , Anon. , Economides, T., & KatsavounisAnon. (2023). Probabilistic community detection in social networks. IEEE Access, 11, 25629–25641. https://doi.org/10.1109/ACCESS.2023.3257021
25. SouravlasAnon. , Anon. , & Kostoglou, I. (2022). A novel method for general hierarchical system modeling via colored Petri nets based on transition extractions from real data sets. Applied Sciences, 13(1), 339. https://doi.org/10.3390/app13010339
26. SouravlasAnon. , Anon. , Tantalaki, N., & KatsavounisAnon. (2022). A fair, dynamic load balanced task distribution strategy for heterogeneous cloud platforms based on Markov process modeling. IEEE Access, 10, 26149–26162. https://doi.org/10.1109/ACCESS.2022.3157435
27. Vinothkumar, J. , & Karunamurthy, A. (2023). Recent advancements in artificial intelligence technology: Trends and implications. International Journal of Multidisciplinary Scientific Research and Development, 2(1), 01–11. https://doi.org/10.54368/qijmsrd.2.1.0003
28. WachterAnon. (2024). Limitations and loopholes in the EU AI Act and AI liability directives. SSRN, 26(3), 671–718. https://doi.org/10.2139/ssrn.4924553
29. Wischmeyer, T. (2020). Artificial intelligence and transparency: Open- ing the black box. In Regulating Artificial Intelligence (pp. 75–101). Springer. https://doi.org/10.1007/978–3–030–32361–5_4
30. Wong, A. (2020). The laws and regulation of AI and autonomous sys- tems. In L. Strous, R. Johnson, D. Grier, & D. Swade (Eds.), Unimagined Futures – ICT Opportunities and Challenges (Vol. 555, pp. 38–54). Springer. https://doi.org/10.1007/978–3–030–64246–4_4
31. Xi, R. (2025). On emerging technologies: The old regime, and the pro- activity. Cardozo International & Comparative Law Review, 8(1).
32. Zaidan, E. , & Ibrahim, I. A. (2024). AI governance in a complex and rapidly changing regulatory landscape: A global perspective. Humanities & Social Sciences Communications, 11(1), 1121. https://doi.org/10.1057/s41599–024–03560–x
33. Bailenson, J. N. (2018). Experience on demand: What virtual reality is, how it works, and what it can do. W. W. Norton & Company.
34. Cummings, M. L. (2018). Artificial intelligence and the future of education. AI Magazine, 39(3), 25–36. https://doi.org/10.1609/aimag.v39i3.2807
35. Dignum, V. (2019). Responsible Artificial Intelligence: How to develop and use AI in a responsible way. Springer. https://doi.org/10.1007/978– 3–030–30371–6
36. FloridiAnon. , & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1
37. Koene, A. , Dowthwaite. , SethAnon., & Webb, H. (2021). Algorithmic bias: Impact on children's mental health. The Lancet Child & Adolescent Health, 5(3), 174–176. https://doi.org/10.1016/S2352–4642(21)00013–5
38. Mittelstadt, B. D. , Allo, P., Taddeo, M., WachterAnon., & FloridiAnon. (2016). The ethics of algorithms: Mapping the debate. Big Data & Soci- ety, 3(2). https://doi.org/10.1177/2053951716679679
39. Riedl, M. O. (2020). Human–centered artificial intelligence and narrative intelligence. Communications of the ACM, 63(12), 62–71. https://doi.org/10.1145/3381831
40. Risser, H. M. , & Bottoms, B. L. (2020). Emotion recognition, AI, and children's rights. Journal of Children and Media, 14(3), 325–341. https://doi.org/10.1080/17482798.2020.1778487
41. Anon. , & Kougias, I. (2021a). The legalhood of artificial intelligence: AI applications as energy services. Journal of Artificial Intelligence and Systems, 3, 83–92.

42. Anon. , & Kougias, I. (2021b). Category theory as interpretation law model in artificial intelligence era. Journal of Artificial Intelligence and Systems, 3, 35–47.
43. Sharkey, N. (2019). Autonomous weapons systems, killer robots and human dignity. Law, Innovation and Technology, 11(1), 107–128.
44. Anon. , & Anon. (2025). Autopoietic co–evolution of AI and law. In S. Anon. , A. Masouras, & L. Anastasiadis (Eds.), Modern Perspectives on Artificial Intelligence and Law (pp. 81–88). IGI Global Scientific Publishing.
45. Kalogera, C. , Anon. , Anon., & Anon. (2025). The digital DNA of mod- ern workforce. In L. , P. LiargovasAnon. Anastasiadis, & S. (Eds.), Harnessing Business Intelligence for Modern Talent Management (pp. 379–390). IGI Global.
46. Anon. , Anon. , AnastasiadisAnon., & Anon. (2025). Legal entropy in AI governance. In S. Anon. , A. Masouras, & L. Anastasiadis (Eds.), Modern Perspectives on Artificial Intelligence and Law (pp. 193–206). IGI Global Scientific Publishing.
47. Anon. , AnastasiadisAnon. , & Anon. (2024). The role of data in envi- ronmental, social, and governance strategy. In C. PapademetriouAnon. , K. Ragazou, A. Garefalakis, & S. Papalexandris (Eds.), ESG and Total Quality Management in Human Resources (pp. 201–218). IGI Global Scientific Publishing.
48. SouravlasAnon. , & Anon. (2021). Pipelined dynamic scheduling of big data streams. Applied Sciences, 10(14), 4796.
49. SouravlasAnon. , & Anon. (2022). On implementing social community clouds based on Markov models. IEEE Transactions on Computational Social Systems.
50. SouravlasAnon. , Anon. , & KatsavounisAnon. (2021a). More on pipelined dynamic scheduling of big data streams. Applied Sciences, 11(1), 61.
51. SouravlasAnon. , Anon. , & KatsavounisAnon. (2021b). A survey on the recent advances of deep community detection. Applied Sciences, 11(16), 7179.
52. Tantalaki, N. , SouravlasAnon. , Roumeliotis, M., & KatsavounisAnon. (2020). Pipeline–based linear scheduling of big data streams in the cloud. IEEE Access, 8, 117182–117202.
53. SouravlasAnon. , KatsavounisAnon. , & Anon. (2021c). On modeling and simulation of resource allocation policies in cloud computing using colored Petri nets. Applied Sciences, 10(16), 5644.
54. SouravlasAnon. , Anon. , Tantalaki, N., & KatsavounisAnon. (2022). A fair, dynamic load balanced task distribution strategy for heterogeneous cloud platforms based on Markov process modeling. IEEE Access, 10, 26149–26162.
55. SouravlasAnon. , Anon. , & Kostoglou, I. (2023a). A novel method for general hierarchical system modeling via colored Petri nets based on transition extractions from real data sets. Applied Sciences, 13, 339.

Augmented Pedagogy in Special Education: Integrating AI, Robotics, and Adaptive Technologies for Inclusive Learning

Spyridon Kontis[1[0009-0009-7892-6966]], Lambrini Seremeti[2[0000-0002-0663-5408]], and Sofia Anastasiadou[1[0000-0001-6404-5003]]

[1] University of Western Macedonia, Greece,
[2] Agricultural University of Athens, Greece,
dmw00034@uowm.gr, lseremeti@aua.gr, sanastasiadou@uowm.gr

Abstract. The rapid advancement of emerging educational technologies has opened new avenues for inclusion and innovation in special education. This paper presents findings from a mixed-methods study conducted in Cyprus, focusing on the integration of Artificial Intelligence (AI), educational robotics, virtual/augmented reality (VR/AR), and adaptive learning software in classroom environments supporting students with special educational needs (SEN). Drawing on both qualitative and quantitative data collected from over 120 participants, including teachers, parents, and learners, the research investigates the pedagogical impact, user acceptance, and challenges of deploying such technologies in real-world educational settings [1, 3].

Observational data and structured interviews highlight how these tools foster motivation, autonomy, and engagement, particularly among students with learning disabilities, autism spectrum conditions, or visual/hearing impairments. The study further identifies barriers such as lack of teacher training, infrastructural constraints, limited digital content in native languages, and the need for ethical considerations when implementing AI-driven tools [4–6].

To address these gaps, the paper proposes the “Assistive EdTech Adoption Framework”, a phased roadmap for schools aiming to deploy inclusive educational technologies aligned with the Universal Design for Learning (UDL) model [3, 10]. This framework emphasizes scalable implementation, interdisciplinary collaboration, and alignment with the European Digital Competence Framework (DigCompEdu). The paper concludes with policy and practice recommendations to enhance digital equity and empower teachers in the use of smart educational ecosystems [13], as illustrated in Figure 1.

Keywords: Special education, Artificial Intelligence, Educational robotics, VR/AR, Inclusive learning, Assistive technologies, DigCompEdu, Universal Design for Learning.

1 Introduction

Inclusive education has become a fundamental principle for 21st-century schooling, calling for flexible, responsive, and highly personalized teaching approaches that address the diverse needs of all learners. Across the globe, policymakers and educators

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.139K
Part of ISBN: 978-86-89755-37-4

are striving to remove barriers and promote equitable opportunities for students with disabilities or additional learning needs.

In this context, the rapid advancement of artificial intelligence (AI), educational robotics, and immersive technologies such as virtual and augmented reality (VR/AR) has opened new possibilities for tailoring pedagogy and curriculum to the unique profiles of each learner. These tools can enable more adaptive, interactive, and data-informed instruction, helping teachers to better support cognitive, emotional, and social development [4],[5].

While the theoretical promise of technology-driven inclusion is well documented, real-world application remains complex. This study explores how AI-powered and digital tools are currently being used in special education settings in Cyprus, examining both their practical benefits and the challenges of scaling such innovations. The research aims to provide insights for practitioners and policymakers interested in maximizing the positive impact of technology in inclusive classrooms [1],[3].

2 Theoretical Background

The field of special education increasingly relies on a broad spectrum of technological solutions designed to meet the individualized needs of students with diverse abilities and learning profiles. Effective inclusion requires tools that facilitate sensory access, enable differentiated instruction, and promote learner autonomy, ensuring that every student can participate meaningfully and achieve their full potential within the learning process [4], [5], [6].

Artificial intelligence (AI) offers significant capabilities in personalizing content delivery, assessment, and feedback, adapting resources dynamically to the learner's abilities, pace, and evolving progress [4], [10]. AI-powered analytics can identify learning gaps, recommend targeted interventions, and even predict potential learning difficulties, thus enabling preventive pedagogical measures. Similarly, virtual and augmented reality (VR/AR) environments provide multisensory, immersive experiences that enhance engagement, scaffold complex concepts, and create simulated scenarios for practicing real-life skills in safe, controlled environments. This is particularly valuable for students with attention disorders or sensory processing differences, who may benefit from tailored sensory inputs and adjustable interaction levels [5].

Educational robotics combines physical interaction with problem-solving and gamification, fostering not only academic learning but also social-emotional development, collaboration, and communication skills [6]. Robots can act as mediators in peer interactions, enabling inclusive participation for students with autism spectrum conditions or speech/language impairments.

This section draws upon international literature to review the pedagogical affordances of these technologies, highlighting both opportunities, such as enhanced engagement, personalization, and inclusivity, and potential pitfalls, including ethical considerations, accessibility barriers, and the digital divide. It also underscores the pivotal role of structured frameworks, notably Universal Design for Learning (UDL) [10] and the European Digital Competence Framework for Educators (DigCompEdu) [3], which provide evidence-based guidelines for integrating digital tools effectively into diverse educational contexts. However, despite promising results, gaps remain in understanding how these technologies can be sustainably scaled across diverse educational contexts, particularly in small or resource-constrained systems. Addressing these gaps is essential to inform policy, guide teacher training, and ensure that innovations are equitably accessible to all learners

3 Methodology

This study employed a convergent mixed-methods design, combining quantitative survey data with qualitative interviews and classroom observations to explore the pedagogical impact of emerging technologies in special education.

Participants The sample comprised 120 participants in Cyprus, including 85 educators (special and mainstream teachers) and 35 parents of students with special educational needs (SEN). The educators represented primary and secondary school levels, with teaching experience ranging from less than 5 years to more than 20 years. Parents came from diverse socioeconomic backgrounds, reflecting the heterogeneity of the SEN community.

Instrument Data were collected using a structured questionnaire developed and validated during a previous MSc research project (Kontis, 2024). The instrument consisted of 32 closed-ended items organized into six thematic domains:

1. Perceived effectiveness of technology (e. g. , “The use of AI tools improves my students’ learning outcomes”),
2. Ease of use and accessibility,
3. Impact on motivation and engagement,
4. Support for autonomy and self-regulation,
5. Collaboration and social interaction,
6. Barriers and challenges (e. g. , cost, lack of training, language resources).

All items were rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Example items included: “VR/AR environments help students with learning disabilities engage more actively in lessons” and “Robotics activities improve peer collaboration among students with autism. ”

The questionnaire was subjected to expert validation by three academics in educational technology and special education, ensuring content validity and internal coherence.

Qualitative component To complement survey data, 12 semi-structured interviews were conducted with teachers, alongside 6 non-participant classroom observations. The interviews explored practical experiences, benefits, and difficulties in implementing technologies such as AI, robotics, and VR/AR. Observations documented student engagement, autonomy, and social interaction patterns.

Procedure Surveys were administered both online and in print form between February and May 2024. Interviews lasted 30–45 minutes and were audio-recorded with participant consent. Observations took place in inclusive classrooms during normal teaching hours.

Data Analysis Quantitative data were analyzed using SPSS v. 28, applying descriptive statistics (means, standard deviations, and frequency distributions) and cross-tabulations by participant group (teachers vs. parents). Qualitative data were analyzed thematically following Braun & Clarke's (2006) six-step framework, enabling identification of recurring themes and triangulation with quantitative results

4 Results

The quantitative findings revealed clear trends regarding the pedagogical impact of emerging technologies in special education. Descriptive statistics were computed in SPSS (v. 28), including frequencies, percentages, and means across the main domains of the instrument

Table 1. Perceived impact of emerging technologies on student outcomes (N = 120)

Indicator	% of respondents reporting improvement
Student engagement	82%
Task completion rates	74%
Motivation and emotional confidence	67%
Peer collaboration and interaction	71%

These results suggest that teachers and parents perceived significant benefits in terms of engagement, collaboration, and motivation, particularly for students with learning disabilities and autism spectrum conditions.

Table 2. Reported barriers to implementation (N = 120)

Barrier	% of respondents identifying barrier
Lack of teacher training	63%
Financial costs of equipment	58%
Limited digital resources in Greek	41%
Infrastructure/technical support	37%

The most frequently cited barrier was insufficient teacher training, followed by the cost of acquiring and maintaining advanced technologies.

Qualitative insights further enriched these findings. Teachers described how robotics facilitated social interaction among students with autism, while adaptive reading software supported learners with dyslexia in overcoming literacy challenges. As one participant noted: "For the first time, I saw my student initiate a conversation with a peer without prompting. " These narratives highlight how emerging technologies not only improved academic outcomes but also contributed to autonomy, confidence, and social inclusion.

Teachers emphasized that these technologies fostered greater autonomy, self-advocacy, and metacognitive awareness in students, while also stressing the importance of ongoing guidance, culturally relevant content, and personalized instructional approaches [7], [11].

5 Discussion

The findings of this study highlight the transformative potential of emerging technologies in special education, while also exposing critical systemic barriers that must be addressed for sustainable impact. When effectively integrated, AI, robotics, VR/AR, and adaptive software foster engagement, autonomy, and social participation among students with special educational needs [4], [5], [17]. The evidence presented supports the principles of Universal Design for Learning (UDL) and the DigCompEdu framework, demonstrating how structured approaches can guide inclusive digital pedagogy [2], [8], [15].

At the same time, barriers such as insufficient teacher training, financial costs, and lack of accessible resources in native languages remain significant challenges [3], [6]. Addressing these requires a whole-school approach, combining long-term professional development, collaborative design of lessons, and sustained infrastructural support.

From a policy perspective, ministries and school leaders should prioritize strategic investment in digital capacity-building, ensuring that educators are equipped with both the technical and pedagogical skills required to implement new tools effectively [11], [19]. Pilot projects and co-design initiatives involving teachers, learners, and families can further ensure that innovations reflect real classroom needs.

Finally, embedding AI ethics throughout the process, from system design to classroom practice, is essential to safeguard transparency, privacy, and fairness [9], [10], [14]. Only through ethically responsible and pedagogically grounded integration can emerging technologies contribute to sustainable inclusion and equity in education.

6 Proposed Framework: Assistive EdTech Adoption

Based on the findings and identified needs, this study proposes a three-phase model for the structured and ethical adoption of assistive educational technologies in special education (Figure 1). The framework was developed by the authors and is grounded in the principles of Universal Design for Learning (UDL) and the European Digital Competence Framework for Educators (DigCompEdu) [2], [8], [15].

Phase 1 – Exploration. Schools begin with low-cost or freely available tools (e. g. , text-to-speech software, browser-based accessibility extensions). The focus is on ease of use, inclusivity, and minimal training, encouraging educators to take first steps toward digital inclusion.

Phase 2 – Integration. As familiarity increases, more advanced solutions are introduced, such as educational robotics, immersive VR/AR scenarios, and adaptive learning management plugins for SEN learners. Teacher training and peer collaboration are emphasized, ensuring that technology use aligns with pedagogical objectives.

Phase 3 – Optimization. Institutions adopt AI-driven personalization systems, wearable sensors, and real-time analytics to tailor interventions and predict learning needs. At this stage, explicit integration of AI ethics—transparency, privacy, and fairness—is required to guarantee safe, equitable, and sustainable adoption.

This phased roadmap promotes a gradual, scalable, and pedagogically aligned transition, allowing both educators and students to thrive in inclusive and ethically responsible digital ecosystems.

7 Conclusions and Recommendations

The findings of this study underscore the urgent need for well-structured policies that guide the effective integration of assistive technologies in special education. Successful implementation requires more than the deployment of tools, it demands a comprehensive strategy encompassing professional development, inclusive instructional design, and sustained support mechanisms.

Policymakers, school leaders, and education ministries should prioritize investment in teacher training programs, co-design initiatives involving both educators and learners, and small-scale pilot projects that reflect the realities of inclusive classrooms. These actions should be aligned with established digital competence frameworks and accessibility standards.

Future research should investigate the long-term impact of these technologies on learner outcomes, comparing implementation strategies across diverse educational systems and cultural contexts. Studies should also examine the factors that influence successful adoption, including teacher readiness, resource availability, and ethical considerations, particularly in relation to AI ethics, to ensure equitable and sustainable integration.

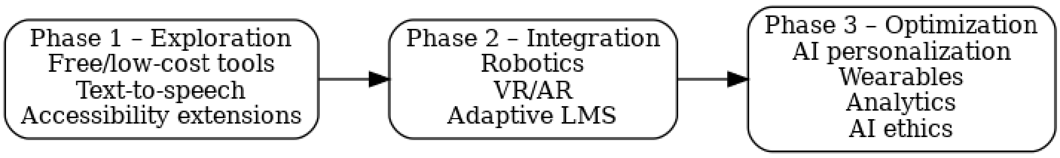


Fig. 1. Assistive EdTech Adoption Framework (Authors' own work)

.

Fig. 2. AR tool used in inclusive classroom context (Source: Pixabay, free to use under Pixabay License)

Disclosure of Interests.

The authors declare that they have no competing interests relevant to the content of this article.

References

1. Kontis, S. (2024). MSc Thesis: Integration of New Technologies in Special Education. Frederick University, Nicosia.
2. European Commission. (2017). DigCompEdu: European Framework for the Digital Competence of Educators. Publications Office of the European Union, Luxembourg.
3. UNESCO. (2022). AI and Inclusion in Education. UNESCO Publishing, Paris.
4. Fung, K. Y. , Fung, K. C. , Lui, T. L. R. , et al. (2025). Exploring the impact of robot interaction on learning engagement: a comparative study of two multi-modal robots. Smart Learning Environments, 12, Article 12.
5. CAST. Center for Applied Special Technology. (accessed: 11 Aug 2025)
6. AEM Center. National Center on Accessible Educational Materials. (accessed: 11 Aug 2025)
7. ISTE. International Society for Technology in Education. (accessed: 11 Aug 2025)
8. Al-Azawei, A. , Serenelli, F. , Lundqvist, K. (2016). Universal Design for Learning (UDL): A content analysis of peer-reviewed journal papers from 2012 to 2015. Journal of the Scholarship of Teaching and Learning, 16(3), 39–56.
9. Anastasiadou, S. D. (2026). Exploring AI ethics through factorial correspondence analysis. In S. Anastasiadou, T. Chadjipadelis, N. Koutsoupias, A. Markos, & G. Menexes (Eds.), Modern International Developments in Data Analysis and Multivariate Statistical Analysis (pp. 495–528). IGI Global Scientific Publishing.
10. Seremeti, L. , Anastasiadou, S. , Masouras, A. , Papalexandris, S. (2025). Addressing the challenges of harmonizing law and artificial intelligence technology in modern society.

IAES International Journal of Artificial Intelligence (IJ-AI), 14(3), 2471–2478. https://doi. org/10. 11591/ijai. v14. i5. pp2471-2478

11. Seremeti, L. , Anastasiadou, S. (2025). Dealing with semantic inconsistencies in AI governance. In S. Anastasiadou, L. Seremeti, A. Masouras, & L. Anastasiadis (Eds.), Modern Perspectives on Artificial Intelligence and Law (pp. 125–134). IGI Global Scientific Publishing.
12. Anastasiadou, S. , Seremeti, L. (2025). Autopoietic co-evolution of AI and law. In S. Anastasiadou, L. Seremeti, A. Masouras, & L. Anastasiadis (Eds.), Modern Perspectives on Artificial Intelligence and Law (pp. 81–88). IGI Global Scientific Publishing.
13. Kalogera, C. , Seremeti, L. , Kontis, S. , Anastasiadou, S. (2025). The digital DNA of modern workforce. In L. Seremeti, P. Liargovas, L. Anastasiadis, & S. Anastasiadou (Eds.), Harnessing Business Intelligence for Modern Talent Management (pp. 379–390). IGI Global.
14. Kontis, S. , Anastasiadou, S. , Anastasiadis, L. , Seremeti, L. (2025). Legal entropy in AI governance. In S. Anastasiadou, L. Seremeti, A. Masouras, & L. Anastasiadis (Eds.), Modern Perspectives on Artificial Intelligence and Law (pp. 193–206). IGI Global Scientific Publishing.
15. Rose, D. H. , Meyer, A. (2006). A Practical Reader in Universal Design for Learning. Harvard Education Press, Cambridge.
16. World Health Organization. (2011). World Report on Disability. WHO Press, Geneva.
17. Zydney, J. M. , Bathke, K. , Hasselbring, T. (2014). Developing technological pedagogical content knowledge (TPACK) of pre-service teachers for inclusive education. Journal of Special Education Technology, 29(2), 1–12.
18. Agorogianni, A. Z. , Zaharis, Z. D. , Anastasiadou, S. D. , Goudos, S. K. (2011). Distance learning technology and service support in Greece: the case study of the Aristotle University over the last decade. Education and Information Technologies, 16(1), 25–39.
19. Seremeti, L. , Anastasiadis, L. , Anastasiadou, S. (2024). The role of data in environmental, social, and governance strategy. In C. Papademetriou, S. Anastasiadou, K. Ragazou, A. Garefalakis, & S. Papalexandris (Eds.), ESG and Total Quality Management in Human Resources (pp. 201–218). IGI Global Scientific Publishing.
20. Souravlas, S. , Anastasiadou, S. (2021). Pipelined dynamic scheduling of big data streams. Applied Sciences, 10(14), 4796.
21. Souravlas, S. , Anastasiadou, S. (2022). On implementing social community clouds based on Markov models. IEEE Transactions on Computational Social Systems. https://doi. org/10. 1109/TCSS. 2022. 3213273
22. Souravlas, S. , Anastasiadou, S. , Katsavounis, S. (2021a). More on pipelined dynamic scheduling of big data streams. Applied Sciences, 11(1), 61.
23. Souravlas, S. , Anastasiadou, S. , Katsavounis, S. (2021b). A survey on the recent advances of deep community detection. Applied Sciences, 11(16), 7179.
24. Tantalaki, N. , Souravlas, S. , Roumeliotis, M. , Katsavounis, S. (2020). Pipeline-based linear scheduling of big data streams in the cloud. IEEE Access, 8, 117182–117202. https://doi. org/10. 1109/ACCESS. 2020. 3004612
25. Souravlas, S. , Katsavounis, S. , Anastasiadou, S. (2021c). On modeling and simulation of resource allocation policies in cloud computing using colored Petri nets. Applied Sciences, 10(16), 5644.
26. Souravlas, S. , Anastasiadou, S. , Tantalaki, N. , Katsavounis, S. (2022). A fair, dynamic load balanced task distribution strategy for heterogeneous cloud platforms based on Markov process modeling. IEEE Access, 10, 26149–26162.
27. Souravlas, S. , Anastasiadou, S. , Kostoglou, I. (2023a). A novel method for general hierarchical system modeling via colored Petri nets based on transition extractions from real data sets. Applied Sciences, 13, 339.

28. Souravlas, S. , Anastasiadou, S. , Economides, T. , Katsavounis, S. (2023b). Probabilistic community detection in social networks. IEEE Access, 11, 25629–25641. https://doi. org/10. 1109/ACCESS. 2023. 3257021

Examining Axiological Assumptions in Machine Learning Publications

Yashpreet Malhotra[1][0009-0009-1561-616X]

[1] University of New Haven, New Haven, Connecticut 06516, USA
yashmalhotra9323@gmail.com

Abstract. This paper presents a study of the values embedded within machine learning research papers. A novel annotation scheme is developed to analyze how values are represented in scholarly documents, focusing on the rationales for research projects, the emphasized attributes of those projects, and the discussion or neglect of potential negative impacts. The method- ology is applied to a corpus of influential papers from top- tier machine learning conferences. The analysis explores the relationship between these encoded values and factors such as institutional affiliations and funding sources, aiming to contribute to a more nuanced understanding of the ethical dimensions of machine learning research.

Keywords: Machine Learning Research, Scientific Values, Ethical Analysis, Research Rationales, Negative Impacts, Institutional Affiliations, Funding Sources, scholarly discourse.

1 Introduction

1.1 The Hidden Values shaping Machine Learning Research

Over the past decade, machine learning (ML) has evolved from a niche academic discipline into a cornerstone of technological advancement, influencing a wide array of industries including healthcare, finance, transportation, and social media. Its rapid proliferation has been accompanied by increasing public attention, government investment, and corporate inter- est. As ML systems become more deeply integrated into the fabric of society, concerns surrounding their ethical, social, and political implications have gained prominence. However, despite these growing concerns, mainstream ML research often continues to treat such issues as peripheral rather than central to the development process. [1], [7]

A dominant narrative persists within the research community that presents ML development as an inherently technical and neutral endeavor—driven by objectivity, performance benchmarks, and empirical rigor. This framing, while useful for fostering scientific progress, tends to obscure the fact that every design choice, evaluation criterion, and research focus reflects an underlying set of values. The choices researchers make—what problems to tackle, which metrics to optimize, whose data to use, and what outcomes to prioritize—are never fully divorced from societal context. [4].

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.148M
Part of ISBN: 978-86-89755-37-4

Yet, these considerations are frequently underexplored in published literature, which often defaults to a narrow set of values such as accuracy, scalability, and novelty.
This study aims to interrogate the implicit value systems embedded in the corpus of ML research. Specifically, it seeks to uncover which attributes are consistently praised, which concerns are systematically excluded, and how these patterns reflect broader institutional forces shaping the field. By conducting a qualitative and thematic analysis of 100 highly cited papers published at top-tier venues—specifically the International Conference on Machine Learning (ICML) and the Conference on Neural Information Processing Systems (NeurIPS)—between 2008 and 2019, this work offers a critical lens into the internal logic of the ML research ecosystem.
These papers were selected not only for their influence, as measured by citation count, but also for their role in setting research agendas and defining norms within the field. The analysis investigates the extent to which the values promoted in these works align with broader societal goals, and how institutional affiliations—whether academic, corporate, or hybrid—affect the framing of research contributions. In doing so, this study contributes to ongoing conversations about the social responsibility of ML research and calls for a more reflective and inclusive approach to innovation in the field. [6], [12]

2 Methodology

To systematically investigate the implicit and explicit value structures embedded within machine learning research, we adopted a multi-phase interpretive methodology grounded in both qualitative analysis and rigorous validation protocols. This hybrid approach enabled a nuanced examination of how research in ML is framed, what kinds of justifications are offered for methodological choices, and which societal considerations are emphasized, downplayed, or omitted altogether. [3]

2.1 Corpus Selection

The dataset consisted of 100 high-impact research papers selected from two of the most prestigious venues in the field of machine learning: the International Conference on Machine Learning (ICML) and the Conference on Neural Information Processing Systems (NeurIPS). These conferences were chosen for their longstanding influence and for setting research trends within the global ML community. Papers were drawn from four strategically chosen years—2008, 2009, 2018, and 2019—thereby capturing both the early development of modern ML techniques and their more recent, industrially-integrated manifestations. Citation count was used as a proxy for impact, ensuring that the papers analyzed were widely read and influential in shaping discourse.

2.2 Annotation Framework and Process

The textual content of each paper was segmented at the sentence level across four core sections: Abstract, Introduction, Discussion, and Conclusion. These sections were chosen because they encapsulate the narrative arc of the paper—from motivation and framing to the articulation of contributions and broader implications.

A team of trained annotators conducted a detailed manual coding of each sentence using a hybrid coding framework that combined deductive and inductive strategies. The deductive component was informed by existing literature on responsible AI and technology ethics, focusing on established normative categories such as fairness, transparency, accountability, safety, and social benefit. This provided a structured lens through which to assess the explicit values referenced in each paper.

In parallel, the inductive approach allowed for the emergence of novel themes not predefined in the coding schema. This was essential for capturing domain-specific justifications or subtle rhetorical strategies that may reflect implicit values. Annotators were instructed to identify and classify sentence- level value statements according to several dimensions: the type of value elevated (e.g., novelty, performance, efficiency), the presence or absence of societal impact acknowledgment, and the rhetorical justification strategies employed (e.g., ap- peals to utility, objectivity, scalability). [4]

2.3 Validation and Reliability Measures

To ensure analytical rigor and mitigate subjectivity, 40% of the dataset underwent dual annotation, followed by a recon- ciliation process to address discrepancies. The inter-annotator agreement was assessed using both raw percentage agreement and statistical metrics. An overall agreement rate of 87% was achieved, demonstrating a high level of consistency in the identification and categorization of value-laden statements.

In addition, a fuzzy Fleiss' kappa score of 0.45 was computed, suggesting moderate agreement across multiple annotators for categorical variables. For ordinal variables such as the degree of societal justification and acknowledgment of potential harms, a weighted Fleiss' kappa exceeding 0.6 was recorded, indicating substantial inter-rater reliability. These validation metrics reinforce the reliability of the annotation framework and support the credibility of subsequent analyses. [12]

2.4 Institutional Attributional Analysis

Beyond textual analysis, each paper was reviewed for meta- data relating to institutional affiliation and funding disclosures. Authors' primary affiliations were classified as academic, corporate, or hybrid (e.g., university-corporate collaborations). Publicly disclosed funding sources were also documented, with particular attention paid to corporate sponsorship or government grants. This layer of analysis was designed to examine whether institutional context influences the values promoted in ML research—either directly through funding priorities or indirectly through organizational norms and incentives.

By triangulating content-level findings with metadata on institutional context, this methodology offers a comprehensive lens through which to understand the value orientations embedded in mainstream ML research. [18]

3 Quantitative Results

This section presents the quantitative findings derived from the analysis of 100 highly cited ML papers. The results are organized around key axes of ethical and societal engagement: justification of research relevance, acknowledgment of potential harms, value prioritization, and institutional trends.

3.1 Justification Distribution

To assess the extent to which authors justify the broader relevance of their research, each paper was classified into one of four categories based on the nature and depth of its societal justification.

Table 1. Classification of Justification Strategies in Analyzed Papers

Justificatory Classification level	Proportion of Papers
No Mention of Societal Relevance	68%
Stated but Unjustified Societal Link	17%
Minimal Societal Justification	11%
Detailed Societal Justification	4%

As shown in Table I, a striking 68% of the analyzed papers made no reference to the societal implications or applications of their work. An additional 17% included vague or symbolic gestures toward societal relevance (e.g., claims of ”real-world applicability”) without substantiating such claims with detailed argumentation or evidence. Only 4% of papers offered a robust societal rationale—highlighting a significant gap in ethical contextualization across the literature.

3.2 Negative Implication Discussion

A separate coding exercise investigated the degree to which authors engaged with potential negative outcomes or ethical risks associated with their contributions.

Table 2. Extent of negative Impact Consideration in Sampled Literature

Engagement with Potential Harm	Proportion of Papers
No Recognition	98%
Brief Mention of Possible Harms	1%
Substantive Risk Discussion	1%
In-Depth Harm Analysis	0%

Table II reveals that nearly all papers (98%) omitted any discussion of potential negative consequences arising from the proposed methodologies. Only 2 papers mentioned possible harms, and even these were restricted to cursory or high-level remarks. Not a single paper in the sample engaged in a deep or systematic exploration of unintended impacts—highlighting a profound asymmetry between technical enthusiasm and ethical foresight. [3]

3.3 Figures and Interpretations

Figure 1 illustrates the dominant values emphasized across the dataset. Performance metrics such as accuracy, precision, and F1 score were by far the most cited benchmarks of success, followed closely by generalization ability and computational efficiency. Conversely, socially relevant values such as fairness, interpretability, and safety were mentioned infrequently, indicating a skew toward optimization-centric paradigms.

Figure 2 highlights a notable temporal shift in the institutional makeup of ML authorship. Papers published in 2008 and 2009 were predominantly authored by researchers from academic institutions. However, by 2018 and 2019, the presence of large technology companies—particularly U.S.-based firms such as Google, Facebook, and Microsoft—had surged. This trend reflects the growing influence of corporate actors in setting research agendas, potentially reinforcing a value system oriented toward scalability, commercial applicability, and proprietary advantage.

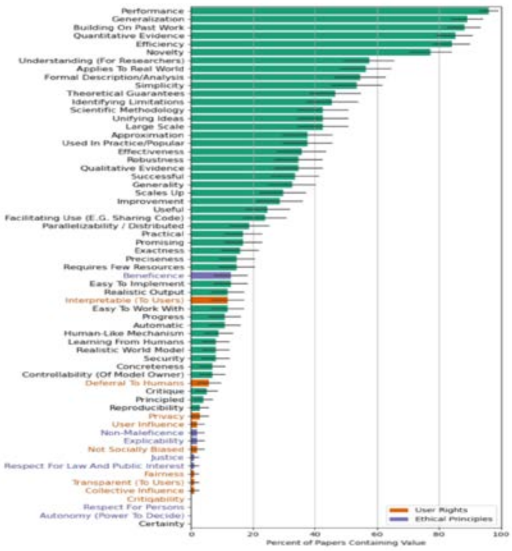


Fig. 1: Frequency distribution of top values (e.g., performance, generalization, efficiency) across reviewed papers.

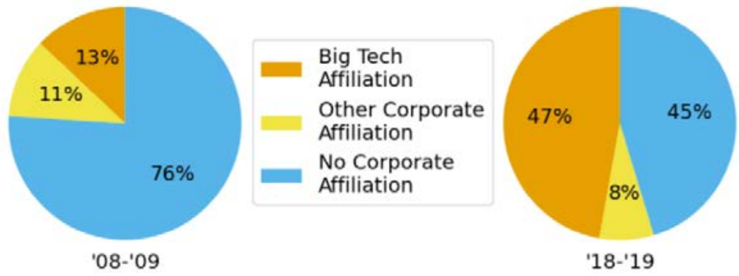


Fig. 2: Temporal shift in paper affiliations: Increased prevalence of Big Tech and elite academic institutions over time.

4 Value Analysis

A thematic exploration of the annotated corpus revealed six dominant values recurrently emphasized in the reviewed ML research: *Performance*, *Generalization*, *Efficiency*, *Building on Prior Work*, *Quantitative Evidence*, and *Novelty*. These were not only frequently mentioned but often operationalized in ways that prioritized technical benchmarks over social impact or inclusivity.

4.1 Performance

Performance was the most cited metric of value, often ex- pressed through improvements in accuracy or superiority over existing methods. This emphasis generally revolved around benchmark datasets and quantitative evaluation metrics such as accuracy, F1 score, or mean squared error. Yet, such evaluations often lacked reflection on the limitations of the metrics used, particularly their inability to capture broader ethical or contextual concerns. Most papers treated accuracy as a proxy for progress, without discussing what such ”improvement” means for real-world stakeholders.

4.2 Generalization

Generalization was primarily framed as the capability of a model to perform consistently across unseen data or varying tasks. However, this concept had almost always evaluated using multiple curated datasets rather than deployment scenarios. Rarely did authors reflect on the sociotechnical implications of transferring models across domains, such as the potential amplification of biases or the ethical concerns of applying ML systems without contextual calibration.

4.3 Efficiency

Efficiency was commonly understood as reduced computational cost, memory usage, or training time. Paradoxically, many papers used the term ”efficient” to indicate the ability to scale to massive data or models—essentially enabling high- resource operations rather than reducing resource consumption. Energy use, environmental impact, or the democratization of access were virtually absent from the discussion, implying that the term “efficiency” primarily served large institutions with abundant resources.

4.4 Building on Prior Work and Novelty

Most papers sought to balance novelty with continuity by demonstrating how their contributions extended or refined existing methods. Novelty was typically framed in algorithmic or architectural terms, while social innovation or problem recontextualization was nearly invisible. Even when prior limitations were mentioned, they were addressed strictly from a performance lens, not through broader critical analysis. [1]

5 Institutional and Corporate Influence

A detailed examination of author affiliations and funding disclosures across the sampled corpus reveals a pronounced concentration of influence among elite academic institutions and major technology corporations. These institutional forces play a pivotal role in shaping not only the direction of research but also the values and priorities embedded within the ML literature. Over the span from 2008 to 2019, this influence intensified, signaling broader structural shifts in the machine learning research ecosystem. [18]

5.1 Authorship Trends

Analysis of authorship patterns across the selected pa- pers demonstrates a marked increase in contributions from corporate-affiliated researchers, particularly those associated with multinational technology companies such as Google, Microsoft, Facebook, and Amazon. Between 2008–2009 and 2018–2019, the proportion of papers with at least one author affiliated with a corporate entity nearly tripled. This rise coincided with a growing dominance of elite academic institutions, including but not limited to Stanford, MIT, Carnegie Mellon, and UC Berkeley.

In contrast, the representation of authors from smaller universities, non-Western institutions, and under-resourced aca- demic settings remained minimal to negligible. This stratification reflects a consolidation of research capital within a narrow band of institutions capable of providing substantial re- sources, infrastructure, and visibility—

factors that significantly increase the likelihood of acceptance at top-tier conferences and high citation impact.

The increasing prevalence of industry-academic collaborations also merits attention. While such partnerships can yield powerful synergies, they often risk aligning academic inquiry with commercial imperatives, particularly when corporate entities contribute funding, data, or compute resources that are inaccessible to independent researchers.

5.2 Funding Patterns

A complementary analysis of funding acknowledgments (where disclosed) reveals a similarly skewed landscape. A substantial proportion of the papers did not explicitly state their sources of financial support. However, through cross- referencing institutional affiliations and known partnerships, we inferred that corporate backing—either directly or indirectly—played a significant role in supporting many of the most-cited works.

Among papers that did disclose funding information, industry sponsors were disproportionately represented. Funding from companies involved in the development and deployment of machine learning systems—especially those with commercial interests in scalability, speed, and competitive advantage—was widespread. This trend suggests a potential narrowing of research agendas, wherein work that aligns with deployable, monetizable outcomes is privileged over research that interrogates ethical trade-offs, social impacts, or long-term risks.

Moreover, the underreporting of funding sources further obscures the extent of corporate influence in shaping the knowledge landscape. This lack of transparency hinders the ability of external observers to critically assess how financial incentives may affect methodological choices, problem framing, or value prioritization in published research.

5.3 Implications for Research Diversity

The concentration of authorship and funding within a small set of elite and corporate institutions has significant implications for epistemic diversity and agenda-setting in machine learning. When research is disproportionately driven by entities with aligned economic or geopolitical interests, the scope of inquiry may contract, prioritizing performance improvements and technical optimization over critical reflection and societal responsiveness.

To foster a more inclusive and balanced research environment, increased attention must be given to diversifying funding pathways, promoting equitable authorship opportunities, and establishing norms around full disclosure of institutional and financial affiliations. Such measures are essential for ensuring that the field evolves not only in terms of technical sophistication, but also in its ethical maturity and global inclusivity.

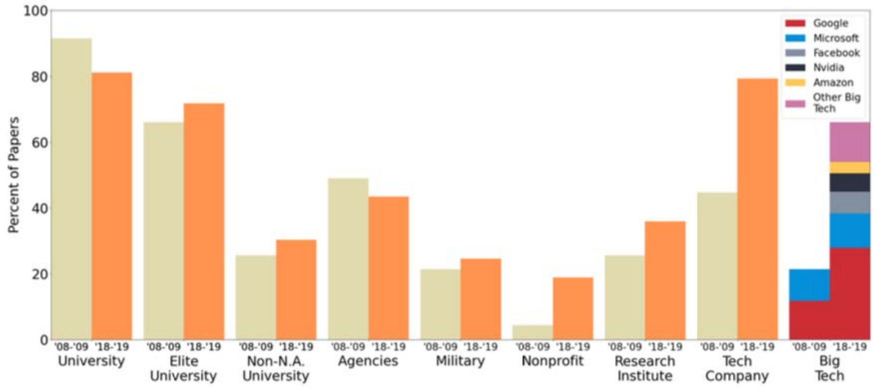


Fig. 3: Corporate and institutional ties of high-impact papers over time.

6 Discussion

The findings of this study offer a substantive challenge to the prevailing narrative that machine learning research is inherently objective, apolitical, or universally beneficial. Instead, our analysis reveals a pronounced value orientation within highly cited ML literature, one that privileges technical metrics, institutional prestige, and deployability, often at the expense of ethical reflection, inclusivity, and social accountability. [14]

One of the most striking patterns observed is the near-total omission of ethical considerations, even in papers addressing high-stakes domains such as healthcare, finance, or surveillance. Core ethical principles fairness, transparency, accountability, autonomy, and human dignity, are rarely acknowledged, let alone meaningfully integrated into the research frameworks. This omission is not merely a matter of oversight but reflects an implicit prioritization of values like performance, novelty, and scalability, which are more easily measured and often more directly aligned with corporate or academic incentives. [3]

While performance metrics such as accuracy, F1 score, or inference time offer convenient benchmarks for progress, their dominance in justifying research contributions may obscure the contextual complexity of real-world deployment. A model achieving state-of-the-art results on benchmark datasets may still reproduce harmful biases, fail under domain shifts, or exacerbate existing inequalities in practical applications. Yet few papers engaged with these risks, and even fewer offered mechanisms for identifying or mitigating them. This signals a systemic undervaluation of harm-centred perspectives in the current ML publication ecosystem. [13]

Furthermore, the growing centrality of elite institutions—both academic and corporate in shaping the research discourse raises concerns about epistemic homogenization and agenda capture. As our institutional analysis indicates, the most visible and impactful research increasingly originates from a concentrated set of organizations with aligned strategic interests. While such institutions undoubtedly contribute valuable resources and infrastructure, their dominance risks marginalizing alternative research paradigms—particularly those emerging from underrepresented regions, community-based organizations, or disciplines outside the core ML com- munity.

This concentration of authorship and funding may also shape the field's incentive structures. Research that advances the goals of commercial scalability, market integration, or technological innovation is often rewarded, while critical, interpretive, or justice-oriented work may struggle for recognition and support. As a result, the boundaries of "valuable" research are defined not merely by scientific rigor but by institutional alignment with dominant economic and political interests.

The absence of diverse voices—including those most affected by the deployment of ML systems—further compounds the problem. Without intentional inclusion of marginalized perspectives, the field risks reifying inequities under the guise of technical progress. Democratizing ML research will there- fore require more than open-source code or broader dataset access; it will necessitate structural changes in publication practices, funding distribution, and community norms that currently privilege a narrow set of actors and values.

Ultimately, these findings underscore the need for a more reflective and inclusive machine learning research culture—one that critically interrogates its assumptions, explicitly acknowledges its limitations, and centers societal well-being as a primary criterion for progress. Future research must work toward rebalancing technical innovation with ethical deliberation, ensuring that the benefits of ML are both equitable and just.

7 Conclusion

This study highlights the inherently normative character of contemporary machine learning research, directly challenging the widespread assumption of its neutrality or objectivity.

Through a systematic content analysis of 100 highly cited papers published between 2008 and 2019 in premier ML venues, we uncovered a consistent and reinforcing pattern: values such as performance optimization, computational efficiency, and generalizability are overwhelmingly prioritized, while ethical, societal, and user-centered considerations are notably marginalized.

These dominant value preferences are not incidental. They are embedded within broader institutional structures, particularly the growing entanglement of elite academic institutions and corporate technology firms, which significantly influence what research is conducted, how it is evaluated, and whose interests it ultimately serves. The result is a research culture where social accountability, transparency, fairness, and harm reduction are treated as peripheral, rather than foundational, to technical progress. [18]

Equally concerning is the near-total absence of rigorous engagement with the potential risks and unintended consequences of ML systems, even in domains where such harms are well-documented. This lack of critical discourse not only weakens the field's ability to safeguard against misuse but also erodes public trust in machine learning technologies.

To address these shortcomings, a redefinition of success within the ML research community is urgently needed. Progress must no longer be equated solely with surpassing benchmark datasets or publishing in prestigious venues. In- stead, evaluation criteria should be broadened to include the social utility, contextual relevance, and ethical integrity of proposed methods. Incentive structures—such as peer review norms, funding priorities, and institutional recognition—must evolve accordingly.

Furthermore, cultivating an inclusive and pluralistic research environment will require intentional efforts to elevate voices and perspectives that have historically been excluded from the ML discourse. This includes researchers from underrepresented regions, disciplines focused on critical theory or social justice, and communities directly impacted by algorithmic systems.

In conclusion, steering machine learning toward a more socially responsible trajectory is not merely a matter of technical refinement. It is a political, institutional, and cultural under- taking—one that demands collective commitment to equity, accountability, and shared human flourishing. Only through such a deliberate shift can ML fulfill its potential as a force for inclusive and ethical innovation.

8 Disclosure of Interest

There was no use of Generative AI to make this research paper or to make this research successful. The author takes the full responsibility for the publication's content.

References

1. S.H.AlHarbi,L.N.Tidjon,andF.Khomh,"Responsible Design Patterns for Machine Learning Pipelines," *arXiv preprint*, May 2023.
2. T. LaCroix and S. J. D. Prince, "Deep Learning and Ethics," in *Under- standing Deep Learning*, arXiv, May 2023.
3. T. K. Gilbert, M. W. Brozek, and A. Brozek, "Beyond Bias and Compliance: Towards Individual Agency and Plurality of Ethics in AI," *arXiv preprint*, Feb. 2023.
4. W. Ma and V. Valton, "Toward an Ethics of AI Belief," *arXiv preprint*, Apr. 2023.
5. M. *, "Mapping the Ethics of Generative AI: A Comprehensive Scoping Review," *Minds and Machines*, 2024.
6. J. Juli, "Ethical Considerations in Artificial Intelligence and Machine Learning," *EasyChair Preprint 12536*, Mar. 2024.

7. R. K. Paul and B. Sarkar, “Generative AI and Ethical Considerations for Trustworthy AI Implementation,” *IJAIML*, vol. 2, no. 1, 2023.
8. S. N. Halder and S. Sarkar, “Ethical application of artificial intelligence and machine learning in research and education,” in *Academic Integrity and Innovation*, Prova Prakashani, 2024.
9. N.HussainandA.Lee,“EthicalConsiderationsinArtificialIntelligence and Machine Learning,” *J. Nonlinear Analysis and Optimization*, vol. 14, no. 1, 2023.
10. Z. Hamid, S. Ajmal, and I. Torshin, “Ethical Considerations in AI and Machine Learning,” ResearchGate, Nov. 2023.
11. S. T. Boppiniti, “Data Ethics in AI: Addressing Challenges in Machine Learning and Data Governance for Responsible Data Science,” *Int. Scientific J. for Res.*, vol.5, no.5, 2023.
12. B. Memarian and T. Doleck, “Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review,” *Computers and Education: Artificial Intelligence*, 2023.
13. S. Ruggieri *et al*., “Can We Trust Fair-AI?,” in *AAAI Conf. on Artificial Intelligence*, Jun. 2023.
14. Castelnovo *et al*., “Fair Enough? A map of the current limitations of the requirements to have ‘fair’ algorithms,” Commun. ACM, 2023.
15. M. R. Islam, “Generative AI, Cybersecurity, and Ethics,” John Wiley Sons, 2024.
16. Gazis *et al*., “Organising AI for safety: Identifying structural vul- nerabilities to guide the design of AI-enhanced socio-technical systems,” *Saf. Sci.*, Apr. 2025.
17. “Transparency and the Black Box Problem: Why We Do Not Trust AI,” *Philosophy Technology*, Dec. 2025.
18. B. Lund, “Standards, frameworks, and legislation for artificial intelli- gence (AI) transparency,” *AI and Ethics*, Jan. 2025.
19. Y. Zhang, C. Dong, W. Guo, J. Dai, and Z. Zhao, “Systems theoretic accident model and process (STAMP): A literature review,” *Saf. Sci.*, 2022–2023.

Prompt Engineering as an AI Literacy Competence: A Framework for Learners and Educators

Dijana Oreški[1] [0000-0002-3820-0126], Alen Kišić[2] [0000-0002-2196-1092], and Maja Rožman[3][0000-0002-8546-4351]

[1] University of Zagreb, Faculty of Organization and Informatics, Pavlinska 2, 42000, Varaždin, Croatia
[2] VERN University, Palmotićeva ul. 82/1, 10000, Zagreb, Croatia
[3] University of Maribor, Faculty of Economics and Business, Razlagova ulica 14, 2000 Maribor, Slovenia
dijana.oreski@foi.hr, alkisic1@vernnet.hr, maja.rozman1@um.si

Abstract. As generative artificial intelligence (GenAI) tools such as ChatGPT, Gemini, and Copilot increasingly enter educational settings, the ability of educators and learners to interact with these systems meaningfully becomes a critical competence. This paper positions prompt engineering - the practice of formulating effective inputs to guide AI outputs - as a foundational skill within the broader concept of AI literacy for educators and learners. This paper suggest a conceptual framework that connects prompt engineering with existing models of AI literacy and examines its relevance for pedagogical design, content generation, feedback, and student interaction. We synthesize recent literature on prompt typologies, outline key dimensions of competence (cognitive, technical, ethical), and propose a structure for integrating prompt engineering into learners' and educators' professional development. By treating prompt engineering as a form of literacy rather than technical know-how, we argue for its central place in future educational practices and policies to enable a smart society in which learners and educators effectively use generative AI.

Keywords: Prompt Engineering; AI literacy; Generative Artificial Intelligence; ChatGPT.

1 Introduction

The integration of GenAI tools such as ChatGPT, Copilot, and Gemini into educational practice is no longer a matter of future potential - it is an active transformation already underway. These tools are increasingly used to generate content, assist in the automation of assessment feedback, support student inquiry, and even assist in curriculum design. As their presence grows, so does the demand for learners and educators to understand and use them effectively and responsibly. This evolution in educational technology calls for a shift in focus from general digital literacy toward a more specific set of competencies known as AI literacy. AI literacy is commonly defined as the capacity to critically understand, use, and evaluate AI systems in varied contexts [1]. Recent frameworks extend this to generative AI, highlighting the importance of discerning tool selection, prompting strategies, and ethical evaluation of

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.160O
Part of ISBN: 978-86-89755-37-4

AI-generated content [2]. Within this broader framework, prompt engineering - the art and practice of formulating effective and purposeful inputs for large language models (LLMs) - emerges as a foundational skill. While often discussed in technical or developer contexts, prompt engineering is increasingly relevant in pedagogical settings, where the quality of an educator's or learner's prompts can significantly influence the relevance, clarity, and ethical integrity of AI-generated content. Recent research emphasizes this relation. [3] demonstrate that a high level of AI literacy among educators directly impacts the success of prompt engineering strategies in learning environments. Their work supports the argument that prompt engineering is not merely a matter of syntax or tool mastery, but rather a pedagogical and epistemic skill that requires educators to understand the affordances, risks, and limitations of AI systems. Despite its growing importance, prompt engineering is not yet systematically addressed in existing learners' or educators' training programs or AI literacy frameworks. Most users encounter these tools without formal guidance, learning through trial and error or relying on templates of uncertain pedagogical value. This lack of structure presents both a challenge and an opportunity: to define prompt engineering as a core element of educators' AI literacy.

This paper aims to provide a conceptual foundation for understanding prompt engineering as part of AI literacy in education. We synthesize existing literature, examine the cognitive, technical, and ethical dimensions of prompt use, and propose a framework for integrating prompt engineering into learners' and educators` development. The paper seeks to inform future curriculum design and policy initiatives, and to open space for critical discussion around the educator's evolving role in AI-mediated learning environments.

The structure of the paper is as follows: Section 2 explores the concept of AI literacy and positions prompt engineering within it. Section 3 outlines key educational applications and typologies of prompts. Section 4 introduces a conceptual framework for prompt engineering competence. Section 5 concludes the paper with recommendations for practice and directions for future research.

2 Theoretical Background

As AI technologies become increasingly embedded in everyday life, the concept of AI literacy has gained prominence as a critical component of digital competence. It encompasses not only technical understanding but also awareness of societal impacts, ethical considerations, and the capacity to make decisions in AI-mediated contexts. Recent frameworks have moved beyond narrow definitions of technical proficiency, emphasizing a holistic view of AI literacy to support individuals - especially educators and students - with the knowledge and skills to critically engage with AI-driven technologies across different settings [4]

AI literacy frameworks demonstrate three main structural approaches (component, competency, and perspective-based), consistently including technical and ethical components but rarely incorporating explicit measurement tools. AI literacy frameworks are developed for varied audiences (K–12, higher education, workforce, and citizens). Furthermore, such frameworks fall into distinct groups. For example, several studies (e.g., [5], [6]) use component or taxonomy-based models, whereas [7] focus on competency-based approaches. Others (e.g. [8], [9]) develop perspective-based frameworks. [9] propose a multidimensional ABCE model with an explicit 32-item assessment tool. That is one out of only few studies (e.g. [10].) integrating clear measurement strategies, while the remainder rely on conceptual or implicit assessments. Most frameworks converge on the inclusion of technical, ethical, and cognitive components. There is a trend toward multidimensionality, with frameworks increasingly incorporating affective, reflective, and collaborative elements. Distinctions arise in: level of granularity (e.g., [7] present detailed sub-competencies vs. broader categories presented in [11], intended audience (universal vs. stage-specific) and degree of empirical validation, where most frameworks are conceptual (as discussed earlier). The lack of standardized assessment imposes cross-contextual comparison and practical implementation.

Several recent studies discuss prompt engineering as a skill within AI literacy (e.g. [3] Some authors (e.g. [12].) distinguish between prompt engineering (the technical skill of formulating prompts) and prompt literacy (the broader ability to refine and critically assess AI outputs), suggesting a need for specific integration within AI literacy frameworks. Foundation for integration are prompt engineering frameworks, which can serve as basic element for integration. Recent literature reported structured frameworks such as CLEAR [13]. CLEAR consists of five principles: Concise, Logical, Explicit, Adaptive and Reflective [13]. Frameworks such as CLEAR highlight the principles needed to formulate purposeful prompts, emphasizing the skills involved in human–AI interaction.

Building on these foundations, the next section explores how prompts function within educational contexts - examining their typologies, roles, and instructional potential in teaching and learning with GenAI.

3 Functions and Typologies of Prompts in Educational Practice

Prompts are widely used to generate and adapt educational content, including lesson plans, problem solutions, explanations, and creative outputs (such as poetry, art, and guided tours). Automated content creation and personalization are common, especially in higher education and science, technology, engineering, and mathematics (STEM) disciplines. Prompts facilitate the design of assessments, automatic grading, and real-time feedback. Examples include generating test items, providing personalized feedback, and supporting self-assessment and reflection ([14]; [15]; [16]). Prompts support student engagement through interactive activities such as Prompt Problems ([17;[18]) role-plays, Socratic questioning, and collaborative problem solving. Prompts are also used to foster critical thinking, reasoning, and dialogue-based learning.

Table 1 summarizes relevant literature insights regarding prompt engineering templates and effectiveness.

Table 1. Prompt engineering templetes and effectivness

Task Category	Prompt Template	Implementation Guidelines	Reported Effectiveness
Content generation	”Generate a lesson plan on [topic] for [grade level].”	Use structured frameworks (PARTS, CLEAR); iterative refinement	Reported to enhance efficiency and personalize content ([14]; [19])
Assessment design	”Create 5 multiple-choice questions on [concept].”	Specify learning objectives; use output formatting	Reported to support assessment redesign and automate grading ([15]; [16])
Feedback provision	”Provide feedback on this student essay.”	Use explicit criteria; enable feedback loops	Reported to improve relevance and quality of feedback [20];[21]
Computational thinking	”Write a prompt that generates code to solve [problem].”	Scaffold with visual representations; evaluate via test cases	Reported to engage students and develop programming skills [17]; [18]
Critical thinking/Socratic	”Act as a Socratic tutor and guide the student to the answer.”	Assign roles; encourage stepwise reasoning	Reported to foster higher-order thinking and engagement ([20]; [16])
Collaborative learning	”Simulate a peer discussion on [topic].”	Assign personas; structure dialogue	Reported to facilitate collaboration and peer learning [22])

By reviewing literature, we found six distinct task categories addressed by prompt engineering in education: content generation, assessment design, feedback provision, computational thinking, critical thinking/Socratic, and collaborative learning. Each task category had unique implementation guidelines. Most reported effectiveness outcomes were unique to each task, except for ”engagement,” which was reported for both computational thinking and critical thinking/Socratic tasks (two studies). Other outcomes (such as efficiency, personalization, assessment redesign, automating grading, relevance and quality of feedback, programming skills, higher-order thinking, collaboration, and peer learning) were each reported in one study.

4 A Conceptual Framework for Prompt Engineering Competence

GenAI become embedded in educational contexts and the ability of educators to construct, refine, and critically assess prompts takes on new significance. Prompt engineering is not simply a technical action - it is a multifaceted competence that encompasses cognitive understanding, practical skill, and ethical awareness. To support educators in developing this ability as part of their broader AI literacy, we propose a three-dimensional conceptual framework for prompt engineering competence, consisting of: (1) cognitive, (2) technical, and (3) ethical-critical dimension.

4.1 Cognitive Dimension: Understanding How Prompts Shape AI Behavior

At the core of effective prompt engineering lies an understanding of how generative AI models operate. This does not require deep expertise in machine learning, but it does involve cognitive insight into the relationship between input and output, including:

- Model behavior and context sensitivity: Educators need to be aware that large language models are sensitive to phrasing, structure, and specificity. For instance, the difference between “Summarize this text” and “Summarize this text in three bullet points focusing on key arguments” is significant in shaping the response.
- Prompt influence on tone, style, and scope: Educators should understand that the design of a prompt directly affects the style and appropriateness of the generated content. Prompts like “Explain this to a 10-year-old” versus “Provide a graduate-level explanation” yield vastly different outputs.
- Mental models of the AI: Teachers often anthropomorphize or misinterpret AI responses. A cognitively competent educator develops an accurate mental model of how AI “thinks,” recognizing it as a probability-based language model, not an intelligent agent.

This cognitive dimension aligns with the AI literacy component of "knowing what AI can and cannot do," including limitations, probabilistic reasoning, and non-deterministic outputs.

4.2 Technical Dimension: Formulating, Iterating, and Reusing Prompts

The second dimension involves the practical ability to write, test, and improve prompts. This is the most visible aspect of prompt engineering and involves:

- Prompt formulation strategies: Including role-based prompts ("You are a science teacher..."), task-based prompts ("Generate 5 quiz questions about…"), constraints ("Limit your response to 150 words"), and chaining (breaking tasks into subtasks).
- Prompt iteration: Educators must be able to assess the quality of AI outputs and refine their prompts accordingly. This includes experimenting with order, specificity, examples, and scope. Iteration is a skill often acquired through practice and reflection.
- Reusable templates and modular design: Prompt engineering competence includes the creation and adaptation of reusable templates, e.g., prompts for lesson planning, rubrics for feedback generation, or reflection scaffolds for students.
- Multilingual and subject-specific adaptation: Since many educators work in non-English environments, the ability to adapt prompts to different languages and subject-specific terminology is essential.
- Tools and interfaces: Technical competence also includes fluency in the tools used (e.g., ChatGPT's custom instructions, prompt libraries, prompt editors) and awareness of their limitations and updates.

This dimension mirrors the "skills" layer of AI literacy - particularly the ability to use AI tools productively and effectively in varied educational contexts.

4.3 Ethical-Critical Dimension: Responsible and Reflective Prompt Use

The third, often underemphasized dimension of prompt engineering competence concerns ethical awareness and critical reflection. Educators do not only shape the AI's output - they also have responsibility for its pedagogical impact.

- Bias and fairness: Prompts can generate biased content, reinforce stereotypes, or marginalize certain perspectives. Competent educators should be able to detect and address such issues and adjust prompts accordingly.
- Academic integrity: Prompt engineering raises questions of originality and authorship. For instance, if an educator generates feedback or assignments via AI, how should this be disclosed? Where is the line between support and automation?
- Transparency with learners: Prompt use should be transparent to students. Ethical competence involves the ability to explain AI involvement in learning processes and model responsible use.
- Over-reliance and de-skilling: Educators must be able to critically assess when the use of AI is enhancing pedagogy and when it may lead to reduced pedagogical engagement.

- Equity and access: Prompt engineering competence includes awareness of accessibility and inclusivity - designing prompts that consider diverse learners, cultural contexts, and digital divides.

This dimension draws from critical digital pedagogy and responsible AI frameworks. It emphasizes reflection-in-action and reflection-on-action - helping educators become thoughtful users of powerful tools.

4.4 Integrating the Three Dimensions into Learner and Educator Development

The three dimensions outlined above are interdependent. For example, a technically well-defined prompt may still produce ethically problematic content if cognitive understanding of model behavior is lacking. Similarly, ethical reflection is difficult without practical experience or conceptual insight. For integration into educator training, we propose a matrix of competencies aligned with these dimensions. Table 2 would cross-reference pedagogical tasks (e.g., content creation, feedback generation, critical thinking support) with competency areas (cognitive, technical, ethical). This provides a foundation for curriculum design, self-assessment tools, and professional development programs.

We argue that prompt engineering should not be treated as an isolated "trick" or quick fix. Rather, it is a layered competence that must be systematically developed through a blend of conceptual understanding, practical engagement, and ethical deliberation. Its integration into educator education is essential to ensure that educators remain empowered, reflective agents in AI-mediated learning environments.

Table 2. Prompt engineering competence matrix

Educational task	Cognitive competence	Technical competence	Ethical-Critical Competence
Lesson/Content generation	Understand how prompt structure affects output depth, tone, and scope	Formulate structured prompts for curriculum-aligned materials (e.g., quizzes, summaries, slides)	Ensure content accuracy, factual consistency, and absence of bias or misrepresentation
Feedback to Students	Interpret how prompts influence formative vs. summative tone of feedback	Design adaptive prompts for rubric-based or formative feedback	Maintain transparency in AI-generated feedback; avoid depersonalization
Student Support	Anticipate different student needs and learning levels in prompt design	Create prompts for explainers, examples, hints, and adaptive questioning	Prevent reinforcing stereotypes or misalignment with learners' backgrounds

Critical Thinking	Recognize how prompting can scaffold analysis, evaluation, or reflection	Use prompts to stimulate Socratic questioning, debates, or multiperspective reasoning	Avoid oversimplification; ensure inclusion of diverse perspectives
Assessment Design	Predict how prompt framing affects the cognitive demands of tasks	Generate question sets, rubrics, or case studies using varied prompt structures	Ensure fairness, avoid answer leakage, and protect academic integrity
Professional Communication	Understand model behavior in context of formal/informal register, tone, and audience	Use role-based or persona prompts (e.g., "Write an email as a department head explaining...")	Communicate transparently about AI use; avoid manipulation or AI overreach
Tool Mediation (AI in classroom)	Understand how AI mediates knowledge creation and teacher-student interaction	Teach students how to co-create prompts; integrate guided prompting activities into lessons	Model responsible use; discuss AI's limitations and potential social impacts

This table can serve as the foundation for a training curriculum, self-reflection checklist, or teaching portfolio rubric. Each row can be expanded into learning outcomes, micro-credentials, or rubric-based assessments for educator competence in prompt engineering. Also, different versions adapted for primary/secondary vs. higher education can also be developed.

5 Conclusion

GenAI continues to change educational environments and there is the need for educators and learners to develop new competencies. Among these, prompt engineering emerges as a key skill - yet one that is often underestimated, informal, and insufficiently integrated into professional development frameworks. This paper has argued that prompt engineering is not merely a technical tool, but a core component of AI literacy for educators and learners, requiring cognitive insight, practical skill, and ethical reflection. We proposed a three-dimensional conceptual framework of prompt engineering competence - cognitive, technical, and ethical-critical - and illustrated its relevance through common educational tasks. This framework is intended to serve as a foundation for training programs, curriculum design, and self-assessment practices. In doing so, the paper tries to shift the discourse around prompt engineering from anecdotal techniques to a structured pedagogical literacy, grounded in educational theory and aligned with emerging AI literacy standards.

The findings and framework presented in this paper have several implications:

(i) Teacher education and continuous professional development programs should incorporate structured modules on AI literacy, with explicit focus on prompt engineering.
(ii) Educational institutions and policymakers should recognize prompt engineering as a transferable digital-educational skill and support the creation of resources, tools, and guidelines for its responsible application.
(iii) Curriculum developers can use the proposed matrix (Table 3) as a basis to design micro-credentials, workshops, and integrated learning tasks involving generative AI.

This paper contributes to the growing field of AI in education by conceptualizing prompt engineering as a literacy-based competence rather than a purely technical task, aligning prompt engineering with broader theoretical frameworks of AI literacy, proposing a structured model that can inform both pedagogical theory and teacher training design, and providing a typology of educational use cases mapped to specific competency dimensions.

By doing so, the paper addresses a gap in the current literature, which has so far emphasized tool use but has often overlooked the development of underlying educator competencies. While the framework presented is grounded in theory and supported by a synthesis of current literature, several limitations must be acknowledged: the paper does not include empirical validation of the framework through classroom studies, educator surveys, or intervention trials; the proposed competency matrix, while illustrative, may require adaptation for specific contexts, such as early childhood education, vocational training, or multilingual environments; AI tools and prompt strategies are rapidly evolving, which may limit the framework's longevity unless updated regularly. To build upon the ideas presented here, future research should:

(i) Conduct empirical studies testing how different types of educators engage with prompt engineering across disciplines and educational levels;
(ii) Explore learner perspectives on prompts generated or used by educators, particularly regarding transparency, trust, and engagement;
(iii) Develop and validate rubrics and training modules based on the proposed framework;
(iv) Investigate cultural and linguistic dimensions of prompt engineering competence in non-English speaking contexts;
(v) Examine the long-term pedagogical impact of integrating prompt engineering into teaching routines on student learning outcomes and educator agency.

In conclusion, the integration of generative AI into education presents both opportunity and responsibility. Prompt engineering stands at the intersection of pedagogy and technology, requiring not only technical fluency but reflective, ethical, and learner-centered thinking. By framing it as a literacy for the age of AI, we empower educators not only to use these tools - but to shape their use in ways that uphold the values and goals of education itself.

Acknowledgement

This paper is supported by Croatian Science Foundation under the project SIMON: Intelligent system for automatic selection and machine learning algorithms in social sciences, UIP-2020-02-6312.

Declaration on Generative AI

During the preparation of this work, the authors used Claude and Perplexity in order to: edit references, grammar, and spelling check, improve readability, and streamline language. After using tools/services, reviewed and edited the content as needed, and took full responsibility for the publication's content.

References

1. Wang, Y., Wu, X., Zhang, J., Liu, H.: Exploring AI literacy for university students: A framework for critical engagement with artificial intelligence. Information and Learning Sciences, vol. 125, no. 1/2, pp. 45–62 (2023).
2. Zhang, L., Magerko, B.: Developing generative AI literacy: Frameworks for understanding, using, and evaluating large language models. arXiv preprint, (2025).
3. Knoth, N., Tolzin, A., Janson, A., Leimeister, J. M.: AI literacy and its implications for prompt engineering strategies. Computers and Education: Artificial Intelligence, vol. 6, article 100225 (2024).
4. Rožman, M., Oreški, D., Elias, A., Pynadath, M. F., Tominc, P.: AI literacy among university students: A comparative study of three countries – Slovenia, Croatia, and India. IEEE Access, (2025).
5. Chiu, T., Ahmad, Z., Ismailov, Murod, Sanusi, I.: What are artificial intelligence literacy and competency? A comprehensive framework to support them. Computers and Education Open, (2024).
6. MacCallum, K., Parsons, D., Mohaghegh, M.: The Scaffolded AI Literacy (SAIL) Framework for Education. He Rourou, (2024).
7. Chee, Hyunkyung, Ahn, Solmoe, Lee, Jihyun: A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway. British Journal of Educational Technology, (2024).
8. Stolpe, Karin, Hallström, Jonas: Artificial intelligence literacy for technology education. Computers and Education Open, (2024).
9. Gu, Xingjian, Ericson, Barbara J.: AI literacy in K-12 and higher education in the wake of generative AI: An integrative review. arXiv preprint, (2025).
10. Ng, D., Wu, Wenjie, Leung, J., Chiu, T., Chu, S. K.: Design and validation of the AI Literacy Questionnaire: The Affective, Behavioural, Cognitive and Ethical Approach. British Journal of Educational Technology, (2023).
11. Pinski, M., Benlian, A.: AI literacy for users – A comprehensive review and future research directions of learning methods, components, and effects. Computers in Human Behavior, (2024).

12. Hwang, Y., Lee, J. H., Shin, D.: What is prompt literacy? An exploratory study of language learners' development of new literacy skill using generative AI. arXiv preprint arXiv:2311.05373 (2023).
13. Lo, L. S.: The CLEAR path: A framework for enhancing information literacy through prompt engineering. Journal of Academic Librarianship, vol. 49, no. 4, article 102720 (2023).
14. Lee, Daniel, Palmer, Edward: Prompt engineering in higher education: A systematic review to help inform curricula. International Journal of Educational Technology in Higher Education, (2025).
15. Chen, Eason; Wang, Danyang; Xu, Luyi; Cao, Chen; Fang, Xiao; Lin, Jionghao: A systematic review on prompt engineering in large language models for K-12 STEM education. arXiv preprint, (2024).
16. Heston, Thomas F.: Prompt engineering for students of medicine and their teachers. arXiv preprint, (2023).
17. Denny, P., Leinonen, J., Prather, J., Luxton-Reilly, A., Amarouche, T., Becker, B. A., Reeves, B. N.: Promptly: Using prompt problems to teach learners how to effectively utilize AI code generators. arXiv preprint arXiv:2307.16364 (2023).
18. Denny, P., Leinonen, J., Prather, J., Luxton-Reilly, A., Amarouche, T., Becker, B. A., Reeves, B. N.: Prompt Problems: A new programming exercise for the generative AI era. In: Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1, pp. 296–302 (2024, Mar).
19. Skvortsova, S., Symonenko, Tetiana, Hnezdilova, K.: Prompt engineering as a tool for optimizing the professional activities of educators: An experimental study of the effectiveness of generative language models. Cherkasy University Bulletin: Pedagogical Sciences, (2025).
20. Leung, Chi-hong: Promoting optimal learning with ChatGPT: A comprehensive exploration of prompt engineering in education. Asian Journal of Contemporary Education, (2024).
21. Park, Jiyeon, Choo, Sam: Generative AI prompt engineering for educators: Practical strategies. Journal of Special Education Technology, (2024).
22. Polverini, G., Gregorcic, B.: How understanding large language models can inform the use of ChatGPT in physics education. European Journal of Physics, vol. 45, no. 2, article 025701 (2024).

Mentora ChatBot as an Intelligent Recommender System in Education

Olga Dukić[1[0009−0004−1694−9592]] and Marija Blagojević[1[0000−0003−4186−0448]]

University of Kragujevac, Faculty of Technical Sciences Čačak, Department of Information Technologies, Čačak, Serbia
email: dukicolga678@gmail.com, marija.blagojevic@ftn.kg.ac.rs

Abstract. The development and evaluation of the Mentora ChatBot, an educational conversational agent designed to provide personalized learner support, are presented. Mentora combines large language models (LLMs) with sequential modeling of user behavior, relying on a BiLSTM architecture alongside heuristic rules to generate recommendations aligned with learners' needs.
The system was developed and initially tested during preparatory classes for the final exam in mathematics and the Serbian language, conducted from February–June 2025 at a private educational institution. Twenty-two eighth-grade students participated, and more than 600 interaction logs were analyzed. Substantial gains in recommendation precision and acceptance were observed with the integration of the BiLSTM module, relative to a purely heuristic baseline.
An empirically grounded approach is introduced that couples LLM-driven dialogue with personalization of learning trajectories. Directions for further development are outlined, including the adaptation of recommendations to diverse learner profiles and the incorporation of explainable recommendations (explainable AI).

Keywords: Recommender systems · educational chatbot · BiLSTM · log analysis

1 Introduction

Contemporary education increasingly strives for a personalized approach to learning, wherein recommender systems and artificial intelligence (AI) play a growing role in monitoring progress and tailoring content to individual learner needs [1,5]. A particular place in this transformation is occupied by large language models (LLMs) and conversational agents, which enable dialog-based interaction, automated feedback, and intelligent guidance of the learning flow.

The development and evaluation of the **Mentora ChatBot**, an educational conversational assistant that uses a combination of LLMs, log analysis, and sequential modeling to generate learner-tailored recommendations, are presented. The system was developed as part of the broader *Mentora* educational platform, designed for deployment in real school settings and encompassing interactive lessons, quizzes, AI-based assistance, and learner progress tracking.

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.171D
Part of ISBN: 978-86-89755-37-4

Mentora ChatBot combines explicit data (e.g., learners' answers to questions) with implicit signals from behavioral logs during learning sessions to produce personalized recommendations. Within the recommender module, a BiLSTM model is implemented to enable sequential analysis and tracking of learning trajectories over time. In addition, heuristic approaches are employed to obtain a hybrid effect by combining rule-based logic with data-driven learning. The current version of the BiLSTM model operates on synthetically generated data, whereas real user logs are in the process of being collected and expanded.

This study addresses the following research questions:

- How can large language models support the generation of educational recommendations within a conversational interface?
- To what extent does combining explicit data and behavioral logs contribute to higher-quality personalization?
- How can sequential models, such as a BiLSTM architecture, improve the dynamic adaptation of recommendations to the learner over time?

The remainder of this paper is organized as follows. Section 2 reviews relevant literature on educational recommender systems and conversational agents. Section 3 describes the architecture and functionalities of the Mentora system. Sections 4 and 5 detail the development of the recommender module and the analysis of log data, while Section 6 presents the experimental results and evaluation. Concluding remarks and directions for future development are provided in Section 7.

The Mentora system was developed during the 2024/2025 academic year within a private educational institution, as part of broader research into the application of AI in personalized learning. The objective is the enhancement of self-regulated learning by leveraging dialog-based interaction and recommendations grounded in prior learning experiences.

2 Theoretical Background and Literature Review

This section summarizes the theoretical foundations and current research directions that have shaped the development of the *Mentora ChatBot* system. The focus is on (i) personalization of learning through recommender systems, (ii) the role of conversational agents, (iii) analysis of usage logs and sequential modeling (KT/sequence-aware RS), and (iv) explainable recommendations and knowledge graphs. In this way, a conceptual framework is established for a hybrid approach (heuristics + BiLSTM within dialogue).

2.1 Recommender Systems in Education

Recommender systems (RS) in education are most commonly based on content-based filtering (CBF), collaborative filtering (CF), and hybrid approaches that combine multiple signals [9]. In CBF, profiles are formed from attributes of lessons/tasks, whereas CF relies on similarities derived from behavioral patterns across larger groups of learners. Hybrid approaches are particularly suitable in school settings due to data heterogeneity (different topics, tasks, learning styles) and the cold-start problem [5,1]. For working with web and usage logs, as well as evaluation metrics (*Precision*@k, *Recall*@k, NDCG@k), standard data-mining literature was followed [2,4,11]. In line with prior work advocating multi-layer architectures for educational recommender systems, the system was modularized into logging, evaluation, and recommendation components [10].

2.2 Conversational Recommender Systems (CRS)

Conversational recommenders (CRS) fuse recommendation generation with dialog: through multiple turns, preferences and constraints are elicited, explanations are provided, and suggestions are iteratively adapted. Surveys of CRS [3] and recent work in the area point to key components: a user model, a dialog management strategy, integration of background knowledge, and evaluation (offline and user studies). In education, CRS naturally align with *learning pathways*—dialog is used for diagnosis and the selection of next steps [12]. In the present approach, language understanding and generation are handled by an LLM, whereas recommendations are produced in a hybrid manner: heuristics for stability and a BiLSTM for sequential context.

2.3 Log Analysis and Sequential Modeling

Usage logs (interactions with lessons, tasks, and the bot) contain implicit patterns of progression, impasses, and knowledge transfer. Sequential modeling of these sequences (sessions, events, inter-event times) enables anticipation of the learner's next step. The seminal *Deep Knowledge Tracing* line of work has shown that LSTM models over task/answer sequences effectively predict performance [7]. Complementary surveys on sequence-aware recommender systems further support the use of sequential context for improved predictions [8]. In the same spirit, a BiLSTM is employed to couple the learner's "history" and "current context," with operation supported under limited real logs augmented by synthetic sequences (pilot phases and cold start). Empirical studies in learning analytics also document predictive value in real classroom logs [6].

2.4 Knowledge Graphs and Explainable Recommendations

Explainable recommendations are essential for trust and uptake, particularly in education, where both learners and teachers seek the "why" alongside the "what." Knowledge graphs (concepts, prerequisites, learning objectives) and a retrieval-augmented generation (RAG) layer over lessons provide a natural basis for explanations: "recommended because it is a prerequisite for X" or "because recent tasks in Y were challenging." Surveys of explainable RS [13] and work on graph-based learning paths indicate the effectiveness of these approaches in EdTech. In the current system, explanations are grounded in lessons and objectives (RAG), while explicit ontologies are planned for subsequent iterations.

3 System Description and Methodology

3.1 Architecture of the Mentora ChatBot System

Mentora ChatBot is a modular conversational system designed to enable a personalized educational experience in primary and secondary school contexts. The system was developed as part of the broader Mentora platform and is based on the integration of large language models (LLMs), an interaction-tracking layer, a heuristic recommendation layer, and a sequential model built on a BiLSTM architecture.

The architecture of the Mentora ChatBot system, which includes the LLM module, the lesson embedding space, the heuristic layer, answer evaluation, and the BiLSTM module for sequential modeling of learner behavior, is shown in Figure 1.

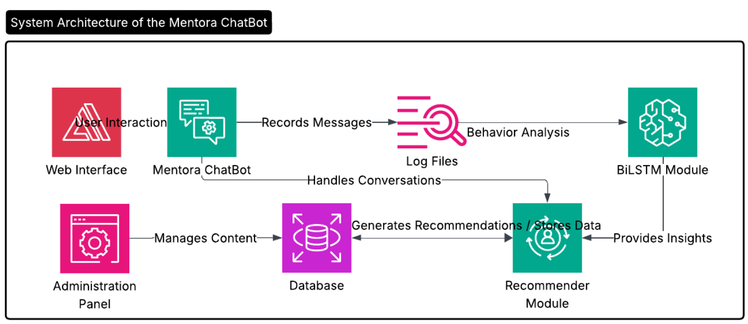


Fig. 1. Architecture of the Mentora ChatBot system

The system components include:

- **LLM module** — the OpenAI GPT-4o API is used for natural-language understanding and generation;
- **Lesson repository and embedding space** — lessons are represented by semantic vectors and tags and are indexed for relevant retrieval;
- **Answer evaluation system** — learner solutions are automatically analyzed (textual, visual, and audio input);
- **Recommendation system** — heuristics are applied based on learner behavior, with an integrated BiLSTM module proposing next learning steps.

A similar multi-layer architecture for e-learning hybrid recommenders has been described, and the present study focuses on the K–12 deployment setting [10].

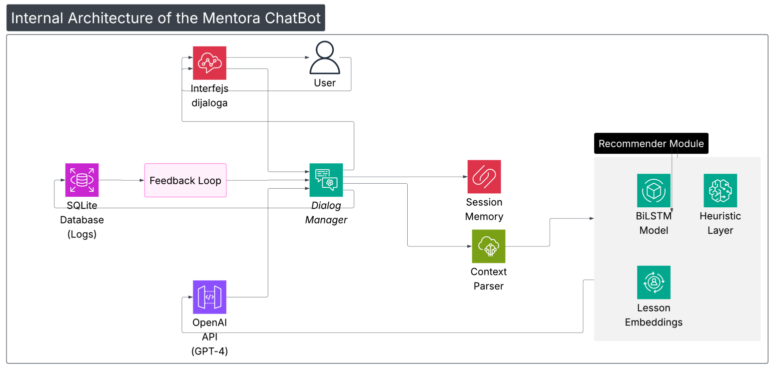


Fig. 2. Internal architecture of the Mentora ChatBot

3.2 Internal Architecture of the Mentora ChatBot

The internal architecture of the Mentora ChatBot is shown in Figure 2. The system consists of multiple interconnected modules organized around user interaction and personalized learning support. The key components are:

- **Input module** — textual, image, or voice input is received and forwarded to the system;
- **Context Parser** — each message is analyzed in the context of prior interactions, with intent and learning phase identified;
- **Dialog Manager** — the course of the conversation is managed and subsequent steps are determined;
- **Session Memory** — prior answers, recommendations, and bottlenecks are retained to ensure continuity;
- **Recommendation module** — individualized suggestions are generated using heuristics and a BiLSTM model;
- **Lesson Embeddings** — semantic retrieval of content is enabled;
- **LLM API layer** — natural-language responses are generated using the GPT-4o model;
- **Feedback loop** — behavioral data are leveraged to improve personalization.

This architecture enables Mentora ChatBot not only to function as a conversational agent but also as an adaptive tutor that learns about the user and monitors progress. The BiLSTM module contributes in particular to understanding behavioral patterns (e.g., recurring errors, returns to prior lessons), thereby producing recommendations aligned with the learner's pace and style.

3.3 Data Collection and Processing

Data are collected directly from learner interactions with the system and include:

- textual answers and questions;
- selected areas of interest;
- task-solving performance;
- session duration and activity order;
- visual inputs (images of tasks, drawings).

All data are stored in a relational database and, in parallel, converted into sequential series for training the BiLSTM module. In the initial phase, the model was tested on synthetic data that emulate typical learner behavior. Starting in September 2025, processing of real data is planned to improve performance and recommendation accuracy.

3.4 System Functionality

Mentora ChatBot integrates multiple functionalities for learning personalization and learner support:

- **Interactive tutor** — questions are answered, learners are guided through lessons, and assistance with problem solving is provided;
- **Recommendations** — subsequent lessons, tasks, and explanations are suggested;
- **Answer evaluation** — textual and visual responses are analyzed and feedback is provided;
- **Progress tracking** — activity is logged and learner progress is measured;
- **Adaptive feedback** — heuristics and the BiLSTM model are used to deliver intelligent recommendations.

The system was implemented using the following technologies:

- **Backend:** Django, PostgreSQL;
- **Frontend:** HTML, CSS, JavaScript, Bootstrap;
- **AI:** GPT-4o API, text-embedding-3-small, prompt engineering;
- **Recommender:** heuristics + BiLSTM in a TensorFlow/Keras environment.

Configuration and API keys are stored in an `.env` file, with scalability supported through a modular architecture.

3.5 Role of the Recommender Module in Dialogue

A modular structure is used to generate recommendations during dialog with the learner. When a need is expressed or a question is posed (e.g., "I do not know how to solve a fractions problem"), the context is recognized and the corresponding part of the recommender module is triggered.

Depending on the quantity and quality of prior interactions with the learner, the following is activated:

- a heuristic approach — in the case of a new user or insufficient data;

– the BiLSTM model — when sufficient information is available about the learner's prior behavior in the system.

The recommendation is then surfaced through the dialog, and multiple options are offered: the suggestion can be accepted, additional explanation requested, or an alternative task provided. When a rich activity history is available, the BiLSTM model is used to generate the next suggestion; otherwise, heuristics are relied upon.

4 Heuristic Layer of the Recommender Module and Personalization

The recommender module within the Mentora ChatBot system was designed to offer learners content that is most relevant at a given moment, in accordance with their prior interactions, achieved results, and interests. At the present stage of development, a heuristic approach is employed that combines explicit and implicit learner data.

This approach constitutes the first layer of the recommendation system and establishes a foundation for later integration with more complex data-driven models.

4.1 Heuristic Strategies

The heuristic layer operates via a set of rules that map observable behavioral patterns to concrete recommendations. Key rules include:

– **Revisiting areas of difficulty** – if repeated incorrect answers are observed within the same topic, additional lessons and tasks from that domain are recommended;
– **Suggesting the next lesson in sequence** – when a lesson is successfully mastered, the next item in the syllabus sequence is suggested;
– **Accounting for prior interest** – if pronounced interest in a topic is evident (e.g., requests for extra examples or questions posed), content related to that concept is prioritized;
– **Adapting to session length and frequency** – for shorter or less frequent sessions, shorter and easier tasks are proposed to help maintain motivation.

4.2 Personalization via Rules

Although not based on machine learning, these heuristics enable a baseline level of personalization and content adaptation. The rules are currently hand-crafted but were informed by insights gained during the pilot testing phase. In subsequent versions, it is planned that rules be generated dynamically from discovered behavioral patterns.

At present, the rules are implemented directly in the system code; however, a management interface is planned to allow teachers or administrators to edit and tailor them to the needs of specific learner cohorts.

4.3 Integration with the Tutor

This heuristic layer is directly coupled with the system's LLM tutor component. Based on the triggered rules, a contextual prompt is constructed and employed within the dialog. For example:

> "Given that the previous lesson on fractions appeared challenging, it is recommended that another example be worked through together. Would you like to proceed?"

In this way, recommendations are not displayed as a separate list but are naturally woven into the flow of conversation, consistent with the principles of conversational recommender systems [3]. This increases the likelihood that the recommendation will be perceived as supportive guidance rather than an algorithmic command, which positively affects engagement.

4.4 Limitations of the Heuristic Approach

The heuristic approach is valuable for rapid testing and basic personalization; however, limitations arise with more complex modeling of learner behavior, especially when multiple factors interact or when learner needs are highly variable.

In addition, heuristics may be insufficiently robust when a mismatch between measured performance and expressed confidence is present, e.g., when tasks are solved correctly but uncertainty or requests for additional explanation are expressed in dialog.

Despite these limitations, heuristics provide a stable and easily adaptable starting layer for recommendations. Their principal advantages lie in simplicity and transparency. In what follows, their extension via sequential modeling of learner behavior is explained.

5 Sequential Modeling of Learner Behavior Using Logs and a BiLSTM Network

5.1 Log Collection and Processing

Log files constitute the foundation for analyzing user behavior in the Mentora system. This approach enables detailed mapping of the user journey and supports advanced personalized learning. Every interaction with the Mentora ChatBot—questions posed, lesson access, correct and incorrect answers, time-on-task, and answer type (textual, multiple-choice, visual input)—is recorded in pseudonymized form in a relational database.

The data are transformed into sequences that represent learning-interaction flows. Each learning session is encoded as a sequence of attribute vectors (e.g., lesson identifier, task type, performance, response time, topic), thereby enabling the application of sequential prediction models.

Prior research has shown that logs from real educational systems encode behavioral patterns that can be mined to improve recommendations and predict learning outcomes [8,6]. A similar strategy is adopted in Mentora to construct the predictive layer. All logs are stored in accordance with privacy and ethical principles: personal identifiers are pseudonymized, and data access is strictly restricted.

5.2 BiLSTM Model Structure

A bidirectional LSTM (BiLSTM) architecture is employed for sequential modeling of learner behavior, well-suited for uncovering patterns in interaction streams.

The model consumes input sequences composed of event vectors with the following components:

- activity type (lesson, quiz, question),
- topic category (e.g., fractions, grammar, logical reasoning),
- activity outcome (correct/incorrect),
- execution time and inter-event pauses,
- prior recommendations and accepted suggestions.

These vectors are normalized and arranged into fixed-length input tensors for training.

The BiLSTM layer processes the sequence in both forward and backward directions, yielding a contextualized representation at each step. In this manner, it is enabled to account for how preceding and subsequent events influence the current timestep, providing a deeper understanding of learning patterns and more precise next-step predictions. Complex trajectories are thus captured, including performance oscillations, topic transitions, and responses to earlier recommendations.

The implementation is based on the PyTorch framework. The model is planned to be trained on sessions collected during the platform's test phase (August–September 2025), with accuracy and F1-score evaluated on a validation set.

5.3 Predicting Learners' Next Steps

On the basis of the BiLSTM output representations, the Mentora system can generate personalized recommendations that anticipate the learner's most likely next step, including:

- recommending the next lesson consistent with prior performance and topic,
- proposing a task type that best suits the learner (visual, textual, interactive),
- assessing the need for a concept explanation or additional guidance,
- flagging to the instructor a potential misunderstanding.

Such recommendations are injected into the Mentora ChatBot dialog, enabling proactive, intelligent system behavior aligned with the goals of conversational recommender systems in education [3,8]. In this way, the system not only reacts to learner actions but also anticipates needs, creating the experience of an intelligent, supportive tutor. Future development includes automated evaluation of recommendations using engagement and performance metrics, as well as experimental comparisons with static heuristic systems. By integrating the BiLSTM model into the Mentora architecture, the platform transitions from a static, rule-based approach to a dynamic system that tracks—and shapes—the learning trajectory in real time.

6 Results and Discussion

6.1 Evaluation Protocol and Statistical Considerations

The pilot was conducted during February–June 2025, including a seven-day intensive testing window, with $n = 22$ eighth-grade students. Learners were assigned to an experimental (BiLSTM) and a control (heuristics) group ($n = 11/n = 11$). Interactions were logged uniformly across groups. Given the pilot scale, we emphasize effect sizes (absolute deltas) and 95% bootstrap confidence intervals (10,000 resamples) rather than null-hypothesis significance testing. Reported gains (Precision@3 +0.17, F1 +0.19; HR 76%) are interpreted as indicative and consistent across sessions.

Compliance with Ethical Standards. All interactions were pseudonymized, and participation was conducted with parental consent, in line with institutional ethical guidelines.

6.2 Experimental Results

The evaluation outcomes are shown in Figure 3. The BiLSTM-based recommendation module was found to outperform the heuristic approach across all metrics. The largest differences were observed in recommendation accuracy and session duration, indicating a higher degree of personalization and greater learner engagement.

The quantitative results of the pilot study are as follows:

- **Precision@3:** 0.71 (BiLSTM) vs. 0.54 (heuristics);
- **F1 score:** 0.68 (BiLSTM) vs. 0.49;
- **Hit Rate:** 76% (BiLSTM);
- **Average session length:** 12.4 minutes (BiLSTM) vs. 9.8 minutes;
- **User satisfaction:** 4.3 out of 5 (Likert-scale survey).

These findings corroborate the initial hypothesis that sequential modeling enables a deeper understanding of learning trajectories and more accurate prediction of next steps. Beyond improved precision, a higher level of learner engagement was observed, reflected in longer sessions and more frequent interaction with the tutor.

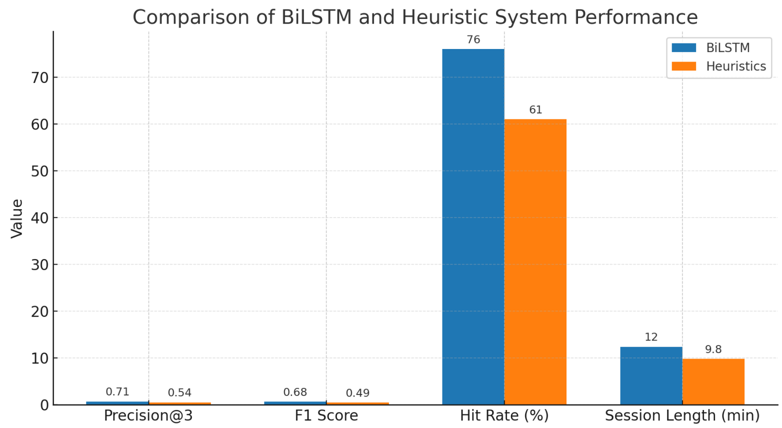


Fig. 3. Performance comparison between the BiLSTM model and the heuristic system

6.3 Discussion and Contributions

The results situate this work within current research on conversational recommender systems (CRS) and sequential modeling in education. In contrast to approaches focused solely on dialog without deep sequential modeling [3], a BiLSTM was integrated to capture temporal dependencies in learning logs. Relative to knowledge tracing that applies LSTM outside the dialog loop [7,6], the contribution lies in embedding the sequential model directly into the dialog flow. Furthermore, by introducing a RAG-based explanation layer, existing efforts on explainable recommendation are extended [13], providing pedagogically meaningful justifications (e.g., "this is a prerequisite for ...", "difficulties were previously observed in ...").

An empirical contribution was established through a pilot evaluation in a real educational setting (22 students, >600 interactions), where measurable gains of the BiLSTM over heuristics were demonstrated (Precision@3 +0.17; F1 +0.19; HR 76%). These results indicate that combining an LLM-driven tutor with a BiLSTM sequence model yields more relevant, adaptive, and explainable recommendations than traditional heuristic approaches.

Limitations. The pilot study was constrained by a relatively small sample and partially synthetic logs. Future work is planned to involve larger cohorts, ablation studies, and enrichment of learner profiles with metadata on motivation, prior knowledge, and learning goals.

7 Conclusion and Future Work

In this paper, the *Mentora ChatBot* was presented—a hybrid conversational recommender system (CRS) for educational settings that combines an LLM-driven tutor, heuristic rules, and a BiLSTM model over learning logs for on-policy decision-making. The system architecture and data flow were described, and a pilot evaluation was conducted under real final-exam preparation conditions (eighth grade), with a focus on recommendation quality and learner acceptance.

Contributions. (i) Integration of *LLM + BiLSTM + RAG* into a single, on-policy decision flow (the LLM manages the dialog, the BiLSTM models sequential context, and the RAG layer provides explanations understandable to learners and teachers); (ii) empirical confirmation that sequential context improves recommendation relevance and uptake over a heuristic baseline (22 students, >600 interactions): *Precision@3* 0.71 vs. 0.54, *F1* 0.68 vs. 0.49, *Hit Rate* 76%; (iii) operationalization of a CRS in a school setting (mathematics and Serbian language) with minimal changes to the existing instructional flow.

Our study has important limitations: a small, convenience sample, irregular study sessions, and partially synthetic logs in the early phase. These factors may bias the results, inflate apparent performance, and materially constrain both internal and external validity; hence, the findings should be viewed as preliminary. In future work we will (i) expand the dataset with real, end-to-end logs from a larger and more diverse student cohort and additional subject domains, (ii) run systematic ablations to isolate component contributions (LLM-only, BiLSTM-only, RAG-only, and combinations), (iii) examine Transformer-based architectures for sequential modeling, and (iv) deepen personalization by learner profile, coupled with richer, pedagogically grounded explanations. To enhance transparency and interpretability, we will include anonymized student–chatbot dialogue snippets in the supplementary material and present additional examples during the conference talk.

In sum, it was demonstrated that the *LLM+BiLSTM+RAG* hybrid can deliver pedagogically meaningful, contextually adaptive, and empirically better-accepted recommendations than static heuristics, laying the groundwork for scalable and transparent learning support in real classroom practice.

Acknowledgments This study was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, and these results are parts of the Grant No. 451-03-136/2025-03/200132, with University of Kragujevac - Faculty of Technical Sciences Čačak.

References

1. Drachsler, H., Kalz, M.: Panorama of recommender systems to support learning. In: Ricci, F., Rokach, L., Shapira, B. (eds.) Recommender Systems Handbook, pp. 421–451. Springer (2015). `https://doi.org/10.1007/978-1-4899-7637-6_12`
2. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Elsevier, 2nd edn. (2006)
3. Jannach, D., Resnick, P., Tuzhilin, A., Zhang, Y.: Conversational recommender systems. Foundations and Trends in Information Retrieval **16**(1), 1–101 (2021). `https://doi.org/10.1561/1500000068`

4. Liu, B.: Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Springer, 2nd edn. (2011)
5. Manouselis, N., Drachsler, H., Verbert, K., Duval, E.: Recommender Systems for Learning. SpringerBriefs in Electrical and Computer Engineering, Springer (2012). https://doi.org/10.1007/978-1-4614-4361-2
6. Okubo, F., Yamashita, T., Shimada, A., Ogata, H.: Learning analytics to improve learning outcomes: A case study of a flipped classroom. Research and Practice in Technology Enhanced Learning **12**(1), 1–23 (2017). https://doi.org/10.1186/s41039-017-0043-6
7. Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., Sohl-Dickstein, J.: Deep knowledge tracing. In: Advances in Neural Information Processing Systems (NeurIPS) (2015)
8. Quadrana, M., Cremonesi, P., Jannach, D.: Sequence-aware recommender systems. ACM Computing Surveys **51**(4), 66:1–66:36 (2018). https://doi.org/10.1145/3190616
9. Ricci, F., Rokach, L., Shapira, B.: Introduction to recommender systems handbook. In: Recommender Systems Handbook, pp. 1–35. Springer (2011). https://doi.org/10.1007/978-0-387-85820-3_1
10. Roux, L., Dagorret, P., Etcheverry, P., Nodenot, T., Marquesuzaa, C., Lopisteguy, P.: A multi-layer architecture for an e-learning hybrid recommender system. In: Proceedings of the 18th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2021). pp. 179–187 (2021)
11. Russell, M.A.: Mining the Social Web. O'Reilly Media, 1st edn. (2013)
12. Sun, Y., Papamitsiou, Z., Giannakos, M.: Design and evaluation of a conversational recommender system for learning pathways. IEEE Transactions on Learning Technologies **16**(1), 95–108 (2023). https://doi.org/10.1109/TLT.2022.3185156
13. Zhang, Y., Chen, X.: Explainable recommendation: A survey and new perspectives. Foundations and Trends in Information Retrieval **14**(1), 1–101 (2020). https://doi.org/10.1561/1500000066

ChatGPT's Capability in Solving Mathematical Modelling Problems

Nemanja Vučićević[1][0000-0002-4903-7280], Aleksandar Milenković[1*][0000-0001-6699-8772], and Marina Svičević[1][0000-0003-2791-3849]

[1] University of Kragujevac, Faculty of Science, Radoja Domanovića 12, 34000 Kragujevac, Serbia
nemanja.vucicevic@pmf.kg.ac.rs, aleksandar.milenkovic@pmf.kg.ac.rs, marina.svicevic@pmf.kg.ac.rs

Abstract. In this paper, we examine the capability of ChatGPT in solving problems based on mathematical modelling. In line with the research objective, we assigned the free version of ChatGPT (GPT-4o) six problems that fourth-year mathematics students, majoring in Applied Mathematics, had solved as part of their exam. We analyzed whether the following steps were properly carried out: constructing a mathematical model, the problem-solving process in the mathematical context, interpreting the solution in relation to the problem, and finally, validation. The results were then compared with the work of the aforementioned students. The findings indicate that, in all parts of the solution and for all tasks, the free version of ChatGPT was either completely or partially successful. Nevertheless, it achieved results that were weaker than those of the students. Based on these findings, we can say that solutions to problems addressed through mathematical modelling – which, in the second step, involve solving first-order differential equations, game theory problems, and linear programming tasks – can be used to some extent as student support. However, it is equally necessary to work on the specialization of LLMs that would solve problems of this level of complexity in accordance with the mathematical modelling process.

Keywords: Mathematical Modelling, ChatGPT, Student Performance Comparison

1 Introduction

In recent times, an increasing number of researchers in the fields of education, engineering, and computer science have been exploring the potential and capabilities of large language models (LLMs) in solving various tasks through AI tools. Contemporary research mainly investigates the extent to which specific tools can successfully address different challenges and solve concrete tasks and problems and often compares the performance of different models in executing specific user instructions. Among these tools, ChatGPT is probably the most widely used.

Mathematical modelling is gaining increasing prominence in mathematics education. On one hand, it is important to introduce future mathematicians – particularly those seeking careers in industry – to the principles of mathematical modelling so that

* Corresponding authors' e-mail: aleksandar.milenkovic@pmf.kg.ac.rs

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.184V
Part of ISBN: 978-86-89755-37-4

they can successfully translate real-world problems from various fields of science and industry into mathematical contexts, solve them, and interpret and discuss the results within the original context. Likewise, prospective mathematics teachers should acquire the same competencies and develop the pedagogical skills needed to teach their students the step-by-step process that characterizes mathematical modelling.

This raises the question of to what extent ChatGPT can effectively solve mathematical modelling problems, to serving as an additional resource in the education of future mathematicians and, in particular, as support in solving problems from various domains through the application of mathematical tools. This study aims to contribute to this emerging area of research.

2 Literature Review

One of the widely accepted processes or cycles of mathematical modelling was proposed by Blum and Leiß [1]. The cycle they suggested consists of six phases and six transitions between these phases. According to their methodology, the first step is to understand the real-world problem and construct the so-called situation model (1). This situation model is then simplified – specifically, the problem is structured, and the necessary assumptions are formulated (2). The next step, which is particularly important in this cycle, involves constructing a mathematical model through a process of mathematization (3). In this step, the process shifts into a mathematical context. The mathematical problem thus formed is then solved, yielding the required mathematical results (4). These mathematical results are then considered and examined in the context of the real-world problem to determine how they can be effectively used to solve the initial problem; through interpretation, results corresponding to the initial real situation are obtained (5). Finally, the results obtained should be validated (6). Within this validation process, the modelling steps should be reconsidered by re-examining the assumptions, the mathematical models, their solutions, and the correctness of the calculations. Although it represents an important aspect of learning, the meta-analysis conducted by Schukajlow et al. [2] found that only 3% of research papers published in leading journals on mathematics education over a six-year period focused on the topic of mathematical modelling. Furthermore, most of these papers were empirical studies with a qualitative research design. Sen Zeytun et al. [3] conducted a qualitative study involving six mathematics students in order to identify the sources of students' difficulties in learning mathematical modelling. Based on their results, it can be concluded that the factors contributing to students' difficulties can be divided into contextual and individual categories. Contextual factors refer to the students' lack of experience in solving problems through mathematical modelling, while the individual factors relate to the time constraints for completing the tasks. Individual factors also include gaps in mathematical knowledge and skills, difficulties in establishing connections between the real-world and mathematical contexts (and vice versa), students' tendency to focus solely on the result, as well as their inability to adequately organize their work [3].

ChatGPT is a Natural Language Processing (NLP) model developed by OpenAI that uses a large dataset to generate responses to students' questions, prompts, and feedback

[4]. There are many studies that explore the capabilities of ChatGPT and other OpenAI models in solving mathematical tasks and problems. For example, OpenAI models have been tested in solving complex mathematical competition problems. Interestingly, when solving problems of lower complexity but with a higher degree of non-standard structure, OpenAI tools perform at the level of average competition participants [5]. However, when tackling more demanding mathematical problems intended for high school students, the o1 and o3-mini models would win awards when competing among the best high school students [6].

ChatGPT can solve problems in the field of linear algebra, compute integrals, and solve differential equations when given appropriate instructions [7]. When it comes to differential equations, it was found that ChatGPT correctly solved 72% of the given first-order differential equations, made minor errors in 20% of them, and in the remaining 8% it made serious computational or methodological errors. In solving second-order differential equations, it solved 64% of the problems completely correctly, made minor errors in 24% of them, and made a serious computational or methodological error in every eighth problem [8]. Zhu et al. [9] explored ChatGPT's compatibility in decision-making when solving game theory problems, which are also widely used in mathematical modelling.

Some research on integrating LLMs into solving mathematical modelling tasks has already been conducted. For instance, Huang et al. [10] introduced a LLM model called Mamo, which has integrated solvers aimed at validating solutions and evaluating the correctness of the entire mathematical modelling process.

3 Research question

Considering the importance of mathematical modelling for the education of mathematicians – regardless of whether their future profession will be in industry or in teaching – as well as the potential of ChatGPT for solving tasks of various levels in terms of mathematical content and complexity, we aimed to examine the accuracy and overall problem-solving process of tasks based on mathematical modelling generated by ChatGPT. This gave rise to the following research questions.

1. To what extent does ChatGPT successfully solve mathematical modelling problems related to the application of differential equations, game theory and linear programming?
2. Are there differences in the performance between mathematics students who attended the Mathematical Modelling course and ChatGPT in terms of accuracy and the structure of solutions within the context of the mathematical modelling cycle?
3. Can the solutions produced by ChatGPT serve students in carrying out validation – the final phase of the mathematical modelling cycle?

4 Method

As part of the course Mathematical Modelling, which is conducted in the seventh semester of the Applied Mathematics program at Faculty of Science, University of Kragujevac, students are expected, upon completion of the course, to be able to apply the principles of mathematical modelling and to formulate mathematical models in various fields of natural and social sciences. This course was offered for the first time in the 2023/2024 academic year. Only four students attended it during that academic year. After completing the theoretical and practical instruction, the students took the written part of the exam, in which they were expected to demonstrate practical knowledge of mathematical modelling. On that occasion, they solved tasks in which they were required to carry out the last four phases of mathematical modelling, specifically from mathematization to validation [1]. The test consisted of 6 tasks. The first task involved a simple theoretical consideration of an exponential growth model with a given constant. The second task focused on the application of a first-order differential equation to a bacterial population in a controlled environment, where students were required to construct an appropriate logistic differential equation using the given data. Using the model they created, they had to determine the size of the population at several specific time points from the initial moment, identify when the population would reach half the carrying capacity, and analyze what happens as the time variable tends to infinity. The third task was also a population model, but with constant relative rates of birth, death, and emigration over time. Students were asked to determine a particular solution to the problem for a given initial value, find the condition under which the model exhibits exponential growth, and identify the condition under which the population remains constant. In the fourth task, students were provided with tabular data related to time and population size and were required to construct a model that fits the given data. Based on the model, they were to explain the meaning of all parameters and their influence on the behavior of the model. The fifth problem dealt with mathematical modelling within decision theory – more specifically, game theory and the analysis of strategies in a competitive environment. The task required modelling the competitive relationship between two manufacturers, identifying the strategies of each player, and determining the optimal decision for the first company in response to various possible moves by its competitor. The final, sixth task focused on the application of linear programming, specifically optimizing production under resource constraints, with two objectives: minimizing production costs under the condition that a certain number of products be manufactured, and maximizing the number of products produced given limited resources.

We gave the same problems that the students had solved to ChatGPT to solve as well. Considering that we wanted to examine whether it could help students in terms of validating their answers by comparing them with the solution produced by ChatGPT, we used the free version of ChatGPT (GPT-4o). When evaluating the solutions provided by the students and the ChatGPT, we assessed the work according to the modelling cycle – that is, to what extent the problem solutions followed the given mathematical modelling approach. The study was conducted in July and August 2025.

5 Results

To address the research questions, we analyzed whether ChatGPT successfully solved the six tasks that the students had worked on. Specifically, for each task, we examined whether the following steps (Fig.1) were successfully carried out: translating the problem into a mathematical context, i.e., the process of mathematization; whether the problem-solving procedure was mathematically correct; interpreting the solution in accordance with the original context; whether and how the validation process was carried out.

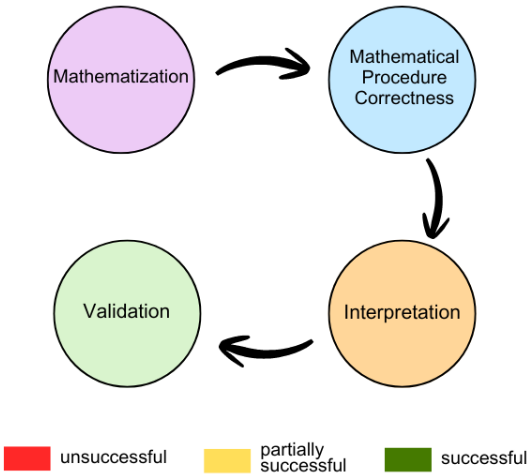


Fig. 1. Illustration of the evaluation of the mathematical modelling process

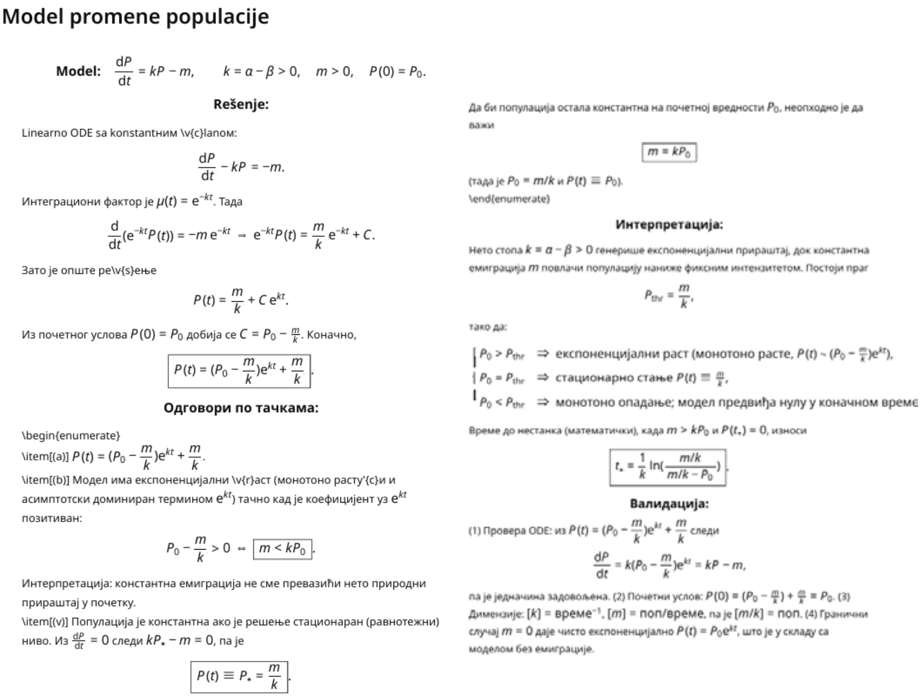
For each aspect and each task, one of three levels of success was assigned: successful, partially successful, and unsuccessful. Solutions categorized as unsuccessful were those in which serious conceptual or procedural errors were identified. The successful category included all solutions that were entirely correct, while partially successful referred to those parts of the solutions that were fundamentally correct but either lacked sufficient detail or precision, or contained a minor computational error in the mathematical result. The following table presents ChatGPT's performance in solving these tasks.

Table 1. ChatGPT's Performance in Solving Mathematical Modelling Tasks

Problem	Mathematization	Mathematical Result	Interpretation	Validation
1	successful	successful	successful	successful
2	successful	partially successful	partially successful	partially successful
3	successful	successful	successful	successful
4	partially successful	successful	partially successful	successful

5	successful	successful	partially successful	successful
6	partially successful	partially successful	partially successful	partially successful

At first glance, it can be seen that none of the four aspects in any of the six tasks was unsuccessful. A detailed analysis of the solutions produced by ChatGPT revealed that the first task was completed successfully. However, in the second task, which concerned a bacterial population model, it made an error in solving the model (although it should be emphasized that the chosen model was appropriate). Specifically, it proceeded with a numerical approach, which led to a computational error that subsequently affected the following steps of the solution. The third task was solved entirely successfully (Fig. 2), whereas in the fourth task – which, unlike the first three, was presented in a tabular form – ChatGPT constructed a model that was not the most suitable solution to the problem. In the fifth task, the processes of mathematization and solution implementation were adequate, but the interpretation of the solution within the appropriate context of game theory could have been better and more refined. The validation process, however, was carried out successfully. In the final task, which was from the field of linear programming, ChatGPT exhibited the greatest weaknesses, especially from the mathematical perspective, as it did not select the most appropriate model, and the objective function was also not well formulated. Nevertheless, the constraints were well defined, and the problem was solved correctly in accordance with them.



Model promene populacije

Model: $\frac{dP}{dt} = kP - m, \quad k = \alpha - \beta > 0, \quad m > 0, \quad P(0) = P_0.$

Rešenje:

Linearno ODE sa konstantним \v{c}lanом:

$$\frac{dP}{dt} - kP = -m.$$

Интеграциони фактор је $\mu(t) = e^{-kt}$. Тада

$$\frac{d}{dt}(e^{-kt}P(t)) = -m\,e^{-kt} \Rightarrow e^{-kt}P(t) = \frac{m}{k}e^{-kt} + C.$$

Зато је опште ре\v{s}ење

$$P(t) = \frac{m}{k} + C\,e^{kt}.$$

Из почетног услова $P(0) = P_0$ добија се $C = P_0 - \frac{m}{k}$. Коначно,

$$\boxed{P(t) = (P_0 - \frac{m}{k})e^{kt} + \frac{m}{k}}.$$

Одговори по тачкама:

\begin{enumerate}
\item[(a)] $P(t) = (P_0 - \frac{m}{k})e^{kt} + \frac{m}{k}$.
\item[(b)] Модел има експоненцијални \v{r}аст (монотоно расту'{c}и и асимптотски доминиран термином e^{kt}) тачно кад је коефицијент уз e^{kt} позитиван:

$$P_0 - \frac{m}{k} > 0 \Leftrightarrow \boxed{m < kP_0}.$$

Интерпретација: константна емиграција не сме превазићи нето природни прираштај у почетку.
\item[(v)] Популација је константна ако је решење стационаран (равнотежни) ниво. Из $\frac{dP}{dt} = 0$ следи $kP_* - m = 0$, па је

$$\boxed{P(t) \equiv P_* = \frac{m}{k}}.$$

Да би популација остала константна на почетној вредности P_0, неопходно је да важи

$$\boxed{m = kP_0}$$

(тада је $P_0 = m/k$ и $P(t) \equiv P_0$).
\end(enumerate)

Интерпретација:

Нето стопа $k = \alpha - \beta > 0$ генерише експоненцијални прираштај, док константна емиграција m повлачи популацију наниже фиксним интензитетом. Постоји праг

$$P_{thr} = \frac{m}{k},$$

тако да:

$P_0 > P_{thr} \Rightarrow$ експоненцијални раст (монотоно расте, $P(t) \sim (P_0 - \frac{m}{k})e^{kt}$),
$P_0 = P_{thr} \Rightarrow$ стационарно стање $P(t) \equiv \frac{m}{k}$,
$P_0 < P_{thr} \Rightarrow$ монотоно опадање; модел предвиђа нулу у коначном време

Време до нестанка (математички), када $m > kP_0$ и $P(t_*) = 0$, износи

$$\boxed{t_* = \frac{1}{k}\ln(\frac{m/k}{m/k - P_0})}.$$

Валидација:

(1) Провера ODE: из $P(t) = (P_0 - \frac{m}{k})e^{kt} + \frac{m}{k}$ следи

$$\frac{dP}{dt} = k(P_0 - \frac{m}{k})e^{kt} = kP - m,$$

па је једначина задовољена. (2) Почетни услов: $P(0) = (P_0 - \frac{m}{k}) + \frac{m}{k} = P_0$. (3) Димензије: $[k]$ = време$^{-1}$, $[m]$ = поп/време, па је $[m/k]$ = поп. (4) Гранични случај $m = 0$ даје чисто експоненцијално $P(t) = P_0e^{kt}$, што је у складу са моделом без емиграције.

Fig. 2. The solution to the third task by ChatGPT in the context of mathematical modelling.

Regarding the students' work, all students solved the first and second tasks completely successfully. In the third task, three students fully carried out the appropriate procedures in all four steps, while one student completed the first two steps entirely successfully but only partially succeeded in interpreting and validating the solution. In the fourth task, two students successfully completed all steps. One student was only partially successful in the mathematization process and in interpreting the solution but fully succeeded in the other two steps. Another student had the same level of success as the previous one in the first three parts but entirely omitted the validation step. In the fifth task, two students completed the task entirely successfully, while the other two were successful in mathematization but only partially successful in the remaining three steps. In the final task, two students again solved it completely successfully. One was partially successful in all steps, while the fourth student was fully successful in the first three steps but partially successful in validation.

To more easily compare the performance of students and ChatGPT, we numerically evaluated their results. Each task was worth 4 points (1 point for each of the four parts of the solution). Each of the four aspects in every task was scored with 1 if that part was fully completed successfully, 0.5 for partially successful, and 0 for unsuccessful. The students scored 24, 22, 20.5, and 20 points, respectively. The arithmetic mean of the students' scores was 21.625 points, with a standard deviation of 1.80 points. Looking at ChatGPT's results on this test, it scored 19 points. Compared to the students' results, the Z-score corresponding to ChatGPT's score is -1.46. This score is relatively low but not extreme, as it falls between one and two standard deviations below the mean. This result indicates that the students still performed better at solving mathematical modelling tasks. On the other hand, some students could benefit from certain aspects of ChatGPT's solutions, particularly in interpreting and validating the results.

6 Discussion and Conclusion

By reviewing the students' work, it can be observed that their shortcomings in mathematical modelling are more pronounced in steps such as interpreting results within the initial context and validating the obtained solution, while weaknesses in mathematizing the problem and carrying out computations are less noticeable. These deficiencies—or factors influencing students' success in mathematical modelling—align with those identified by Sen Zeytun et al. [3], particularly referring to difficulties in mathematization, i.e., establishing connections between the real-world problem and the mathematical context, and vice versa.

When it comes to ChatGPT's performance, we see that it successfully carried out the mathematization process in 4 out of 6 tasks – meaning it understood the problem's context and correctly assigned mathematical meaning to the models. In the same number of cases, it was fully successful in solving mathematical problems, including constructing equations, solving differential equations, and addressing tasks from other mentioned fields. That ChatGPT is successful in solving first-order differential equations was also demonstrated by Koceska et al. [8]. In the remaining two cases, it was partially successful. On the other hand, Li et al. [11] pointed out the shortcomings of ChatGPT in solving linear programming problems compared to an LLM that had been additionally trained to handle that specific type of task. In our research as well, ChatGPT was partially successful when solving the linear programming task. Interestingly, it was only partially successful in both initial steps of the linear programming task. More concerning is that in 4 out of 6 cases, it was only partially successful in interpreting the results, while in the remaining two, it was fully successful. It should be noted that the tasks given to ChatGPT did not include explicit instructions regarding the mathematical modelling process, so future research should examine its solutions in more detail if more specific modelling-related instructions are provided. As for validation, ChatGPT was fully successful in 4 out of 6 cases, while in the remaining two, it was only partially successful.

Based on the obtained results, we can say that ChatGPT largely successfully solved mathematical modelling problems related to the application of differential equations, game theory, and linear programming – though these results could have been better. When comparing ChatGPT's results with those of the students, we observe that all four students achieved better results in terms of solution structure within the mathematical modelling cycle. However, it is also evident that certain parts of ChatGPT's solutions could still be useful to some students, particularly in verifying and validating problem solutions.

Obviously, this research has its limitations. The first limitation concerns the sample size, which stems from the fact that only four students were enrolled in the Mathematical Modelling course during the given academic year. Consequently, our findings cannot be generalized. Furthermore, in this research we examined the capabilities of only one tool - the free version of ChatGPT - because it is accessible to all students without additional subscription costs. These limitations also point to directions for future research. In addition to providing more detailed instructions when assigning tasks within the framework of the mathematical modelling cycle, future studies could also investigate the capabilities of other free tools, such as DeepSeekR1, as well as paid versions of widely adopted tools, including those offered by OpenAI, Claude Opus, Gemini, and others, and compare their performance on tasks completed by larger student samples. Furthermore, future research could focus on developing specialized LLM models designed to address and solve relevant problems in accordance with the principles of mathematical modelling, with the goal of providing higher-quality support to students.

Acknowledgments. This work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through the Agreement No. 451-03-65/2024-03/200122.

Disclosure of Interests. No potential conflict of interest was reported by the authors.

References

[1] Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems. *Mathematical modelling (ICTMA 12): Education, engineering and economics*, 222-231.
[2] Schukajlow, S., Kaiser, G., & Stillman, G. (2018). Empirical research on teaching and learning of mathematical modelling: A survey on the current state-of-the-art. *ZDM*, *50*(1), 5-18.
[3] Sen Zeytun, A., Cetinkaya, B., & Erbas, A. K. (2023). Why do prospective teachers have difficulties in mathematical modelling? Insights from their perspectives. *International Journal of Mathematical Education in Science and Technology*, *55*(1), 94–117. https://doi.org/10.1080/0020739X.2023.2171922
[4] Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., et al. (2023). How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med. Educ. 9*, doi: 10.2196/45312
[5] Svičević, M., Milenković, A., Vučićević, N. & Stanković, M. (2025) Evaluating the Success of AI Tools in Supporting Student Performance in Mathematical Kangaroo Competition, *Computer Applications in Engineering Education 33*(4), https://doi.org/10.1002/cae.70063
[6] Milenković, A., Vučićević, N., & Svičević, M. (2025) Evaluating open ai tools o1 and o3-mini in solving high school problems from Serbian national mathematics competition, *The teaching of mathematics, XXVIII* (1), 14 - 29, DOI:10.57016/TM-EYRQ1676
[7] Helfrich-Schkarbanenko, A. (2023). Mathematik und ChatGPT: ein Rendezvous am Fuße der technologischen Singularität. Springer Spektrum.
[8] Koceska, N., Koceski, S., Lazarova, L. K., Miteva, M., & Zlatanovska, B. (2023). Can ChatGPT be used for solving ordinary differential equations. *Balkan Journal of Applied Mathematics and Informatics*, *6*(2), 103-114.
[9] Zhu, T., Li, X., & Zhang, W. (2023). Applying chatgpt-powered game theory in ransomware negotiations. *Authorea Preprints*.
[10] Huang, X., Shen, Q., Hu, Y., Gao, A., & Wang, B. (2024). Llms for mathematical modeling: Towards bridging the gap between natural and mathematical languages. *arXiv preprint arXiv:2405.13144*.
[11] Li, Q., Zhang, L., & Mak-Hau, V. (2023). Synthesizing mixed-integer linear programming models from natural language descriptions. *arXiv preprint arXiv:2311.15271*.

Convergence of Artificial Intelligence, Ontologies, and Qualia in e-learning re-design

Lamprini Seremeti[1][0000-0002-0663-5408], Charis Kalogera[1][0009-0005-6762-3944], Stylianos Papalexandris[2,3][0000-0001-9248-2067], and Sofia Anastasiadou[3][0000-0001-6404-5003]

[1] Agricultural University of Athens, 75 Iera Odos Street, Votanikos, 118 55, Athens, Greece
[2] Democritus University of Thrace, Greece
[3] University of Western Macedonia, Active Urban Planning Zone (ZEP), 50 150, Kozani, Greece

Abstract. Nowadays, e-learning systems rely heavily on artificial intelligence (AI) as the core operational mechanism to deliver adaptive, personalized, and scalable learning experiences. However, the effectiveness of e-learning depends not only on computational capabilities but also on the conceptual basis of the knowledge under representation. Ontologies provide the formalized frameworks through which AI-based educational systems organize, correlate, and retrieve the educational content. Beyond the technical and structural aspects of e-learning lies the deeper philosophical foundation, that of Qualia, which refer to the phenomenological and subjective learners' sense, such as their perception of continued support, their engagement with the content in a meaningful way, and the quality of sensation on learning. This paper introduces the need for an e-learning entity which is based on the interplay between AI, Ontologies, and Qualia, arguing for an integrated framework that involves computational intelligence, semantic structure, and philosophical foundation.

Keywords: E-learning, Artificial Intelligence, Ontologies, Qualia.

1 Introduction

The ultimate goal of education is the development of individual autonomy, otherwise self-governance in the deepest and broader sense, i.e. intellectual, emotional, ethical, and social self-realization [1], [2]. This core purpose must remain the same also in the context of e-learning [3]. Autonomy does not simply mean the ability to study independently, but the cultivation of critical thinking, self-determination, emotional awareness, and the capacity to make informed decisions [4]. Autonomous learners are guided toward becoming active participants in their own learning path, not passive recipients of information [5]. Designing e-learning systems that recognize this complexity ensures that the digital transformation of education remains faithful to its most human-centered goal.

E-learning has been flourishing in recent years, largely as a result of rapid advancements in artificial intelligence (AI) and the implementation of educational ontologies

Reveiw Paper
DOI: https://doi.org/10.46793/eLearning2025.193S
Part of ISBN: 978-86-89755-37-4

that can be used to organize content in any field and thus make it machine interpreted [6]. Both technologies have revolutionized how learners interact with knowledge, offering adaptive learning paths [7]. Indeed, the extensive body of literature clearly indicates that the design and implementation of e-learning systems are primarily grounded in the achievements of artificial intelligence, often in conjunction with educational ontologies [8 – 13]. These two pillars form the operational and structural backbone of most modern digital learning platforms. Through them, systems model learner behavior, and personalize content. However, there is a scarcity of research that identifies philosophical dimensions [14], [15] as a third foundational element in the development of e-learning environments. The novelty of this paper is the inclusion of Qualia, that is the science of subjective experience, into e-learning design. Even though philosophy of mind has debated qualia for decades, computer science lacks a shared conceptual and formal schema to represent personalized states (e.g., frustrated, overwhelmed, pleased, surprised, etc.) in ways that are computable and interoperable [16]. While AI and Ontologies have provided the operational and structural foundation for e-learning, the integration of Qualia introduces the essential human dimension of how do learners feel. Incorporating such philosophical perspective could radically enhance the way educational technologies relate to learners, not only as cognitive agents but as sentient individuals. Bridging this gap may be essential for the next generation of e-learning.

In this vein, alongside AI (i.e. the operational mechanism) and Ontologies (i.e. the semantic structure), Qualia are introduced as the qualitative basis of e-learning design. The concept of Qualia is not claimed access to consciousness, but standardization of attributions of learners' experience (e.g., confusion, anxiety, curiosity, etc.). Based on the AI-Ontologies-Qualia triad, AI can estimate such attributions from data, Ontologies can guarantee consistency, and Qualia can direct pedagogical interventions toward the quality of learning.

The motivation for writing this paper is to obtain information about the basic terms and the use of AI, ontologies, and qualia in education in order to get better guidance when creating next generation e-learning. In this sense, the aim of this paper is to present these three distinctive but enmeshed domains and give an overview of recent research in the field, in the context of education. The research question of this paper is focused on finding how to combine AI, ontologies, and qualia in e-learning design in order to overcome its current limitations and meet the educational needs of the next generation.

The paper is structured as follows. Section 2 explores the theoretical background of AI, ontologies, and qualia in e-learning. Next, the interplay among the three domains is discussed in Section 3. Future research directions are discussed in Section 4. Finally, Section 6 concludes the paper.

2 Theoretical Background

This section broadly conceptualizes AI, Ontologies, and Qualia as the main theoretical building blocks of next generation e-learning systems design. The aim is not a systematic literature review [17] nor a systematic mapping study [18]; rather is conceptual clarity of the corresponding domains in order to motivate their interrelation in the remainder of the paper.

2.1 Artificial Intelligence as Operational Mechanism

Artificial Intelligence (AI) is the subfield of computer science which offers a spectrum of technologies, including machine learning, deep learning, knowledge representation and reasoning, computer vision, natural language processing, in order to enhance efficiency, productivity, and decision-making in all sectors of everyday life [19].

In the education sector, AI has recently been adopted and used, particularly in the design of digital educational environments [20], [21]. Innovations such as intelligent tutoring systems, personalized and adaptive learning, conversational agents, predictive analytics, gamification, and AI-based assessment confirm the growing impact of AI on learner engagement, instructional design, and performance monitoring [22]. However, the integration of AI in e-learning also presents significant challenges [23]. These include data privacy, algorithmic bias, transparency, and educator resistance [24].

Indeed, bibliography on AI in education is both vast and rapidly expanding, reflecting the increasing stakeholders' interest in leveraging AI to transform learning and teaching. The growing body of literature can broadly be divided into two major categories. The first one focuses on specific applications of AI in e-learning, like natural language processing in detecting misconceptions of learners about a specific concept [25]; intelligent tutoring systems development for providing learners with personalized, adaptive, and interactive instructions [26], [27]; adaptive and personalized learning platforms design that customize content to individual learner progress [28]; and conversational agents which answer questions, explain concepts, provide feedback, guide learners through problem-solving steps, and simulate role-play scenarios for skills development [29]. Studies in this category emphasize mainly technical design, implementation, and performance evaluation, demonstrating how AI technologies can enhance instructional delivery, personalize learning experiences, and improve educational outcomes. The second one addresses the specific challenges associated with AI integration in education, especially ethical concerns [30], such as algorithmic bias [31], data privacy [32], academic integrity [33], [34], teachers' role [35], as well as issues like pedagogical alignment [36], and the potential of over-reliance on automation [37].

This path of research tends to adopt a critical, interdisciplinary perspective, incorporating insights from education, computer science, philosophy and social science. Both categories form a comprehensive framework for understanding the integration of AI in education, where technological innovation and critical evaluation progress in parallel. The application-focused research drives the development of advanced learning systems and tools that can enhance teaching and learning, while the challenge-oriented scholarship ensures these innovations are implemented in ethical and pedagogically sound way.

In a nutshell, AI in education is an interdisciplinary field that studies how AI innovations are integrated into educational settings. However, most research focuses on technological and ethical aspects, which cannot achieve a deep understanding of the complex nature of instant sense-making of learners. Indeed, current AI-based educational platforms, as operational mechanisms, can produce detailed analytics, personalized recommendations, and adaptive feedback, but without addressing the subjective, emotional experiences of learners, these outputs may remain underutilized.

2.2 Ontologies as Knowledge Structure

Ontologies aim at capturing the semantics of domain expertise by deploying knowledge representation schemes, enabling a machine to understand the relationships between concepts in a domain [38].

In short, ontologies are semantic web technologies that are thought as the structure of web-based applications [39]. They explicitly define the relationships between concepts within a specific domain, enabling shared understanding and interoperability among systems and humans [40]. They are rooted in philosophy, especially by using the term in singular, that is "ontology", to refer to a unified and fundamental branch of metaphysics, but then they widely adopted, in plural, in computer science as "domain ontologies", by providing a formal representation for knowledge structure, thus enabling data integration, and reasoning [41]. Nowadays, they are no longer only theoretical approaches; they are semantic tools to facilitate the goals of sharing, and reusing knowledge [42].

Traditionally, studies on ontologies in the context of education are viewed from two perspectives, "ontologies in education" and "educational ontologies". Although these terms are often used interchangeably, they signify distinct research directions. The first term refers broadly to the application of ontology engineering within educational settings. For instance, ontologies were used for curriculum modelling and management, for describing learning domains, learning data, and e-learning services [43], [44]. Recent studies in the field of e-learning refer that ontologies allow systems to adapt content based on learner profiles, preferences, and performance [45]; allow institutions to track performance and enhance decision-making [46]; aligns gamification elements with cognitive models to improve learner engagement [47]; and facilitate curriculum modeling and knowledge alignment across different platforms, fostering standardization [8]. On the other hand, educational ontologies are specific domain ontologies that focus on the standardization of specific learning domains and educational concepts, such as learning objectives through Bloom's taxonomy, instructional strategies, or assessment types [48 – 50].

In summary, ontologies have gained the attention of scientists in e-learning field since they are used as tools for organizing and managing information due to their inherent features, including resource sharing and reuse, knowledge modeling, and inference.

2.3 Qualia as the Philosophical Foundation

Qualia are the felt, subjective "what it is like" aspects of experience [51]. The term became central in late 20th century debates about mind and consciousness [52]. As a theoretical foundation of computer science, Qualia mark the difference between processing information and experiencing it [53 – 55].

Indeed, scholars and engineers have already explored the conceptual, ethical, and architectural implications of qualia within AI. One of the dominant themes in the latest literature is that while AI systems have made significant progress in simulating emotional expressions [56], the problem of simulating subjective feelings remain unresolved [57]. Another notable area of recent expansion lies in affective computing, particularly with respect to ethical implications, such as data privacy and user satisfaction [58 – 60]. In a broad sense, the research on qualia in computer science, and, in particular in AI, reflects a rich multidisciplinary dialogue. The field is advancing rapidly, from abstract philosophical studies to concrete architectural experiments [61].

Especially, in the context of education, qualia are rarely discussed explicitly in e-learning literature; they are indirectly related with AI-based learning systems. For instance, the design of emotion-sensitive intelligent tutoring systems relies on tracking facial expressions, or voice tone to infer user emotions and adapt content accordingly [62]. These systems approximate qualia by interpreting behavioral patterns [63]. Despite substantial progress in e-learning domain, by relating education, cognitive science and neuroscience [64], qualia remain unutilized in explaining the e-learning processes that are accompanied by felt experience [65]. Unfortunately, for e-learning designers, many of the desirable traits are not inherently measurable via traditional, quantitative means, but they are emergent properties dependent on qualitative aspects of mechanistic, unconscious reactions.

3 The Interplay of AI, Ontologies, and Qualia in E-learning

Re-designing and modernizing e-learning is now imperative to satisfy the diverse disclosure requirements of learners, educators, institutions, regulators, and industry partners. This goal requires a coherent convergence of AI for operational intelligence and personalization, Ontologies for structured and interoperable knowledge, and Qualia to capture the phenomenological and affective dimensions of learning.

Although there is no existing body of literature that explicitly focuses on the intersection of AI, Ontologies, and Qualia in e-learning, a number of studies across these domains provide conceptual and technical insights for their interrelation. These works, while often studying the concepts in isolation, collectively contribute to a broader understanding of how subjective experience (Qualia), structured knowledge representation (Ontologies), and intelligent adaptation (AI) may converge in the design of next-generation e-learning. More precisely, AI is increasingly used to personalize learning experiences through adaptive algorithms and intelligent tutoring systems [66], [67] while ontologies serve as structured frameworks that formalize domain knowledge and learners' models [68], [69]. Simultaneously, the concept of qualia has gained traction in affective computing and user experience design, particularly in emotionally responsive learning environments [70]. Taken together, these fields open new possibilities for modeling and enhancing not only the cognitive but also the emotional and reactive dimensions of digital learning.

This paper proposes a novel, interdisciplinary theoretical framework for understanding the relational interaction between learners and advanced e-learning models through the lens of the triad AI-Ontologies-Qualia.

The following Table 1 gives an overview of the three pillars work together in designing next generation e-learning.

Table 1. AI-Ontologies-Qualia triad for e-learning.

	Definition in e-learning	Scope in e-learning	Interdependencies
Artificial Intelligence	The application of data-driven algorithms that personalize instruction, automate feedback, and optimize learning pathways at scale	Operational mechanism (augment teaching and personalize learning by using technological advancements)	AI depends on Ontologies for clear meaning of context and on Qualia to prioritize which experiences matter
Ontologies	Formal, shared models of concepts and their relationships that structure educational knowledge for interoperability, personalization, and reasoning	Knowledge structure (formal, shared semantic models that define domains, learner profiles, and instructional processes, enabling interoperability, precise metadata, adaptive personalization, and explainable analytics across platforms)	Ontologies depend on AI to populate instances and on Qualia to formalize experiences
Qualia	The subjective qualities of learners' experience that shape motivation and decision-making	Theoretical foundation (recognition and modelling of the learners' lived experience)	Qualia depend on AI for operationalization and on Ontologies for standardization

The integration of AI, Ontologies, and Qualia-like modeling, although still in the early stages of development, is reshaping e-learning. Ontologies provide the semantic structure of the model, given that they provide the formal knowledge representation of learner preferences and content and AI operates as the main mechanism that enables adaptive and context-sensitive interaction, through the use of machine learning technologies [71], [72]. Qualia, while not directly programmable, as theoretical foundation of the model, serve as motivational goals for creating emotionally and ethically aware learning environments. Though still largely theoretical, such proposal indicates a direction where Qualia, AI, and Ontologies convergence could support e-learning re-design.

4 Future Directions

As future work, the Qualia Ontology for e-learning will be constructed, aligned with existing educational ontologies and in accordance with AI advancements. Its core will include classes and relations standardizing attributions of learner experience with explicit dimensions, modalities, context, and evidence. For instance, the class of LearningEpisode refers to a time-bounded instructional activity for a specific learner, such as watching a video, solving a problem, or taking a quiz. Each LearningEpisode is characterized by a Modality (e.g., visual, auditory, or textual) within which experience arises, and by PhenomenalQuality terms that capture what the LearningEpisode felt like (e.g., clarity, perceived complexity, or novelty). The qualia core is represented by the class AffectiveState, which encodes states such as confusion, curiosity, boredom, anxiety.

This ontology is under construction and its purpose is to encode what learners are likely experiencing a learning path in a consistent and interoperable way. By completing and using this ontology in an e-learning environment, mechanism (i.e. what the system does), meaning (i.e. how things are named and related), and experience (i.e. why it matters for learning) are clearly interwined, thus preventing pedagogical errors and enabling transparent and controllable content adaptation to learners' aspirations.

5 Conclusion

Next generation e-learning seems to stand at the confluence of technological innovation, semantic structure, and philosophical research. AI provides the technological mechanisms for adaptability and efficiency; ontologies offer the conceptual schemes for knowledge representation; and qualia explain the grounds of private feelings. Their interplay can lead to intelligent, coherent, and private thinking enriched e-learning environments. This triadic form may mark a shift toward e-learning systems that respect the spirit of democracy in learning such as, the objectivity of knowledge, the subjectivity of experience, and the equal access to learning.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Cuypers, S.E.: Is personal autonomy the first principle of education? Journal of Philosophy of Education **26**(1), 5-18 (1992)

2. Neufeld, A., Kassam, A.: Why respect for learner autonomy is an ethical priority. The Clinical Teacher **22**(2) (2025)
3. Yang, Y., Chen, J., Zhuang, X.: Self-determination theory and the influence of social support, self-regulated learning, and flow experience on student learning engagement in self-directed e-learning. Frontiers in Psychology **16**, 01-15 (2025)
4. Chong, S.W., Reinders, H.: Autonomy of English language learners: A scoping review of research and practice. Language Teaching Research **29**(2), 607-632 (2025)
5. Drumm, S.: Applying individual strategies enhances learning in asynchronous learning paths. Computers and Education Open **8**, 100257 (2025)
6. Bhadra, P.: The impact of web 3.0 technologies in teaching learning process. In: Biswas, P., Chattopadhyay, S., Biswas, S., Pal, D., Khawas, K. (eds.) Modern Technology in Education, pp. 155-166, Gungun Publishing House (2025)
7. Rathod, A.O., Jadhav, A., Shelke, S.S., Labade, S.B., Kale, V.S., Adhav, S.: Towards smarter e-learning: A machine learning approach to emotion recognition and text analysis. In: Proceedings of the Global Conference in Emerging Technology, pp. 1-6, IEEE, India (2025)
8. Bousalem, S., Benchikha, F., Marir, N.: Personalized learning through MBTI prediction: A deep learning approach integrated with learner profile ontology. IEEE Access **13**, 30861-30874 (2025)
9. El Fezazi, N.E., El Miloud, S., Oumaira, I., Daoudi, M.: Adaptive m=learning content on the moodle platform using m-learner modelling approach, ontologies, and machine learning techniques. Educational Process: International Journal **16**, e2025218 (2025)
10. Dey, A., Ganguly, A., Banik, I.R., Bhuiya, S., Sengupta, S., Das, R.: Smart recommendation system in e-learning using machine learning and data analytics. SN Computer Science **6**, 706 (2025)
11. Maheswari, M., Gayathri, M., Reddy, M.Y., Shabu, S.L.J., Refonaa, J., Dhamodaran: Collaborative recommendation systems for e-learning sources. In: Proceedings of the 7th International Conference on Intelligent Sustainable Systems, pp. 1171-1179, IEEE, India (2025)
12. Kommers, P., Smyrnova-Trybulska, E.: Third decade in the twenty-first century skills and A.I. in education. In: Smyrnova-Trybulska, E., Chen, N.S., Kommers, P., Morze, N. (eds.) E-learning and Enhancing Soft Skills, pp. 1-13, Springer, Cham (2025)
13. Amastini, F., Sekarhati, D.K.S., Puspitasari, M.: Advancing adaptive and personalized e-learning systems: A systematic literature review. Advance Sustainable Science, Engineering and Technology **7**(2), 1-14 (2025)
14. Salimova, R., Nurmanbetova, J., Kozhamzharova, M., Manassova, M., Aubakirova, S.: Philosophy of education in a changing digital environment: An epistemological scope of the problem. AI & Society **40**, 1655-1666 (2025)
15. Masveta, D., Manyangara, M.E.: The UX-UI continuum: Exploring the interplay between user experience and user interface in e-learning platforms. Cogent Education 12(1), 2536531 (2025)
16. Yildiz, T.: The minds we make: A philosophical inquiry into theory of mind and artificial intelligence. Integrative Psychological and Behavioral Science **59**(1) (2025)
17. Carrera-Rivera, A., Ochoa, W., Larrigana, F., Lasa, G.: How-to conduct a systematic literature review: A quick guide for computer science research. MethodsX **9**, 101895 (2022)
18. Taipalus, T.: Systematic mapping study in information systems research. Journal of the Midwest Association for Information Systems **2023**(1), 9-24 (2023)
19. Khaleel, M., Jebrel, A., Shwehdy, D.M.: Artificial intelligence in computer science. International Journal of Electrical Engineering and Sustainability **2**(2), 01-21 (2025)

20. Qin, Q., Zhang, S.: Visualizing the knowledge mapping of artificial intelligence in education: A systematic review. Education and Information Technologies **30**, 449-483 (2025)
21. Samala, A.D., Rawas, S., Wang, T., Reed, J.M., Kim, J., Howard, N.J., Ertz, M.: Unveiling the landscape of generative artificial intelligence in education: A comprehensive taxonomy of applications, challenges, and future prospects. Education and Information Technologies **30**, 3239-3278 (2025)
22. Idris-Tajudeen, R., Akin-Olayemi, T.H., Akinsiku, A.M.: Artificial intelligence in e-learning: A systematic literature review of 21^{st} century trends and innovations. Tech-Sphere Journal for Pure and Applied Sciences **2**(1) (2025)
23. Singh, A.K., Kiriti, M.K., Singh, H.: Education AI: Exploring the impact of artificial intelligence on education in the digital age. International Journal of System Assurance Engineering and Management **16**, 1424-1437 (2025)
24. Anastasiadou, S. D: Exploring AI Ethics Through Factorial Correspondence Analysis. In: Anastasiadou, S., Chadjipadelis, T., Koutsoupias, N., Markos, A., Menexes, G. (eds.) Modern International Developments in Data Analysis and Multivariate Statistical Analysis, pp. 495-528, IGI Global Scientific Publishing (2026)
25. Tlili, A., Saqer, K., Salha, S., Huang, R.: Investigating the effect of artificial intelligence in education (AIEd) on learning achievement: A meta-analysis and research synthesis. Information Development **41**(3), 825-842 (2025)
26. Kokver, Y., Pektas, H.M., Celik, H.: Artificial intelligence applications in education: Natural language processing in detecting misconceptions **30**, 3035-3066 (2025)
27. Letourneau, A., Deslandes Martineau, M., Charland, P., Karran, J.A., Boasen, J., Leger, P.M.: A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. Npj Science of Learning **10**(29) (2025)
28. Zhong, X., Zhan, Z.: An intelligent tutoring system for programming education based on informative tutoring feedback: System development, algorithmic design, and empirical study. Interactive Technology and Smart Education **22**(1), 3-24 (2025)
29. Rizvi, I., Bose, C., Tripathi, N.: Transforming education: Adaptive learning, AI, and online platforms for personalization. Technology for Societal Transformation, 45-62 (2025)
30. Yusuf, H., Money, A., Daylamani-Zad, D.: Pedagogical AI conversational agents in higher education: A conceptual framework and survey of the state of the art. Educational Technology Research and Development **73**, 815-874 (2025)
31. Rohrl, S., Staufer, S., Bugert, F., Nadimpalli, V.K., Hauser, F., Grabinger, L.: Ethical considerations of AI in education: A case study based on Pythia learning environment system. IEEE Access **13**, 115331-115353 (2025)
32. Boateng, O., Boateng, B.: Algorithmic bias in educational systems: Examining the impact of AI-driven decision making in modern education. World Journal of Advanced Research and Reviews **25**(01), 2012-2017 (2025)
33. Mienye, I.D., Swart, T.G.: ChatGPT in education: A review of ethical challenges and approaches to enhancing transparency and privacy. Procedia Computer Science **254**, 181-190 (2025)
34. Seremeti, L., Anastasiadou, S., Masouras, A., Papalexandris, A.: Addressing the challenges of harmonizing law and artificial intelligence technology in modern society. IAES International Journal of Artificial Intelligence 13(3), 2471–2478 (2025)
35. Balalle, H., Pannilage, S.: Reassessing academic integrity in the age of AI: A systematic literature review on AI and academic integrity. Social Sciences & Humanities Open **11**, 101299 (2025)

36. Mouta, A., Torrecilla-Sanchez, E.M., Pinto-Liorente, A.M.: Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. Education and Information Technologies **30**, 3343-3387 (2025)
37. Dai, Y.: Integrating unplugged and plugged activities for holistic AI education: An embodied constructionist pedagogical approach. Education and Information Technologies **30**, 6741-6764 (2025)
38. Khlaif, Z.N., Hamamra, B., Hussein, E.T.: AI paradox in higher education: Understanding over-reliance, its impact, and sustainable integration. Research Square 1-31 (2025)
39. Mayer, R., Antunes, G., Caetano, A., Bakhshandeh, M., Rauber, A., Borbinha, J.: Using ontologies to capture the semantics of a (business) process for digital preservation. International Journal on Digital Libraries **15**, 129-152 (2015)
40. Ding, L., Kolari, P., Ding, Z., Avancha, S.: Using ontologies in the semantic web: A survey. In: Sharman, R., Kishore, R., Ramesh, R. (eds.) Ontologies. Integrated Series in Information Systems **14**, 79-113. Springer, Boston, MA (2007)
41. Zhang, X., Zhao, C., Wang, X.: A survey on knowledge representation in materials science and engineering: An ontological perspective. Computers in Industry **73**, 8-22 (2015)
42. Busse, J., Humm, B.G., Lubbert, C., Moelter, F., Reibold, A., Rewald, M., Schluter, V., Seiler, B., Tegtmeier, E., Zeh, T.: Actually, what does "ontology" mean? A term coined by philosophy in the light of different scientific disciplines. Journal of Computing and Information Technology **23**(1), 29 (2015)
43. Tan, S.Z.K., Baksi, S., Bjerregaard, T.G., Elangovan, P., Gopalakrishnan, T.K., Hric, D., Joumaa, J., Li, B., Rabbani, K., Venkatesan, S.K., Valdez, J.D., Kuriakose, S.V.: Digital evolution: Novo Nordisk's shift to ontology-based data management. Journal of Biomedical Semantics **16**(6) (2025)
44. Kougias, I., Seremeti, L., Kalogeras, D.: Ontology-based knowledge management in NGEEs. International Journal of Pure and Applied Sciences and Technology **2**(1), 54-62 (2011)
45. Stancin, K., Poscic, P., Jaksic, D.: Ontologies in education – state of the art. Education and Information Technologies **25**, 5301-5320 (2020)
46. Yuen, H.K., Man, H., Cheung, R.: Rule-based reasoning for personalized learning in a web-based programming environment. In: Proceedings of the 2025 IEEE 14th International Conference on Engineering Education, Malaysia (2025)
47. Musa, M.H., Salam, S., Norasikin, M.A., Shabarudin, M.S., Lestari, U.: A systematic agile framework for test-driven ontology validation in academic performance analytics and decision-making. Journal of Theoretical and Applied Information Technology **103**(11) (2025)
48. Gomaa, Y., Lahoud, C., Abel, M.H., Moussa, S.: Managing data heterogeneity for ontology-driven models: Application to gamified e-learning contexts. In: Proceedings of the 17th International Conference on Computer Supported Education, pp. 726-736, SciTePress (2025)
49. Seremeti, L., Aggelopoulou, N., Pierrakeas, C., Kameas, A.: Binary relations in educational ontologies. In: Badica, A., Colhon, M. (eds.) Proceedings of the 2015 Balkan Conference on Informatics: Advances in ICT, pp. 45-50 (2015)
50. Conde, A., Larranaga, M., Arruarte, A., Elorriaga, J.A.: A combined approach for eliciting relationships for educational ontologies using general-purpose knowledge bases. IEEE Access **7**, 48339-48355 (2019)
51. Villegas-Ch, W., Garcia-Ortiz, J.: Enhancing learning personalization in educational environments through ontology-based knowledge representation. Computers **12**(10) (2023)
52. Chumley, L.H., Harkness, N.: Introduction: QUALIA. Anthropological Theory **13**, 1-2 (2013)

53. Kauffman, S.A., Roli, A.: What is consciousness? Artificial intelligence, real intelligence, quantum mind and qualia. Biological Journal of the Linnean Society **139**(4), 530-538 (2023)
54. Haikonen, P.O.: Qualia, consciousness and artificial intelligence. Journal of Artificial Intelligence and Consciousness **09**(03), 409-418 (2022)
55. Pandey, K.K., Jokhakar, V.: Research and applications of qualia – The unit of subjective experience. International Journal of Innovative Technology and Exploring Engineering **9**(7), 207-214 (2020)
56. Gomez, J.D.: A harder problem of consciousness: Reflections on a 50-year quest for the alchemy of qualia. Frontiers in Psychology **16**, 01-04 (2025)
57. Ross, P., Atkinson, A.P.: Expanding simulation models of emotional understanding: The case for different modalities, body-state simulation prominence, and developmental trajectories. Frontiers in Psychology **11** (2020)
58. Pagan, A., Loveland, K.A., Acierno, R.: Evaluating the emotional accuracy of AI-generated facial expressions in neurotypical individuals. Discover Computing **28**, 116 (2025)
59. Devillers, L., Cowie, R.: Ethical considerations on affective computing: An overview. Proceedings of the IEEE **111**(10), 1445-1458 (2023)
60. Wu, J.: Social and ethical impact of emotional AI advancement: The rise of pseudo-intimacy relationships and challenges in human interactions. Frontiers in Psychology **15** (2024)
61. Baker, D., Tippireddy, M.K.R., Farhan, A., Ahmed, B.: Ethical considerations in emotion recognition research. Psychology International **7**(2), 43 (2025)
62. Wang, X.: The possibility of artificial qualia. Communications in Humanities Research **6**(1), 67-71 (2023)
63. Petrovica, S., Anohina-Naumeca, A., Ekenel, H.K.: Emotion recognition in affective tutoring systems: Collection of ground-truth data. Procedia Computer Science **104**, 437-444 (2017)
64. Fernandez-Herrero, J.: Evaluating recent advances in affective intelligent tutoring systems: A scoping review of educational impacts and future prospects. Education Sciences **14**(8), 839 (2024)
65. Alkhassawneh, S., Al Sharif, H.: Perspectives of brain research (educational neuroscience) on the design and implementation of teaching strategies in educational technology. Journal of Neuroeducation **5**(2), 14-24 (2025)
66. Luo, Y., Sun, L.: The effects of social and spatial presence on learning engagement in sustainable e-learning. Sustainability **17**, 4082 (2025)
67. Saragih, R.I.E.: Ai-powered education: Transforming learning through personalized and scalable solutions. International Journal of Information System and Innovative Technology **3**, 1-9 (2025)
68. Naayini, P.: AI and the future of education: Advancing personalized learning and intelligent tutoring systems. Frontiers in Educational Innovation and Research **1**(1), 29-40 (2025)
69. Wang, Y., Wang, Y.: A survey of ontologies and their applications in e-learning environments. Journal of Web Engineering **20**(6), 1675-1720 (2021)
70. Alseny, C., Ilham, D., El Kahatir, H.: Revolutionizing distance learning: The impact of ontology and the semantic web. Computer Sciences and Mathematics Forum **10**(1), 16 (2025)
71. Gutierrez, R., Villegas-Ch, W., Govea, J.: Development of adaptive and emotionally intelligent educational assistants based on conversational AI. Frontiers in Computer Science **7**, 1628104 (2025)
72. Abdullah, S.R.A., Al-Azawei, A.: Predicting online learners' performance through ontologies: A systematic literature review. The International Review of Research in Open and Distributed Learning **26**(1) (2025)

73. Seghroucheni, O.Z., Lazaar, M., Al Achhab, M.: Systematic review and framework for AI-driven tacit knowledge conversion methods and machine learning algorithms for ontology-based in e-learning platforms. International Journal of Interactive Mobile Technologies **19**(1), 126 (2025)

A Comparative Analysis of Chatgtp-4/4.5 and Human-written Summaries in Linguistic Research

Aleksa Stošić[1[0009-0005-4601-0401]], Marija Milojković[1[0000-0003-4787-1017]], and Aleksandra Branković[1[0009-0007-7810-5581]]

[1] Belgrade Metropolitan University, Tadeuša Košćuška 63, Belgrade 11158, Serbia
aleksa.stosic@metropolitan.ac.rs
marija.milojkovic@metropolitan.ac.rs
aleksandra.brankovic@metropolitan.ac.rs

Abstract. This study evaluates the potential of ChatGPT-4 and ChatGPT-4.5 as research assistants in applied linguistics (AL) by examining their ability to generate annotated bibliographies of research articles. Five AL papers on technology in English pronunciation and speaking instruction were summarized by both models and by human researchers, producing 25 summaries. Fourteen expert raters assessed the summaries for quality and judged their authorship. Results show that both models produced factually accurate and structurally faithful summaries. However, both models lacked critical selectiveness, could only provide generalized statements on relevance, and relied on surface-level markers to assess credibility. Quantitative analysis indicated that ChatGPT summaries were rated as comparable in quality to human-authored ones, though inter-rater agreement was low and a bias against texts perceived as AI-generated was observed. Qualitative findings revealed that experts distinguished AI from human summaries based on information density, word choice, stylistic naturalness, and evaluative engagement. Overall, ChatGPT proved advantageous in accuracy, structural consistency, and efficiency, but its weaknesses in evaluative depth and authenticity suggest that, while it can accelerate the early stages of literature review, it cannot substitute for the nuanced judgment and interpretive reasoning required in applied linguistics.

Keywords: ChatGPT, applied linguistics, research assistant, research summarization, annotated bibliography.

1 Introduction

The rapid advancement of generative artificial intelligence (AI), especially through large language models (LLMs) such as ChatGPT, has had a wide impact on the academic community and the process of conducting research. Since its public release in November 2022, ChatGPT has been employed for tasks such as generating abstracts and summarizing research articles [1], while some authors have even used it as a co-author [2]. ChatGPT has undergone several iterations, and its current versions at the time of writing, ChatGPT-4 and its successor ChatGPT-4.5, have much enhanced

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.205S
Part of ISBN: 978-86-89755-37-4

capabilities compared to previous iterations such as ChatGPT-3 [3][4]. These developments suggest an expanding potential for using ChatGPT as a research assistant.

So far, ChatGPT has been used as a research assistant in diverse fields, including medicine for generating research article abstracts [5], international business and management for summarizing from citations, retrieving citations from those summaries, and linking them back to the original abstracts [6]. In the field of linguistics, Bae's [7] demonstration study illustrated how ChatGPT-3.5 and ChatGPT-4 could assist experimental linguistics by suggesting sources for literature reviews, supporting experimental design, and facilitating statistical analyses. The author concluded that the tools could save researchers time and streamline data preparation, but constant verification was necessary as the tools tended to exhibit shallow reasoning and produce hallucinations[1] and outdated information. Furthermore, Uchida [8] assessed ChatGPT-3.5 for core corpus-linguistic tasks such as frequency analysis, collocation identification, and genre classification, finding partial alignment with established corpus data and highlighting significant limitations, particularly in genre analysis, suggesting that the tool should be used as an auxiliary rather than primary tool for rigorous research. However, there is a notable gap in the literature as to how ChatGPT can be used as an assistant for summarizing research articles. Its use could greatly assist researchers during the time-consuming stage of literature preview, which is foundational for subsequent stages of the research process. Reluctance to use ChatGPT for such purposes likely stems from its tendency to generate inaccurate information and hallucinate, which puts its ability to reliably summarize information into question [1] [9]. Ethical concerns have also been raised, with several studies demonstrating that ChatGPT can produce plagiarized information when used as a research assistant [10] [11] [12]. Nonetheless, if the accuracy of ChatGPT's latest models continues to improve, ChatGPT could substantially benefit researchers by reducing the time spent on reviewing literature, and allow for more time to be invested in conducting practical linguistic research. While field-specific applications vary, this study focuses on applied linguistics (AL), with the expectation that insights into ChatGPT's summarization capacity may yield benefits across disciplines.

Therefore, this study seeks to evaluate the performance of ChatGPT-4 and ChatGPT-4.5 in summarizing AL research papers by addressing the following research questions:

1. To what extent do ChatGPT-4 and ChatGPT-4.5 produce accurate summaries that capture key findings while adhering to the required structural conventions?
2. In what respects do ChatGPT-generated summaries differ from human-authored summaries?

By evaluating the performance of these models, this study aims to provide insight into the advantages and limitations of using ChatGPT for summarizing linguistic research papers, while also offering potential insights for the broader development of artificial intelligence. To the best of our knowledge, as of August 2025, no peer-reviewed

[1] AI generated content that is fluent and confident but factually incorrect, fabricated, or unsupported by its training data.

study has systematically examined ChatGPT's ability to summarize AL research articles. This study addresses that gap by analyzing summaries of AL papers generated by ChatGPT-4 and ChatGPT-4.5 and comparing them with those produced by human authors, drawing on expert evaluations to determine relative quality and reliability.

2 Previous Research

2.1 Technology Acceptance Models and the Use of ChatGPT as a Research Assistant

Previous research on the acceptance of ChatGPT has examined the factors that influence its adoption as a research assistant, but concerns about its accuracy and reliability for summarization tasks have also been raised. These studies primarily rely on the factors outlined in the Technology Acceptance Model (TAM) [13] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [14]. These theories are closely connected as TAM focuses on perceived usefulness and ease of use, while UTAUT refines these factors and expands them with performance expectancy, effort expectancy, social influence, and facilitating conditions. Empirical studies have shown that there is a high acceptance rate of LLMs among academics, and that its adoption is primarily shaped by usefulness, ease of use, and researcher competence, alongside factors such as AI's perceived intelligence, enjoyment, social influence, and institutional practices [15]. Balaskas et al. [16] reported similar findings, adding that age and prior AI experience also play a moderate role. Nevertheless, concerns about accuracy persist. Salleh [6] documented frequent errors of omission (excluding relevant information), and commission (including irrelevant information) when ChatGPT is used as a research assistant. They noted that while ChatGPT can generate relevant summaries from citations, it often fails to understand the papers' theoretical significance. Moreover, it often mistakenly attributed the work to other authors, and often exhibited the Matthew Effect by showing a bias towards more prominent authors. These findings suggest that ChatGPT can help in summarization, but remains unreliable for citation and theoretical integration. Similarly, Rahman et al. [17] found that, even though ChatGPT-3.5 could generate acceptable abstracts if given precise prompts, it struggled to evaluate the studies critically and connect their significance when used for literature review. Therefore, in addition to examining whether ChatGPT-4 and ChatGPT-4.5 can accurately summarize AL research papers, this study also assesses their ability to critically evaluate research papers as an essential component of the literature review process.

2.2 Evaluation Criteria for ChatGPT-generated Summaries

When evaluating the effectiveness of ChatGPT in generating summaries of research papers to support AL research, the accuracy and faithfulness of the information in the generated output constitute a sine qua non condition and therefore represent the primary criterion in the present study.

Previous research has frequently assessed ChatGPT-generated summaries using ROUGE (Recall-Oriented Understudy for Gisting Evaluation), a set of automatic evaluation metrics used to measure lexical overlap between generated and reference summaries to determine how much of the essential content is captured [18]. Goyal et al. [19] reported that ChatGPT-3 exhibited lower ROUGE scores than traditional summarization methods when summarizing news articles and other similar content. In contrast, Yang et al. [20] found ChatGPT-3's performance on summarizing news, dialog transcripts, and Reddit posts to be on pair with traditional methods. Zhang et al. [21] investigated the performance of the more advanced model ChatGPT-3.5 in extractive summarization of diverse text types, including news articles, scientific articles, and government reports. Their findings also showed that ChatGPT-3.5 performed worse in terms of ROUGE scores when compared to traditional methods, but yielded better results in terms of faithfulness when provided with step-by-step summarization instructions. Such procedural prompting reduced the occurrence of hallucinations and errors, which earlier studies had identified as recurrent problems in ChatGPT-3 and ChatGPT-3.5 [19] [21] [22].

Although ROUGE remains widely used, Fabbri et al. [23] demonstrated that it was insufficient for evaluating academic or scientific summarization, as it failed to capture deeper qualities such as factual accuracy and coherence, and called for more refined metrics and the inclusion of human judgment. Responding to this limitation, Hake et al. [5] evaluated ChatGPT-3.5's summarization ability in medical research by asking experts to assess summaries or research articles generated from abstracts. Their findings confirmed that the tool could be of assistance since study participants claimed that summaries were high in quality, had high accuracy, and low bias, and they were often able to classify if the articles were relevant for different medical specialties. Nonetheless, the authors stressed that full-text evaluation remained necessary before making final research selection. To date, no comparable human-judgment studies have been conducted in applied linguistics, and the present study seeks to address this gap by drawing on the expertise of professional linguists to evaluate the effectiveness of ChatGPT-4 and ChatGPT-4.5 for AL research summarization.

Apart from accuracy and faithfulness, the quality of a summary also depends on its adherence to the conventions of academic writing. As a subgenre of academic discourse, summaries are expected to follow a recognizable structure and exhibit characteristic linguistic features [24]. These features include conciseness, high information density, explicit cohesion achieved through signaling devices, and the use of evaluative language to highlight significance or limitations. In the context of applied linguistics research, this further requires the consistent use of appropriate disciplinary terminology, since precision in technical language is essential for accurately conveying research findings and theoretical constructs. Accordingly, this study investigates whether ChatGPT-generated summarizes demonstrate these linguistic qualities.

As the structure of a summary may vary depending on particular research goals, the quality of ChatGPT-generated summaries should be judged by whether the tool can adhere to the required format. In this study, evaluation is based on a modified annotated bibliography format, as outlined by the Purdue Online Writing Lab of the College of Liberal Arts of Purdue University, though other formats could also be used when prompting ChatGPT to summarize research papers. The term annotated bibliography refers to a list of research sources where each entry includes the full citation and a short paragraph that describes the source's content and evaluates its relevance and quality, while also providing a statement why the source is relevant for the researcher's own work. It is thus a more robust format than a summary, which only condenses the main ideas of a single article without providing an evaluation. Therefore, a further criterion for assessing the effectiveness of ChatGPT-4 and ChatGPT-4.5 in AL research summarization is their ability to follow the prescribed structure specified in the prompt. The annotated bibliography format was selected because it not only requires presentation of the topic, methodology, and key findings of the research article, but also demands a critical evaluation of the article's credibility and its relevance to the researcher's topic and goals. As Rahman et al. [17] and Salleh [6] found, ChatGPT-3.5 lacks the ability to critically evaluate information for literature review, so employing the annotated bibliography format provides a means of testing whether the current models, ChatGPT-4 and ChatGPT-4.5, demonstrate improved capacity for critical evaluation alongside summarization.

3 Methodology

Given the insights from previous research and the evaluation criteria identified, the present study adopts a mixed-method approach to assess the summarization performance of ChatGPT-4 and ChatGPT-4.5 in applied linguistics and determine whether the two models have improved capabilities and become more suitable to be used as a research assistant.

3.1 Topic Selection and Corpus Selection

A research topic in the field of applied linguistics, the integration of technology in English language pronunciation and speaking instruction, was chosen. This topic was selected for practical reasons, as the researchers involved in this study possess professional expertise in this area and are therefore well positioned to evaluate the quality of the generated summaries. It is assumed that findings concerning the factual accuracy, faithfulness, structural conformity, and academic language features of ChatGPT-generated summaries in this domain are likely generalizable to other areas of applied linguistics.

Five research papers published between 2022 and 2024 were selected to form the study corpus [25] [26] [27] [28] [29]. These papers were selected on the basis that the subject matter of these papers closely aligns with our own interests in AL research, and because the survey participants would also be familiar with these topics. They were drawn from peer-reviewed journals to ensure scholarly credibility and rigor. All were experimental studies, as this methodological uniformity helped minimize variability in the source material and enabled a fairer comparison between ChatGPT-generated and human-authored summaries. This corpus size was deemed sufficient to obtain meaningful findings, while acknowledging that larger corpora might yield additional insights beyond the scope of the present study.

3.2 Human-written Summaries

The researchers in this study (three junior lecturers in Applied English and Linguistics at Belgrade Metropolitan University who have produced peer-reviewed publications in the field of applied linguistics) independently produced summaries of the five articles using the modified Purdue OWL annotated-bibliography template. The annotation procedure comprised the following steps:

- State what the topic of the paper is and what it is about.
- Summarize the methodology and most important results of the paper (if applicable).
- Evaluate the credibility of the cited works.
- State how the findings or methodology of the paper are applicable to our own research paper.

3.3 ChatGPT-generated Summaries

Each of the five articles was uploaded separately to ChatGPT-4 and ChatGPT-4.5. In line with Zhang et al. [21], both models received step-by-step instructions specifying the annotated-bibliography format and required elements to minimize hallucinations and improve faithfulness. The same prompt was used for both models to ensure comparability. The exact prompt text is provided below.

- You are working on a research paper that discusses the integration of technology and AI into the ESL classroom. You want to summarize the following research paper and write the summary in the style of an annotated bibliography. You need to write the summary using the following template. Do not omit any information.
- Provide a citation for this paper following the APA7 style of citation.
- State what the paper is about and what topic it discusses.
- State what the methodology of the paper is and how the research was conducted.
- State what the most important results of the paper were.
- Provide an evaluation of the paper's credibility and explain why it is or is not a reliable source.
- State how the paper is relevant for your own paper and research.

Additional contextual information about the research topic was intentionally withheld in order to test how the two models would generate critical links to previous literature and to determine whether their performance showed any improvements.

In total, 25 summaries were produced: 15 written by the researchers (five each) and 10 generated by ChatGPT (five by ChatGPT-4 and five by ChatGPT-4.5).

3.4 Summary Evaluation Survey Questions and Participants

Fourteen university professors teaching English linguistics at Belgrade Metropolitan University and the Faculty of Philosophy of the University of Niš participated in this study as expert raters. All participants teach applied linguistics courses, and were selected through convenience sampling based on their relevant teaching and research expertise. Each expert rated the overall quality of all 25 summaries on a five-point Likert scale (very poor, poor, acceptable, good, very good) and answered whether they thought each summary was AI-generated. To avoid priming effects, no explicit criteria for judging the quality of the summaries were provided, as raters were expected to rely on their professional judgment. An optional open-ended question followed each summary, allowing participants to elaborate on their ratings and their decisions regarding AI authorship.

4 Data Analysis and Results

4.1 Summarization Quality Analysis of ChatGPT-generated Summaries

The first stage of analysis focused on evaluating the summaries produced by ChatGPT-4 and ChatGPT-4.5 in terms of accuracy, faithfulness to the original texts, and adherence to the required structural format. This analysis was carried out manually by all three researchers, who systematically examined each summary against the source article and the prescribed annotated bibliography template.

ChatGPT-4

Upon first examination of the five summaries, it was noticed that the summary of Article 5 contained several obvious inaccuracies and hallucinations. The information in the summary appeared to have been merged with the content of a different study. In order to eliminate any possible mistakes made by the researchers when prompting ChatGPT, this output was discarded and regenerated in a new chat with temporary chat mode enabled. The prompt was not changed nor were the model settings altered in any way. The regenerated summary of Article 5 no longer contained hallucinations, suggesting the error resulted from context window confusion between the documents rather than an inherent model failure.

Closer inspection of the final set of summaries revealed that they did not contain any information that was not included in the articles themselves. In each summary, ChatGPT had successfully identified the topic of the paper, described the methodology accurately, and presented the results in a way that did not omit any important information. However, there was one discrepancy in how it identified the topic in its summary of Article 4.

Summary 4: *This paper explores how technology-enhanced learning (TEL) contributes to the improvement of English pronunciation and overall language proficiency, emphasizing the role of digital tools such as speech recognition software, mobile applications, and interactive learning platforms. It investigates how blending tradi-*

tional instruction with modern technology can help learners overcome pronunciation challenges, receive real-time feedback, and enhance self-directed language practice.

The research focus of the original article was on the use of digital tools to enhance pronunciation, fluency, and comprehension skills, while the results from interviews indicated that the advantages of these systems lay in their ability to provide instant feedback and provide guidance for further improvement in real time. The summary, on the other hand, presented these elements as research aims rather than reported outcomes. We regard this imprecision as a minor error that is unlikely to influence a researcher's decision to consult the full paper. Overall, the factual accuracy and faithfulness of the analyzed ChatGPT-4 summaries were found to be high, and the outputs can be considered reliable in assisting researchers in deciding whether to engage with the original texts further. These results are consistent with Zhang et al. [21], who reported that hallucinations and inaccuracies were greatly reduced when models were provided with structured, step-by-step instructions.

Regarding the evaluation of the articles' credibility, several patterns were observed. Four of the summaries (all except the fifth) cited publication in a peer-reviewed journal as evidence of credibility, while all five referred to the inclusion of detailed methodology or statistical analysis. Additional elements also appeared: the summary of Article 1 mentioned the study's limitations and the authors' affiliations, Summary 2 highlighted ethical considerations, and Summaries 2 and 3 pointed to the use of theoretical frameworks. While all of these factors are valid indicators of credibility, human-authored annotated bibliographies might not typically state them all explicitly. The findings indicate that ChatGPT-4 lacks the critical selectiveness of human researchers, however, it might produce different output if provided with more specific criteria focusing on article credibility in the prompt.

The final part of each summary contained one or two sentences that outline the relevance of the article. However, these sentences were found to be overly general and provided no specifics as to how these articles fit into the researcher's own work. This further illustrates ChatGPT-4's limited capacity for critical evaluation and corroborates earlier findings reported in the literature [6] [17].

In regard to whether ChatGPT-4 was able to follow the requested annotated bibliography structure as defined in the prompt, the results indicate that the model performed successfully. All five summaries followed the prescribed format, as illustrated in the table below.

Table 1. Structure of ChatGPT-4 Generated Summaries

	Total word count	% of words describing the topic, methodology, and findings	% of words providing an evaluation of credibility and own research usefulness
Article 1	299	65%	35%
Article 2	304	65%	35%
Article 3	299	79%	21%
Article 4	302	73%	27%
Article 5	219	66%	34%

ChatGPT-4.5

The same analytical procedure was applied to the summaries generated by ChatGPT-4.5. As with ChatGPT-4, the summaries did not include any information absent from the original articles. Each summary stated the topic faithfully, and both the methodology and results were described accurately, though the ChatGPT-4 summaries tended to provide more detail in these sections. The output of ChatGPT-4.5 were also comparatively shorter than those produced by ChatGPT-4.

With respect to the evaluation of the articles' credibility, patterns similar to those in ChatGPT-4 were observed. All five summaries referred to methodology and data analyses as evidence of credibility. Summaries 1, 3, and 5 additionally mentioned publication in a peer-reviewed journal, with Summary 1 explicitly naming the journal. Summary 5 further included information about the author and the article's DOI.

Summary 5: *Authored by a scholar with a solid publication record in language education, it appears in a peer-reviewed academic journal with a clear DOI.*

Additionally, Summaries 2 and 4 included sample size, Summaries 3 and 4 referenced theoretical frameworks, and Summaries 2 and 5 noted limitations as further evidence of credibility.

Comparing the two models, it can be concluded that both ChatGPT-4 and ChatGPT-4.5 consistently referred to aspects of methodology and data analyses as primary indicators of credibility, which aligns with accepted standards of academic reliability. Both models also frequently mentioned peer review status as evidence, even though this need not be stated explicitly when articles have already been selected from peer-reviewed journals. The same applies to references to author information, which is typically not noted in annotated bibliographies. Other forms of evidence offered by the models appeared somewhat random, though still reasonable and acceptable. These findings suggest that ChatGPT-4.5 does not demonstrate greater critical selectiveness than ChatGPT-4.

Finally, the last section of the ChatGPT-4.5 summaries addressed the articles' relevance to the personal study. As with ChatGPT-4, these sections consisted of one or two sentences making general links between the findings and the study at hand, showing no notable improvement in this respect over the earlier model.

As with ChatGPT-4, all summaries generated by ChatGPT-4.5 adhered to the requested structure, as shown in Table 2, indicating that both models are consistent in following structural instructions when clearly specified in the prompt.

Table 2. Structure of ChatGPT-4.5 Generated Summaries

	Total word count	% of words describing the topic, methodology, and findings	% of words providing an evaluation of credibility and own research usefulness
Article 1	218	61%	39%
Article 2	230	53%	47%
Article 3	233	70%	30%
Article 4	258	71%	29%
Article 5	220	60%	40%

The distribution of content across topic, methodology, findings, and evaluation was relatively similar in both models, with ChatGPT-4 devoting 65–79% of words to description and ChatGPT-4.5 allocating 53–71%. Given the questionable information occasionally used as credibility evidence and the generally vague statements regarding article relevance, ChatGPT-4 appears to be marginally more useful than ChatGPT-4.5 for generating annotated bibliographies.

Taken together, these observations directly address the first research question of this study, indicating that both models can produce accurate and structurally consistent summaries, with ChatGPT-4 showing marginally greater usefulness than ChatGPT-4.5.

4.2 Survey Results

Quantitative Analysis Results

Table 3 presents the quantitative results of the survey. All 14 participants provided ratings and AI-authorship judgments for each of the 25 summaries.

Table 3. Survey Results

Author	Summary	Rating			Perceived AI			
		Mean	SD	Count of *yes* answers	Count of *no* answers	Count of both *yes* and *no* answers	Mean	SD
Human 1	1	3.57	0.94	2	12	1	0.18	0.37
	2	3.79	0.70	2	12	0	0.14	0.36
	3	3.79	0.58	2	11	1	0.18	0.37
	4	4.07	0.62	2	11	1	0.18	0.37
	5	4.21	0.36	4	8	2	0.36	0.46
Human 2	1	4.21	0.89	9	5	0	0.64	0.50
	2	4.00	0.68	6	8	0	0.43	0.51
	3	4.07	0.62	10	3	1	0.75	0.43
	4	3.93	0.73	9	5	0	0.64	0.50
	5	4.14	0.53	7	5	2	0.57	0.47
Human 3	1	3.29	0.73	1	12	1	0.11	0.29
	2	3.93	0.92	0	14	0	0.00	0.00
	3	2.86	0.77	4	9	1	0.33	0.46
	4	3.86	0.95	3	11	0	0.21	0.43
	5	3.86	0.77	1	13	0	0.07	0.27
ChatGPT-4	1	4.00	1.04	11	2	1	0.82	0.37
	2	4.36	0.84	10	4	0	0.71	0.47
	3	4.00	1.04	9	4	1	0.68	0.46
	4	3.93	1.21	8	4	2	0.64	0.46
	5	3.86	0.95	12	2	0	0.86	0.36
ChatGPT-	1	3.86	0.77	8	6	0	0.57	0.51

4.5	2	4.14	0.86	8	6	0	0.57	0.51
	3	3.93	0.92	11	2	1	0.82	0.37
	4	4.21	0.89	6	7	1	0.46	0.50
	5	3.71	0.64	9	5	0	0.64	0.50

Since all the participants were asked to evaluate the same summaries, the degree of agreement or consistency among them was first calculated using the Interclass Correlation Coefficient (ICC). The result was ICC(2.k)=0.39 indicating poor-to-fair consistency among the 14 experts. This suggests that a considerable portion of the variability in ratings was attributable to differences among raters rather than true differences in the quality of the summaries. To further explore rater variability, mean ratings and standard deviations were calculated for each rater, as shown in Figure 1.

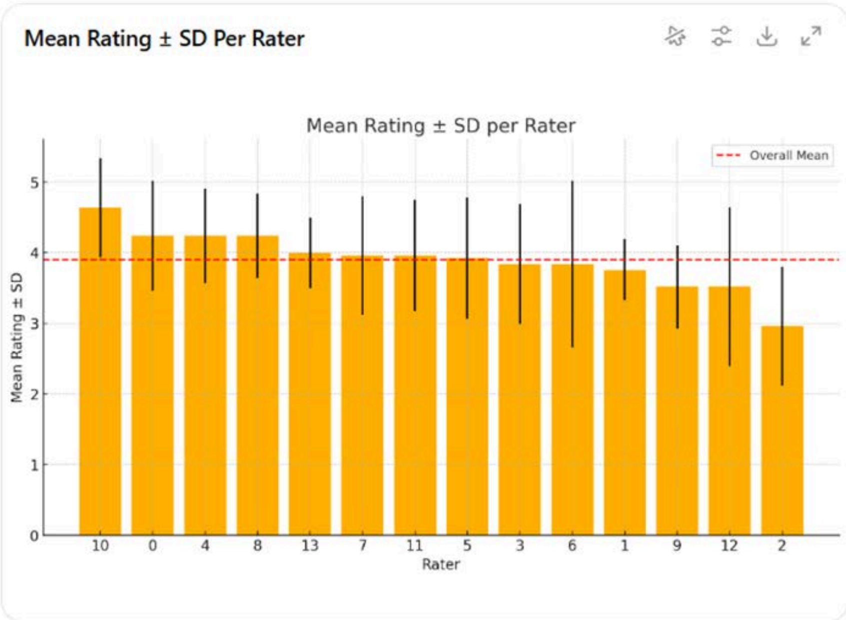


Fig. 1. Mean Rating and Standard Deviation per Rater.

It was found that Participant 10 was consistently more lenient (M = 4.64), while Participant 2 was systematically stricter (M = 2.96). Additionally, Participants 6 and 12 demonstrated the highest rating variability (SD > 1.1). A one-way ANOVA confirmed significant differences across raters ($p < 0.001$), and Tukey's HSD post-hoc tests revealed that multiple rater pairs differed significantly in their mean ratings. These results indicate that participants applied different internal criteria for assessing the quality of the summaries, and that the overall ratings cannot be considered highly reliable as an objective measure of summary quality.

To address this issue, analytical approaches were employed to reduce the influence of rater bias and obtain a more stable estimate of central tendency (i.e. what most raters thought on average even if they disagreed on the details). The data was reanalyzed using median ratings per summary, which revealed that ChatGPT-4 (Mdn = 4.0) and ChatGPT-4.5 (Mdn = 4.0) summaries received comparable evaluations, while Human2 summaries tended to receive slightly higher ratings (Mdn = 4.1) and Human1 and Human3 summaries slightly lower ratings (Mdn = 3.8). Grouping summaries into three categories (Human, ChatGPT-4, ChatGPT-4.5) showed no significant differences. ChatGPT-4 and ChatGPT-4.5 summaries both received median ratings of 4.0, while Human summaries received a median rating of 3.9. Although Human summaries displayed greater variability (SD = 0.39) than ChatGPT-4 and ChatGPT-4.5 (SD = 0.00 for both), the mixed-effects model revealed that these differences were not statistically significant. These findings suggest that ChatGPT-4 and ChatGPT-4.5 produced summaries that were evaluated as being of comparable quality to human-authored summaries, with no significant performance differences between the two models. However, the low inter-rater agreement raises questions about the subjectivity of expert evaluations.

In addition to rating quality, experts were also asked to judge whether each summary was generated by AI, and to offer additional insights into how they differentiated between human and AI-authored texts. Their overall detection accuracy was calculated using a majority vote criterion (≥50% of raters identifying a summary as AI). The result was 56%, suggesting that experts performed only slightly better than chance (50%). Detection accuracy differed by true authorship, whereby experts correctly identified 70% of AI-authored summaries but only 47% of human-authored summaries. This indicates that participants were substantially more accurate at detecting AI-generated summaries than at correctly identifying human-written ones, suggesting the ChatGPT outputs displayed identifiable characteristics.

Finally, a linear regression analysis was conducted to examine whether perceived AI authorship influenced quality ratings. The results revealed a significant effect of perception, whereby summaries perceived as AI-generated received ratings that were, on average, 0.32 points lower than those perceived as human (or uncertain) ($p < 0.001$). This suggests that bias associated with perceived AI authorship negatively impacted ratings regardless of the true authorship of the summary, and may have contributed to the low inter-rater agreement among the participants.

Qualitative Analysis Results

Responses to the open-ended survey questions were analyzed using thematic analysis following Braun and Clarke's [30] framework. An inductive, semantic coding approach was adopted, meaning that codes were generated directly from participants' responses without imposing pre-existing categories, and the analysis focused on the explicit content of the data. Codes were then iteratively grouped into broader themes through comparison and refinement. To structure the analysis and enhance interpretability, themes were organized around four analytical cases: (1) correct classification of human-written texts, (2) misclassification of human-written texts, (3) correct classification of ChatGPT-authored texts, and (4) misclassification of ChatGPT-authored texts. This approach provided a systematic way to capture both the criteria raters used in their judgments and the implications of those judgments for evaluating ChatGPT's effectiveness as a research assistant.

Cases When Human-written Texts Were Correctly Classified

Within the first case, several themes emerged that reflect the criteria raters relied on to distinguish human from AI authorship: information density, word choice and style, structure and organization, credibility and relevance, and the presence of errors.

Information density. Most responses (n = 9) in this category focused on the amount of information presented. Six participants highlighted omissions in methodology or results as indicators of human authorship: "*Not very detailed (no mention of Likert scales, etc.)*" (ART2 HUM3 P7); "*The results and conclusions… are not precise and comprehensive enough*" (ART3 HUM3 P11). Three participants, however, regarded conciseness as a strength: "*Includes appropriate specifics without information overload*" (ART5 HUM3 P13). Overall, participants tended to expect ChatGPT to provide extensive descriptive detail, whereas human authors were expected to exercise greater critical selectiveness. This indicates that ChatGPT is effective for descriptive summarization but less aligned with academic expectations of concise and evaluative writing, which aligns with our findings from the qualitative analysis of the ChatGPT-generated summaries.

Word choice and style. Informal expressions, hedges, and clumsy phrasing were frequently interpreted as markers of human authorship. For example: "*The report hedges at the end with 'The results of this study appear to be credible,' which I think an AI would avoid*" (ART1 HUM1 P8). Language errors further reinforced this perception: "*Also, there are some language mistakes, i.e. wrong choices*" (ART1 HUM3 P4). By contrast, ChatGPT's strengths lie in grammatical accuracy, consistent formality, and the correct use of disciplinary terminology, but its lack of hedging and imperfection may undermine perceptions of authenticity.

Structure and organization. Human-written texts were often perceived as disorganized: "*The summary is very chaotic, not clear(ly organized)*" (ART1 HUM1 P7). In contrast, ChatGPT's rigid adherence to structural conventions can be advantageous for clarity and for facilitating the accurate interpretation of information, though this same rigidity may reduce the stylistic naturalness typically associated with human writing.

Credibility and relevance. Engagement with personal research was a strong indicator of human authorship: "*The conclusion demonstrates actual engagement with how this relates to the writer's own research*" (ART5 HUM1 P13). ChatGPT, by contrast, lacks the ability to establish genuine connections between articles and a researcher's individual project, underscoring a core limitation of the tool as a research assistant.

Errors. Spelling and grammatical mistakes were typically attributed to human authorship: "*Contains small mistakes like 'inlcuded' instead of 'included'*" (ART4 HUM3 P13). ChatGPT, by contrast, largely eliminates such errors, which can enhance efficiency in research-related tasks. However, this very perfection paradoxically makes its outputs more easily identifiable as AI-generated.

Cases When Human-written Texts Were Misclassified

In the second case, themes similar to those identified in the first case re-emerged, particularly regarding word choice, information density, relevance, and errors. This highlights the overlapping criteria participants used and the difficulties of reliably distinguishing between human and AI summaries.

Word choice. Repetitive phrasing and vague expressions led participants to attribute some human-authored summaries to AI: "*Great summary, but wording such as 'well-grounded' and 'valuable insights' are vague… sounding professional*" (ART3 HUM2 P12). This illustrates that stylistic overlap exists, as humans occasionally produce "AI-like" phrasing. For ChatGPT as a research assistant, this suggests that its outputs can blend into established academic conventions, but at the same time may reinforce perceptions of formulaic or overly generic language.

Information density and relevance. A lack of detail regarding how an article related to personal research was often taken as a sign of AI authorship: "*It is impersonal and shows that AI can summarize well but does not relate anything to personal interest*" (ART1 HUM2 P3). When human authors omitted such evaluative context, their summaries were sometimes misclassified as AI-generated. This highlights that while ChatGPT cannot yet convincingly perform evaluative tasks, human writers may also neglect this dimension, contributing to overlap and misclassification.

Errors. Factual mistakes and grammatical slips were at times attributed to AI: "*The summary… combination of generic phrasing and minor grammatical inconsistencies suggests AI*" (ART1 HUM1 P13). This reflects a bias in which participants expected AI to produce errors resembling those made by humans, illustrating how perceptions of AI unreliability can distort evaluations and lead to misclassification.

Cases When ChatGPT-authored Texts Were Correctly Classified

In the third case, participants highlighted themes similar to those observed in the first and second cases, particularly word choice, structure, credibility, and errors, though here these features were more consistently associated with AI authorship.

Word choice and style. Six participants described AI language as "mechanical" or characterized by "typical AI phrasing." For example: "*This one appears to be fully generated by AI. It sounds much more mechanical*" (ART1 ChatGPT-4 P10). Others observed: "*The text is repetitive*" (ART1 ChatGPT-4.5 P12). Word choice was the most frequently cited criterion across all cases, indicating that while ChatGPT ensures consistency and correctness, its predictability and limited lexical variety remain its most prominent stylistic weaknesses.

Structure and information density. ChatGPT was frequently recognized for its rigid organization: "*Extremely systematic progression through all elements*" (ART2 ChatGPT-4 P13). It was also described as detailed and clear: "*Very well organized, detailed and clear*" (ART1 ChatGPT-4 P7). This consistency represents a strength, making ChatGPT reliable for descriptive tasks. However, its rigidity also reduces naturalness, reinforcing perceptions of mechanical style.

Credibility. Formulaic credibility sections were frequently identified as AI markers: "*The final paragraph about credibility reads like a standard AI template*" (ART4 ChatGPT-4 P13). While ChatGPT can reliably insert credibility indicators, they often lack nuance and may appear artificial, limiting its usefulness in evaluative dimensions of research assistance. This observation highlights and reinforces the findings reported earlier, where both models demonstrated limited critical selectiveness in assessing credibility.

Syntactic complexity and punctuation. Participants frequently noted the simplicity of sentence patterns and the presence of “AI-like” punctuation: “*Sentences are mostly simple and to the point… typical AI punctuation*” (ART5 ChatGPT-4 P12). These perceptions indicate that while ChatGPT’s preference for clarity and formula supports readability, it also reduces stylistic authenticity and makes its outputs more readily identifiable as AI-generated.

Cases When ChatGPT-authored Texts Were Misclassified

In the fourth case, the only recurring theme concerned credibility and relevance. The main reason AI-generated summaries were mistaken for human was when they included statements about personal research: “*This one has a reference to ‘my research’*” (ART3 ChatGPT-4.5 P11). This shows that ChatGPT can mimic evaluative engagement, but such personalization remains surface-level rather than genuine.

Additional insight emerged from participants’ comments, indicating that personal attitudes toward AI influenced their evaluations. One participant explicitly acknowledged: “*My obviously negative attitudes towards the use of AI… influence my evaluation of abstract quality*.” This demonstrates that predispositions toward AI can bias perceptions, undermining the fairness of comparative evaluations. For ChatGPT as a research assistant, this highlights that user trust is as critical a factor as textual quality in determining its acceptance and effectiveness.

5 Discussion

The purpose of this study was to evaluate the performance of ChatGPT-4 and ChatGPT-4.5 when summarizing applied linguistics (AL) research papers and to examine if these models could function as reliable research assistants. The study addressed two research questions: (1) To what extent do ChatGPT-4 and ChatGPT-4.5 produce accurate summaries that capture key findings while adhering to the required structural conventions? and (2) In what respects do ChatGPT-generated summaries differ from human-authored summaries?

With respect to Research Question 1, both ChatGPT-4 and ChatGPT-4.5 produced summaries that were factually accurate and faithful to the original texts, with no hallucinations when the articles were uploaded individually and the models were guided with explicit, step-by-step prompts. This suggests that the newer versions of ChatGPT have improved in terms of accuracy and supports earlier observations that procedural prompting can reduce hallucinations [21] [22]. ChatGPT-4 tended to provide more detailed accounts of methodology and results, while ChatGPT-4.5 generated shorter and more concise outputs. Both models consistently followed the annotated bibliography structure, demonstrating their reliability for descriptive summarization and structural conformity. However, they showed limited critical selectiveness, often relying on surface-level credibility markers and offering only generalized statements of relevance, confirming limitations highlighted in prior studies [6] [17].

Turning to Research Question 2, the findings indicate that ChatGPT-generated summaries were rated as being of comparable quality to human-authored ones, with no significant performance differences observed across the groups. This suggests that, in terms of overall perceived quality, ChatGPT-4 and ChatGPT-4.5 can already perform at a level similar to human researchers when tasked with descriptive summarization in applied linguistics. This aligns with Hake et al.'s [5] study in medicine, where experts judged ChatGPT-3.5 summaries as high in accuracy and low in bias. However, the low inter-rater agreement highlights the subjective nature of expert evaluations, suggesting that raters relied on diverse criteria and that judgments of summary quality are not straightforward. Thematic analysis helped clarify these underlying criteria. Participants often associated information density, repetitive or mechanical phrasing, rigid structure, and formulaic credibility statements with ChatGPT, while hedging, stylistic variety, evaluative engagement, and occasional errors were taken as indicators of human authorship. This reflects broader concerns raised in applied linguistics and corpus studies that ChatGPT outputs risk appearing formulaic or lacking nuance [8].

The detection task further clarified these distinctions. Experts were considerably better at recognizing AI-authored summaries (70%) than human-authored ones (47%), indicating that ChatGPT output retained identifiable stylistic features that distinguished them from human writing. While this ability to detect AI-authored texts may safeguard against over-reliance on ChatGPT outputs, it also suggests that the models are not yet fully capable of blending seamlessly with human writing styles in academic contexts. Importantly, regression analysis revealed that perceived AI authorship negatively influenced ratings. This bias may be an impediment toward using ChatGPT as a research assistant, echoing concerns in acceptance studies that user perceptions and attitudes significantly shape the adoption of AI tools [15] [16].

The advantages and disadvantages of ChatGPT, as revealed in the findings, become especially clear when compared with human-authored summaries. As shown in Table 4, ChatGPT demonstrates notable strengths in accuracy, structural consistency, grammatical correctness, and efficiency, making it particularly effective for descriptive tasks and the initial stages of literature review. At the same time, its limitations, most notably its inability to engage critically with the material, restrict ChatGPT's usefulness for tasks that demand selectiveness and nuanced interpretation. These results align with prior studies emphasizing ChatGPT's potential as a time-saving tool [31], while confirming its insufficiency for tasks requiring deeper evaluation and theoretical integration [6] [17].

Table 4. Advantages and Disadvantages of using ChatGPT in AL Research Summarization

Dimension	ChatGPT (4 / 4.5)	Human Authors
Accuracy & Faithfulness	Generally accurate and faithful to source texts, few hallucinations when prompted properly.	May omit details or include errors due to oversight or subjectivity.

Dimension	ChatGPT (4 / 4.5)	Human Authors
Structure & Organization	Consistently follows annotated bibliography format. Is rigid and systematic.	Structure sometimes inconsistent or disorganized, but allows flexibility and nuance.
Information Density	Provides comprehensive descriptive detail, reliable for factual coverage.	More selective and concise, exercising critical judgment in deciding what to include.
Word Choice & Style	Grammatically correct, formal, and terminologically precise, but mechanical, repetitive, and formulaic.	Varied, natural, and authentic, but prone to hedging, vagueness, and clumsy phrasing.
Credibility Evaluation	Reliably cites surface-level markers (peer review, methodology, author info). Formulaic and lacking nuance.	Offers more critical selectiveness and nuanced evaluation of credibility.
Relevance & Engagement	General, impersonal statements. Cannot genuinely connect findings to personal research.	Can link articles to own research goals and contexts, showing authentic engagement.
Errors	Few to none (spelling, grammar, factual slips rare). Perfection can make outputs identifiable as AI.	Occasional spelling, grammar, or factual mistakes. Paradoxically perceived as more "human."
Expert Evaluation	Rated as comparable in quality to human summaries (Mdn = 4.0). Identifiable stylistic markers make AI easier to detect. Bias against perceived AI lowered ratings.	Rated similarly (Mdn = 3.9), but more variable in quality. Harder to detect as human when vague or impersonal.
Overall Usefulness	Efficient for descriptive summarization, initial scanning, and standardized outputs.	Stronger in evaluative depth, contextualization, and authentic voice.

6 Conclusion

This study investigated the extent to which ChatGPT-4 and ChatGPT-4.5 can serve as research assistants by generating annotated bibliographies of research articles related to applied linguistics (AL). The findings showed that both models are highly reliable in terms of factual accuracy and structural adherence, as they produced summaries that were comparable in quality to those written by human experts. Their ability to faithfully summarize topics, methodologies, and findings demonstrates they hold potential for tasks that require descriptive summarization and standardized formats.

At the same time, both models exhibited limitations when providing critical evaluation. While they consistently highlighted peer-review status, methodological detail, or author credentials as markers of credibility, the evaluations were formulaic and lacked critical selectiveness. Similarly, statements about the studies' relevance were typically general and impersonal, and failed to demonstrate genuine engagement with research contexts. These weaknesses suggest that even though ChatGPT can assist in providing accurate summaries of the methods and results of AL research articles, it cannot engage critically with them nor provide nuanced reasoning as to why they might be relevant for the researchers' own work. Therefore, human expertise remains invaluable for theory-driven interpretation.

7 Limitations and Further Research

Several limitations of the present study should be acknowledged. First, the dataset consisted of only five articles per model, which does not capture the full range of ChatGPT's performance across different types of articles or subfields of AL. Second, although 14 experts participated in the evaluation, inter-rater reliability was low, which reflects the subjective nature of expert judgments and the influence of biases toward or against AI authorship. This limits the generalizability of the quantitative findings and highlights the need for further research into the criteria experts use when assessing summary quality and identifying AI authorship. Third, the study only tested the summarization of annotated bibliographies. Thus, testing how ChatGPT summarizes other formats such as structured abstracts or critical reviews may reveal different strengths and weaknesses. Fourth, prompting strategies were standardized but not systematically varied, even though prior research suggests that prompting strongly influences ChatGPT's performance [21] [22]. Finally, the analysis was restricted to ChatGPT-4 and ChatGPT-4.5, and other iterations of LLMs may address some of the limitations observed here.

Future research should expand the dataset to include a larger and more diverse corpus of AL research articles and experiment with multiple summary formats. Studies should also systematically vary prompting strategies to assess whether evaluative depth and critical selectiveness can be improved. Further research should also explicitly investigate the criteria on which experts base their perceptions of AI-generated versus human-authored texts, as understanding these criteria would help clarify sources of bias and improve both model development and evaluation practices. Finally, comparative studies across disciplines would help determine whether the findings observed in applied linguistics generalize to other fields, thereby clarifying the broader potential and limitations of ChatGPT in academic research.

Disclosure of Interests.

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., & Ramesh, S. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. npj *Digital Medicine*, *6*(1), 75. doi:10.1038/s41746-023-00819-6
2. O'Connor, S., & ChatGPT. (2023, January). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice, 66*, 103537. doi:10.1016/j.nepr.2022.103537
3. OpenAI. (2023). GPT-4 technical report. https://openai.com/research/gpt-4
4. OpenAI (2025). Introducing GPT-4.5. https://openai.com/index/introducing-gpt-4-5/
5. Hake, J., Crowley, M., Coy, A., Shanks, D., Eoff, A., Kirmer-Voss, K., Dhanda, G., & Parente, D. J. (2024). Quality, accuracy, and bias in CHATGPT-based summarization of Medical Abstracts. *The Annals of Family Medicine, 22*(2), 113–120. https://doi.org/10.1370/afm.3075
6. Salleh, H. M. (2023). Errors of commission and omission in artificial intelligence: Contextual biases and voids of chatgpt as a research assistant. *Digital Economy and Sustainable Development, 1*(1). https://doi.org/10.1007/s44265-023-00015-0
7. Bae, H. (2023). *CHATGPT as a Research Assistant in Experimental Linguistics*. https://doi.org/10.2139/ssrn.4585546
8. Uchida, S. (2024). Using early LLMS for corpus linguistics: Examining chatgpt's potential and limitations. *Applied Corpus Linguistics, 4*(1), 100089. https://doi.org/10.1016/j.acorp.2024.100089
9. Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in CHATGPT: Implications in scientific writing. *Cureus*. https://doi.org/10.7759/cureus.35179
10. Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using CHATGPT in academic writing is (not) a form of plagiarism: What does the literature say? *Online Journal of Communication and Media Technologies, 13*(4). https://doi.org/10.30935/ojcmt/13572
11. Meyer, J. G., Urbanowicz, R. J., Martin, P. C., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). CHATGPT and large language models in academia: Opportunities and challenges. *BioData Mining, 16*(1). https://doi.org/10.1186/s13040-023-00339-9
12. Rahardyan, T. M., Susilo, C. H., Iswara, A. M., & Hartono, M. L. (2024). CHATGPT: The future research assistant or an academic fraud? [A case study on a state university located in Jakarta, Indonesia]. *Asia Pacific Fraud Journal, 9*(2), 275–293. https://doi.org/10.21532/apfjournal.v9i2.347
13. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of Information Technology. *MIS Quarterly, 13*(3), 319. https://doi.org/10.2307/249008
14. Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425. https://doi.org/10.2307/30036540
15. Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended Tam based acceptance of ai-powered chatgpt for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon, 10*(8). https://doi.org/10.1016/j.heliyon.2024.e29317
16. Balaskas, S., Tsiantos, V., Chatzifotiou, S., & Rigou, M. (2025). Determinants of CHATGPT adoption intention in higher education: Expanding on Tam with the mediating roles of trust and risk. *Information, 16*(2), 82. https://doi.org/10.3390/info16020082
17. Rahman, M., Terano, H., Rahman, N., Salamzadeh, A., & Rahaman, S. (2023). CHATGPT and academic research: A review and recommendations based on practical ex-

amples. *Journal of Education, Management and Development Studies, 3*(1), 1–12. https://doi.org/10.52631/jemds.v3i1.175
18. Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out* (pp. 74–81). Association for Computational Linguistics.
19. Goyal. T., Li, J. J., & Durrett, G. (2022). *News summarization and evaluation in the era of gpt-3*. (arXiv:2209.12356). https://arxiv.org/abs/2209.12356
20. Yang, X., Li, Y., Zhang, X., Chen, H., & Cheng, W. (2023). *Exploring the limits of ChatGPT for query or aspect-based text summarization* (arXiv:2302.08081). arXiv. https://arxiv.org/abs/2302.08081
21. Zhang, H., Liu, X., & Zhang, J. (2023). *Extractive summarization via ChatGPT for faithful summary generation* (arXiv:2304.04193). arXiv. https://arxiv.org/abs/2304.04193
22. Ma, Y., Liu, J., Yi, F., Cheng, Q., Huang, Y., Lu, W., & Liu, X. (2023). *AI vs. human – Differentiation analysis of scientific content generation*. arXiv. https://doi.org/10.48550/arXiv.2301.10416
23. Fabbri, A. R., Kryściński, W., McCann, B., Xiong, C., Socher, R., & Radev, D. (2021). Summeval: Re-evaluating summarization evaluation. *Transactions of the Association for Computational Linguistics, 9,* 391–409. https://doi.org/10.1162/tacl_a_00373
24. Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press. https://doi.org/10.1017/9781108686136
25. Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022). 'Look, I can speak correctly': Learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning, 37*(5–6), 1335–1363. https://doi.org/10.1080/09588221.2022.2080230
26. Dennis, N. K. (2024). Using AI-powered speech recognition technology to improve English pronunciation and speaking skills. *IAFOR Journal of Education: Technology in Education, 12*(2), 107–126. https://doi.org/10.1016/j.caeai.2024.100230
27. Hsu, H.-W. (2024). An examination of automatic speech recognition (ASR)-based computer-assisted pronunciation training (CAPT) for less-proficient EFL students using the Technology Acceptance Model. *International Journal of Technology in Education, 7*(3), 456–473. https://doi.org/10.46328/ijte.681
28. Lyu, J., & Andi, H. K. (2024). The role of technology-enhanced learning in improving English pronunciation and language proficiency. *Sciences of Conservation and Archaeology, 36*(4), 96–102. https://doi.org/10.48141/sci-arch-36.4.24.9
29. Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. LLT Journal: *A Journal on Language and Language Teaching, 27*(1), 98–115. https://doi.org/10.24071/llt.v27i1.8151
30. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
31. Patel, S. B., & Lam, K. (2023). Chatgpt: The future of discharge summaries? *The Lancet Digital Health, 5*(3). https://doi.org/10.1016/s2589-7500(23)00021-3

Reusable Unity-based Puzzles for the Escapp Educational Escape Room Platform

Maja Cosic[1[/]], Sonsoles López-Pernas[2[0000-0002-9621-1392]], Aldo Gordillo[3[0000-0001-9785-4827]], Enrique Barra[3[0000-0001-9532-8962]], Alexandra Santamaría-Urbieta[4[0000-0003-0935-0616]], Milos Kostic[1[0009-0005-0912-9518]], Miljan Milosevic[1,5[0000-0003-3789-2404]]

[1] Belgrade Metropolitan University, 11000 Belgrade, Serbia
[2] University of Eastern Finland, Joensuu 80130, Finland
[3] Universidad Politécnica de Madrid, 28040 Madrid, Spain,
[4] Universidad Internacional de La Rioja, 26006 Logroño, La Rioja, Spain
[5] Institute for Information Technologies, 34000 Kragujevac, Serbia
miljan.milosevic@metropolitan.ac.rs

Abstract. Educators have shown growing interest in escape rooms, as they engage students while fostering teamwork, leadership, creativity, and communication. Consequently, educational escape rooms are emerging as a novel form of learning activity that seeks to enhance students' learning through immersive and highly engaging experiences. This paper presents a novel approach to designing reusable educational puzzles within Unity and integrating them seamlessly into the Escapp platform for creating educational escape rooms, as part of the IGLUE project. We describe the workflow that enables educators to configure customizable puzzle instances via Escapp's interface—each instance accessible through a unique URL that exposes configuration data in JSON format. Four distinct Unity puzzles have been developed according to a flexible, parameter-driven architecture, enabling them to adapt dynamically based on educator-defined configurations. During runtime, the Unity game fetches the JSON data to instantiate and render the puzzle, while solution validation occurs via secure API communication with Escapp—puzzle attempts are sent back to the platform for validation, and correct solutions trigger automatic notifications of completion. This bidirectional integration supports modularity, reusability, and real-time feedback and learning analytics, enhancing the design of escape-room-style learning scenarios in a scalable and extensible way.

Keywords: educational escape rooms, Escapp platform, reusable puzzles, game-based learning.

1 Introduction

In recent years, game-based learning has gained increasing attention as an effective approach to enhance learner engagement and motivation in digital education. Among the many forms of game-based learning, educational escape rooms have emerged as a particularly promising method, as they combine immersive storytelling with problem-solving activities that require creativity, critical thinking, and collaboration [1]. Educational escape rooms situate learners in scenarios where they must solve

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.225C
Part of ISBN: 978-86-89755-37-4

challenges to progress and offer opportunities to reinforce knowledge acquisition while simultaneously fostering transversal skills such as teamwork and communication [2].

Though leisure escape rooms —those designed primarily for entertainment— have been mostly physical, the digitalization of the concept has contributed to the adoption of educational escape rooms, since they can be applied in both classroom and remote learning contexts, provide multiple levels of difficulty, and allow for easy reusability and scalability across different subject domains [2].

However, most existing digital educational escape rooms rely on static designs, which limits flexibility for educators and the possibility to adapt them to the peculiarities of their classroom [3]. The IGLUE project addresses this challenge through the Escapp platform [3], a web-based environment that allows educators to create and orchestrate escape rooms, configure puzzle parameters, and track student progress (https://escapp.es/). The contribution of the present study focuses on developing reusable Unity-based puzzles integrated within Escapp. Each puzzle is dynamically configured via JSON files generated by the platform and validated through API communication, ensuring adaptability and security. Four Unity puzzles have been implemented as proof-of-concept, demonstrating how reusable design can enhance flexibility and scalability of educational escape rooms.

The aim of this paper is to present our work on the design and integration of reusable Unity puzzles into the Escapp platform. We describe the educational and technical background of escape rooms, introduce the Escapp platform, detail the architecture for puzzle configuration and solution validation, and present four specific puzzle implementations. Finally, we discuss lessons learned and potential directions for future development of reusable and scalable digital escape rooms.

2 Background

2.1 Educational escape rooms

Educational escape rooms are a type of educational game that combines problem-solving, collaboration, and critical thinking within a narrative framework [1]. Unlike traditional escape rooms, which focus purely on entertainment, educational escape rooms are designed to meet specific pedagogical objectives [2]. Students are placed in immersive scenarios where they must solve puzzles and challenges in order to progress, thus reinforcing learning outcomes in a playful yet structured manner. Key benefits identified in the literature include increased learner engagement, development of transversal skills such as communication, collaboration, creativity, and critical thinking [4], [5], [6], [7], as well as the possibility of integrating domain-specific content in areas such as healthcare [8], or STEM [9]. Digital escape rooms further expand these possibilities by allowing both classroom and remote participation, adaptability to different difficulty levels, and the reuse of puzzle templates across multiple educational contexts [10].

A growing body of reviews illustrates the breadth of these applications. Early surveys highlighted escape rooms as promising tools to enhance student motivation and engagement across disciplines [11], while later analyses examined design principles linking puzzles, narratives, and learning goals [1]. Research has also emphasized the importance of curricular integration and 21st-century skill development, though with limited theoretical grounding [12], a gap that has been noted in recent scholarship [13]. Empirical syntheses provide further evidence: meta-analyses report that educational escape rooms are generally effective across levels of education and modes of delivery [2], while domain-specific reviews in STEM [9] and medical education [14] indicate both motivational benefits and measurable knowledge gains.

At the same time, educational escape rooms are challenging to design, create, and conduct. Developing pedagogically sound activities requires instructors to construct an engaging narrative alongside a sequence of puzzles that are logically connected, employ meaningful game mechanics, and target the desired learning outcomes [15]. Such design processes demand imagination, creativity, and a strong pedagogical orientation. Beyond the conceptual stage, the practical implementation of puzzles and interactive elements can be difficult and time-consuming, particularly for instructors who lack technical expertise. This tension between the pedagogical potential of escape rooms and the practical challenges of their design motivates the need for platforms that lower barriers to creation and provide reusable puzzle templates.

2.2 The Escapp platform

Escapp is a web-based platform specifically developed to support the creation, orchestration, and execution of educational escape rooms [3]. It enables educators to design escape room scenarios without requiring advanced technical skills by offering a catalog of configurable puzzles, a scenario editor, and tools for monitoring student progress. One of Escapp's strengths is its modular architecture, which allows external applications—such as Unity-based puzzles—to be seamlessly integrated into escape room narratives. Each puzzle can be parameterized through a JSON configuration generated within Escapp's interface, ensuring flexibility and reusability. In addition, Escapp provides APIs for solution verification and event notifications, making it possible to synchronize game logic between the Unity puzzles and the platform. The tight coupling between customization, gameplay, and progress tracking makes Escapp a powerful tool for educators who aim to leverage game-based learning in both face-to-face and online environments (see Fig. 1).

Fig. 1. Screenshot of the Escapp platform team interface.

2.3 The IGLUE project

Educational escape rooms often face limitations such as the static nature of puzzles, restricted opportunities for reuse, and the high technical effort required for customization. Addressing this limitation requires a framework that supports the reusability and configurability of puzzles. This is the core idea behind the IGLUE project, in which the Escapp platform has been developed as a web-based environment to design, orchestrate, and manage educational escape rooms. Escapp enables educators to create customized puzzles and scenarios without advanced technical knowledge, while also providing APIs for solution verification and player progress tracking.

Within this project, our contribution focuses on the development of reusable Unity-based puzzles that are directly integrated with the Escapp platform. Each puzzle is configured dynamically through a JSON file generated by Escapp and validated via API communication. In this way, puzzles can be easily adapted to different educational contexts while ensuring security and modularity. Four Unity puzzles have been designed and implemented as part of this approach, serving as concrete examples of how a reusable architecture can enhance the effectiveness of digital escape rooms.

3 Creating Reusable Unity-based Puzzles for Escapp

3.1 Reusable Puzzle Design in Unity

The development of reusable puzzles in Unity followed a modular and parameter-driven approach, ensuring that a single puzzle template could be adapted to multiple learning contexts. Instead of embedding fixed content or solutions directly into the Unity code, puzzles were designed to load their configuration at runtime. This was achieved through a JSON-based mechanism, where external parameters define the behavior and appearance of each puzzle instance.

The configuration file may contain elements such as the puzzle's textual content, difficulty level, number of interactive components, or hints to be displayed to learners. Once the Unity puzzle is launched, it retrieves the JSON file from a unique URL generated by the Escapp platform, parses the data, and instantiates the puzzle accordingly. This approach separates the pedagogical content, defined by the educator, from the technical implementation, designed by the developer.

3.2 Integration Between Unity Puzzles and Escapp

The integration between Unity puzzles and the Escapp platform is based on a two-way communication model. In the first direction, JSON → Unity, the puzzle reads the configuration file generated by Escapp. This ensures that when a student launches a puzzle within an escape room scenario, the Unity application automatically adapts to the parameters chosen by the educator. The configuration stage provides flexibility while maintaining a consistent workflow inside the platform.

In the opposite direction, Unity → Escapp, the interaction focuses on solution validation and progress tracking. When a learner attempts to solve a puzzle, Unity does not perform the validation internally. Instead, it sends the proposed solution to Escapp's API, which checks correctness and returns a boolean result. If the answer is valid, Unity notifies Escapp of puzzle completion, enabling the platform to register progress and orchestrate the sequence of the escape room. This mechanism—implemented through REST calls and potentially WebSocket communication—ensures security, prevents unauthorized access to solutions, and guarantees that the educator retains full control over the learning flow.

3.3 About reusable puzzles

In this section we present the reusable puzzles created for the purpose of the IGLUE project: Caesar cipher, Bomb, Radio and Safebox.

Caesar cipher puzzle. The Caesar cipher puzzle relies on a JSON configuration file that specifies two parameters: a string representing the original code and an integer that defines the shift value. Based on this input, the Unity application generates a cipher that players are required to decode.

```
{
  "code": "ABC 123",
  "shift": -2
}
```

Fig. 2. Code of Caesar cipher.

For example, if the code is ABC and the shift is +2, the encrypted output becomes CDE. Conversely, a negative shift produces a reverse transformation; with the same code ABC and a shift of –2, the resulting cipher is YZA.

The game interface is divided into two sections: the upper part displays the encoded text, while the lower part provides the input area for players' attempts, Fig. 3. Input is performed through a virtual keyboard positioned below the screen, with additional options to reset or correct entries via backspace. The system also allows users to select specific letter positions when constructing their solution. The visual design emulates a retro-style handheld device, reinforcing the puzzle's mysterious atmosphere. All graphical elements were hand-drawn to enhance usability and player immersion. The development process followed a two-step approach: an initial prototype was created to validate interaction and gameplay mechanics, which was later refined into the final version implemented in Unity.

Fig. 3. Caesar cipher puzzle made in Unity.

Bomb puzzle. The bomb puzzle is a 3D puzzle that consists of four distinct smaller puzzles, each contributing to the overall puzzle system, Fig. 4a. To successfully complete the puzzle the player needs to stop the bomb timer by pressing the stop button.

a b

Fig. 4. Bomb puzzle example: a) Overall puzzle system, b) Sliders puzzle.

The stop button is covered by a plastic lid, which can only be lifted after the player completes all the smaller puzzles. For the purpose of connecting the puzzle to Escapp, the smaller puzzles must be solved in a specific order. The four smaller puzzles in order are: the sliders puzzle, color button puzzle, wire cut puzzle and number pad puzzle. Each of the puzzles has a lightbulb that indicates if the puzzle is active. If the lightbulb shines yellow it means that the puzzle has been activated, but if the lightbulb is just gray it means that the puzzle is inactive. When the puzzle is completed correctly the lightbulb will turn green. However the player makes a mistake the lightbulb will blink red and reset.

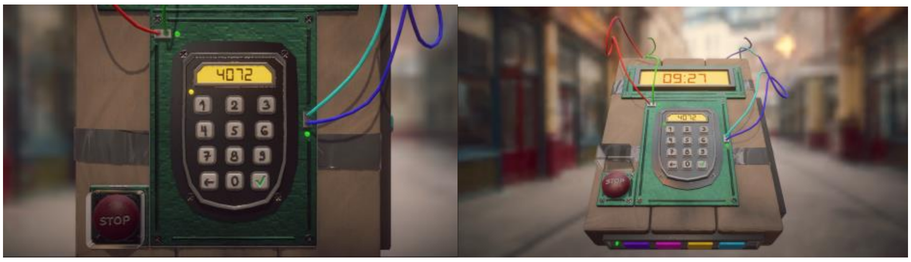
The first puzzle to be completed is the sliders puzzle, Fig. 4b. This mini puzzle features two separate sliders positioned on the sides of the bomb. The player must match each of their values to the corresponding values entered in Escapp. To change the value, the player needs to grab the slider's handle and drag it to the correct position.

a b

Fig. 5. Bomb puzzle example: a) Color button puzzle, b) Wire cut puzzle.

The next mini puzzle is the color button puzzle, Fig. 5a. This puzzle consists of four differently colored buttons that are placed at the bottom part of the bomb. The buttons themselves need to be pressed in a predefined sequence. The order of the buttons and their color are entered through Escapp. Furthermore, the button sequence can repeat the same button multiple times or exclude some buttons. The length of the sequence can also be changed to fit each escape room's purposes. If the player makes a mistake during the sequence, it will reset and the player will have to start from the beginning of the sequence.

The third mini puzzle is the wire cut puzzle, Fig. 5c. This puzzle consists of four wires that connect the front of the bomb to its top. Each wire has a specific color and must be cut in a predefined order, which can be entered through Escapp. Since wires are being cut, each wire can only appear once during the sequence, and the sequence can only range from 1 to 4. When the player hovers the mouse over a wire, the cursor changes to a wire cutter icon to simulate cutting. Once a wire is cut, it will visually appear cut in the game. If the player cuts a wire in the wrong order, the sequence resets, and they must start again from the beginning.



a b

Fig. 6. Bomb puzzle example: a) Number pad puzzle, b) Completed bomb puzzle.

The last mini puzzle is the number pad puzzle, Fig. 6a. This puzzle consists of a number pad with digits, a backspace, and an enter button, located at the center of the bomb. The player needs to enter the correct number code, which can be set through Escapp. Once the player has inputted the code, they must press enter to check if it is correct. If the code is incorrect, the player will need to reenter a new code.

Finally, once all of the mini puzzles are completed the plastic lid covering the button will lift up and the player can press the stop button to stop the timer, Fig. 6b. With that the player has successfully completed the whole puzzle.

Safebox puzzle. The safebox puzzle is a 3D puzzle that requires the player to enter a code using a combination lock, Fig. 7. The lock can be spun both clockwise and counterclockwise to the desired digit. Pressing the key Q on the keyboard turns the rotation counterclockwise, while pressing the key E turns the rotation clockwise. For easier use, the player is using a circular slider to turn the handle to the correct number. The code for the safebox is entered through Escapp, as well as the direction of the lock. Once the correct code has been entered, the player can click on the safebox handle to open it. If the player enters a wrong digit while inputting the code, the lock resets and the player must start over. Inside the safebox there is either a note or an image that provides an additional clue or story element to the game. The content inside can be added through Escapp, including whether it should be a note or an image.

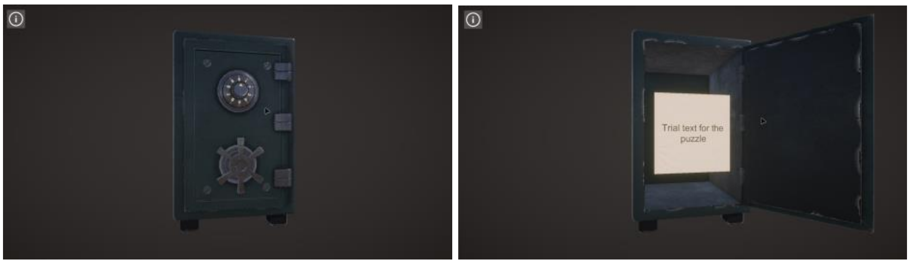


Fig. 7. Safebox puzzle example: a) Locked safebox, b) Open safebox.

Radio puzzle. The radio puzzle is a 3D puzzle that requires the player to adjust the volume and the frequency of a sound to match the reference wave, Fig. 8. In the first place, the player needs to turn on the radio and then use the volume and frequency handles to adjust the sound itself. The radio screen displays the current volume and frequency values of the sound. In addition to that, the end goal of the puzzle is to align the sound wave with the reference wave shown behind it. Just like in the other puzzles, the correct values and the sound itself should be entered trough Escapp. Once the sound is matched correctly, the puzzle is completed. The player can readjust the audio as much as they need to until they have found the correct values.

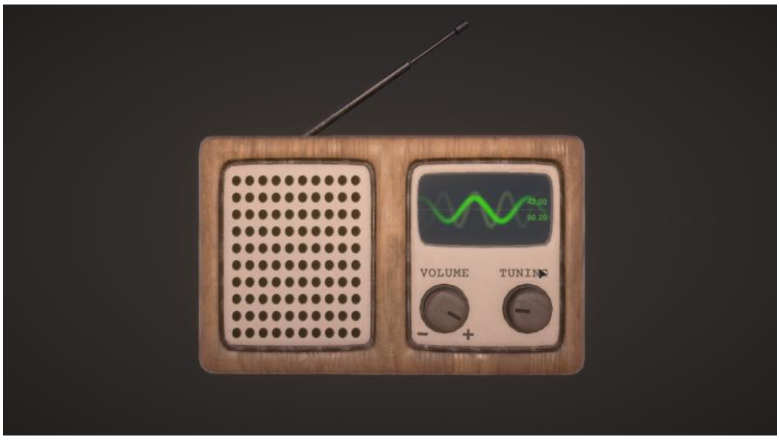


Fig. 8. Radio puzzle made in Unity.

Discussion and conclusion

This work shows how reusable Unity-based puzzles can be integrated into the Escapp platform to enrich educational escape rooms. By using JSON configuration files, puzzle content is separated from logic, allowing educators to easily adapt puzzles to different topics without additional programming.

The integration ensures secure solution validation through Escapp's API and smooth tracking of student progress. This modular design reduces development effort, supports scalability, and encourages puzzle reuse across contexts. Overall, the approach provides a flexible and reusable framework that strengthens the role of digital escape rooms in e-learning.

Acknowledgments. This work has been co-funded by the European Union through the Erasmus+ Programme as part of the project "Implementation of game-based learning using escape rooms" (IGLUE) (Reference: 2024-1-ES01-KA220-HED-000256356). Co-funded by the European Union. The opinions and viewpoints expressed are solely those of the authors and do not necessarily reflect those of the European Union or the Spanish Service for the Internationalization of Education (SEPIE). Neither the European Union nor the granting authority can be held responsible for them.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Veldkamp, A., van de Grint, L., Knippels, M.-C. P. J., & van Joolingen, W. R. (2020). Escape education: A systematic review on escape rooms in education. Edu-cational Research Review, 31, 100364. https://doi.org/10.1016/j.edurev.2020.100364
2. López-Pernas, S. (2024). Educational escape rooms are effective learning activities across educational levels and contexts: A meta-analysis. IEEE Transactions on Learning Technologies, 17, 711–724. https://doi.org/10.1109/tlt.2023.3328913.

3. Lopez-Pernas, S., Gordillo, A., Barra, E., & Quemada, J. (2021). Escapp: A web platform for conducting educational escape rooms. IEEE Access: Practical Innova-tions, Open Solutions, 9, 38062–38077. https://doi.org/10.1109/access.2021.3063711
4. Pan, R., Lo, H., & Neustaedter, C. (2017). Collaboration, awareness, and communi-cation in real-life escape rooms. Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17. the 2017 Conference, Edinburgh, United Kingdom. https://doi.org/10.1145/3064663.3064767
5. Warmelink, H., Mayer, I., Weber, J., Heijligers, B., Haggis, M., Peters, E., & Louwerse, M. (2017, October 15). AMELIO. Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play. CHI PLAY '17: The annual symposium on Computer-Human Interaction in Play, Amsterdam The Neth-erlands. https://doi.org/10.1145/3130859.3131436
6. Williams, P. (2018, March). Using escape room-like puzzles to teach undergraduate students effective and efficient group process skills. 2018 IEEE Integrated STEM Education Conference (ISEC). 2018 IEEE Integrated STEM Education Conference (ISEC), Princeton, NJ, USA. https://doi.org/10.1109/isecon.2018.8340495
7. Wu, C., Wagenschutz, H., & Hein, J. (2018). Promoting leadership and teamwork development through Escape Rooms. Medical Education, 52(5), 561–562. https://doi.org/10.1111/medu.13557.
8. Guckian, J., Eveson, L., & May, H. (2020). The great escape? The rise of the escape room in medical education. Future Healthcare Journal, 7(2), 112–115. https://doi.org/10.7861/fhj.2020-0032
9. Lathwesen, C., & Belova, N. (2021). Escape Rooms in STEM Teaching and Learn-ing—Prospective Field or Declining Trend? A Literature Review. Education Scienc-es, 11(6), 308. https://doi.org/10.3390/educsci11060308.
10. Makri, A., Vlachopoulos, D., & Martina, R. A. (2021). Digital Escape Rooms as In-novative Pedagogical Tools in Education: A Systematic Literature Review. Sustaina-bility: Science Practice and Policy, 13(8), 4587. https://doi.org/10.3390/su13084587
11. Fotaris, P., & Mastoras, T. (2019). Escape rooms for learning: A systematic review. In Lars Elbaek, Gunver Majgaard, Andrea Valente, Saifuddin Khalid (Ed.), Proceed-ings of the 13th European Conference on Games Based Learning (ECGBL 2019) (pp. 235–243). Academic Conferences and Publishing International Limited.
12. Taraldsen, L. H., Haara, F. O., Lysne, M. S., Jensen, P. R., & Jenssen, E. S. (2022). A review on use of escape rooms in education – touching the void. Education In-quiry, 13(2), 169–184. https://doi.org/10.1080/20004508.2020.1860284
13. Vorderobermeier, A., Abel, J., & Sailer, M. (2024). Theoretical foundations and ap-proaches in research on educational escape rooms: A systematic review. Education-al Research ReviewVolume 44, August 2024, 100625.
14. Reinkemeyer, E. A., Chrisman, M., & Patel, S. E. (2022). Escape rooms in nursing education: An integrative review of their use, outcomes, and barriers to implementation. Nurse Education Today, 119(105571), 105571. https://doi.org/10.1016/j.nedt.2022.105571
15. López-Pernas, S., Gordillo, A., Barra, E., & Quemada, J. (2019a). Examining the Use of an Educational Escape Room for Teaching Programming in a Higher Educa-tion Setting. IEEE Access, 7, 31723–31737. https://doi.org/10.1109/access.2019.2902976.

Using prompt engineering to foster critical thinking in higher education

Ana Mirković Moguš

University of Josip Juraj Strossmayer, Faculty of Education, Cara Hadrijana 10, 31000 Osijek, Croatia
amirkovic@foozos.hr

Abstract. This paper presents a systematic review of the latest research (2023–2025) concerning the contribution of prompt design to the development of critical thinking skills in the field of AI-supported eval practices. The review combines studies of empirical and conceptual nature that are examining various prompting strategies, such as role-based, Socratic, scenario-driven, iterative, and chain-of-thought approaches. According to the findings, well designed prompts essentially stimulate the higher-order cognitive processes, which are the ones that involve the critical thinking skills of analysis, creation, and reflective reasoning and also that they lead to increased engagement and metacognitive awareness. The results additionally point to the possibility of prompt engineering to fundamentally change the role of AI as a mere informational resource to that of a co-teacher with active engagement in pedagogy, albeit with some issues still existing in the area of teacher preparation, curriculum modification, and AI literacy. In conclusion, prompt engineering is just as much a teaching practice as it is a technical skill and it is at the core of students' critical and reflective abilities development in modern learning environments.

Keywords: AI in education, prompt design, metacognition

1 Introduction

In a time when the use of AI in higher education is entering very fast, a thorough knowledge of prompt engineering skills could have a positive effect on student's capabilities [1] and would also be able to strengthen their critical thinking skills [2]. There are projects that specifically aim to introduce AI literacy in school level [3]. The primary aim of these projects is to make the students realize the responsible usage of AI, however, they can only do so if they understand how to 'communicate' with an AI to get the desired result. It, therefore, means that students need to be competent prompt engineers [2]. Prompt engineering is basically a clever and deliberate selection of inputs that will bring about the desired response or the right conduct by an AI system. As such, in education, it means the creation of such prompts that would be able to attract students' attention as well as to challenge their critical and creative thinking. The qualities of prompt engineering are found in its capacity to change the AI from a simple information source into a user-friendly tool which encourages more profound learning and

Reveiw Paper
DOI: https://doi.org/10.46793/eLearning2025.235MM
Part of ISBN: 978-86-89755-37-4

comprehension [4]. In prompt engineering, students come up with, verify and improve prompt that would result in desired AI responses, thus immersing them in elaborate problem-solving and reflective thinking.

2 Prompt engineering and critical thinking

Existing research provides various views, some of which warn that AI applications may slow down the development of critical skills, whereas others maintain the position that AI may enhance learning if properly put in practice [5].

Methods of critical thinking are frequently discussed in the literature. Authors of [1] observe that if a learner asks an AI system to prepare a paper it deprives the student of the chance to engage in critical thinking. With the implementation of ChatGPT, there are questions about the decline of the critical thinking skills and the possible obstruction of human relations [6]. Excessive reliance on ChatGPT could result in the lowering of students' cognitive abilities, especially in critical thinking [7]. Critical thinking, by definition is the capability to assess data, reason, and come up with logical solutions, is among the crucial components of higher education, that perhaps AI application could have an influence on [8]. One of the essential skills for learners is the ability to critically think. This is especially true in college and universities where students are supposed to be independent, solve problems, and analyze the given data [9]. But, prompt engineering may be a factor to motivate students to apply their analytical and critical thinking skills, as shown in higher education contexts [10-12], in creative or experiential settings [13-15], and in classroom-based case studies [16]. By using prompts that aim to delve deeper into the topic, produce intriguing questions, or come up with practical solutions can result students to be actively engaged which in turn will higher their cognitive skills [12]. Author of [17] referred to "pivotal components" as the three elements which are the basis for an effective prompt engineering process: Content Knowledge, Critical Thinking, and Iterative Design. He states that one of the most important skills that the user must have when dealing with AI tools is critical thinking which "isn't a mere accessory but a necessity" and also say that the function of the referred skill is to assess, validate and challenge the AI tool's results. The skills that need to be developed here are the aptitude to critically analyze the content of AI responses and to accurately identify when hallucinations, biases, inaccuracies, or any other kinds of inappropriate content are present. These skills are the very basic ones that allow the user to later be able to modify prompts in a proper way.

Working with students to create efficient prompts for large language models such as ChatGPT can sharpen critical thinking capabilities [18, 19]). Moreover, ChatGPT supports user's critical thinking skills as it helps the user with self-regulation by providing frameworks for evaluating responses, thereby reinforcing reflective reasoning [20]. It is important to approach to AI from a critical perspective. One of the ways to develop the skills of prompt engineering is to carry out a critical thinking exercise [2]. The stages of constant assessment and redesigning of instructions, as shown in the prompt engineering templates found in the literature, are main features of the skill development process for human-AI interaction. Author of [21] invites to reflect on the potential application of his CLEAR framework in teaching and asserts that this method will cultivate the critical thinking abilities of learners.

Research indicates that prompt engineering can support inclusivity and equity [2, 22] and enhance reflective or critical questioning approaches [23, 24]. Prompting strategies have been investigated across multiple disciplines, from higher education pedagogy [10-12] to creative and applied learning environments [13-15], with newer work extending into technological innovation and classroom case studies [16, 2].

A study by [24] demonstrated that AI chat models, when integrated with engineered prompts, significantly improved undergraduate students' critical thinking skills, with an average increase of 12 points in critical thinking assessments. Participants also reported high satisfaction and engagement with the AI prompts, indicating their effectiveness in stimulating deeper cognitive engagement [24]. The study also demonstrates that engineered AI prompts can measurably improve students' abilities in analysis, evaluation, and synthesis which are core components of critical thinking. Prompts functioned as structured scaffolds that moved learners from passive knowledge reception to active engagement and finally to reflective judgment.

Other research explored various strategies for prompt engineering, such as role-playing prompts and Socratic prompts, which are designed to enhance learning experiences with ChatGPT. These strategies help tailor prompts to individual needs, fostering engagement and promoting critical thinking skills [22, 25]. Study concludes that prompt engineering acts as both a pedagogical and cognitive scaffold: close prompts build foundational knowledge, open prompts stimulate higher-order reasoning, role-play prompts foster applied decision-making and empathy [22]. Socratic prompts systematically push learners into critical evaluation and reflective reasoning. The survey evidence shows teachers perceived substantial improvements in critical thinking, analytical skills, and metacognition when prompts were carefully designed and aligned with objectives.

Author of [2] emphasized the necessity of AI literacy and prompt engineering proficiency in modern education. The study discussed strategies for embedding these skills within educational curricula to enrich educational experiences and promote critical thinking [2]. It also emphasizes that prompt engineering is not only a technical skill but also a pedagogical strategy. By shifting from simple input/output prompts to chain of thought prompts, tree of thought prompts, and expert prompting, educators can design AI interactions that scaffold analysis, evaluation, reflection, and ethical reasoning which are all core aspects of critical thinking.

Authors of [23] presented the Socratic Playground for Learning (SPL), a GPT-4-based intelligent tutoring system that engages the Socratic method to stimulate the critical thinking skills. By means of prompt engineering, SPL creates educational situations, supports lesson contexts, and manages multi-turn conversations. Experimenting Socratic prompts motivate the students to ask questions, think critically, and reflect.

While these studies highlight the benefits of prompt engineering in fostering critical thinking, they also point to challenges such as the need for comprehensive educator training and curriculum adaptation. The integration of AI in education requires careful consideration of societal impacts and the development of AI literacy among educators and students alike [2].

Even with the increasing interest on AI powered educational solutions, there still remains a substantial gap between the employment of AI to generally improve teaching and learning processes and its use to develop student critical thinking skills via prompting only. This research comprehensively aims to scrutinize how prompt engineering shapes the critical thinking skills in higher education. It seeks to delineate the connection between prompt design and students' logical thinking, assessment of information, and active engagement in critical thinking. The question is which types of prompt engineering strategies in AI-supported learning environments are designed? How are they related to critical thinking outcomes?

3 Methodology

This research used a systematic literature review approach which adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework involving four main stages: identification, screening, eligibility, and inclusion [26]. A literature search was conducted using the Scopus, Web of Science and ERIC databases. The search covered publications considering time interval of the last three years (from 2023 to 2025). The goal of this systematic review is to outline the effects of prompt engineering on the development of students' critical thinking skills. Moreover, this research is intended to grasp the effect of prompt engineering on learners' capacity to critically and independently evaluate and assess information. Databases were searched by keywords: prompt engineering, critical thinking, AI in education.

Following the recognition of relevant documents, it was necessary to select the articles that qualified the inclusion parameters of the study. In order to carry out the initial screening process, various criteria were set, such as eliminating duplicate documents, implementing the exclusion and inclusion criteria, concentrating on items that were pertinent to the objective of the research and restricting the choice to English language papers openly accessible. There were 57 records identified through database searching, 29 full text articles assessed for eligibility and 12 articles included in the study. In order to understand the content of selected articles for this review a deductive thematic analysis approach was employed which consisted of defining and grouping themes on the basis of the research questions [27].

Data were extracted on: prompt types used, reported outcomes on reasoning or critical thinking, identified best practices. In final, findings were synthesized into a framework linking prompt design and critical thinking outcomes. The limitations refer on synthesizing studies with different research designs (conceptual vs. empirical).

4 Results and discussion

Following table presents synthesized types of prompt design and related outcomes on critical thinking (Table 1).

Table 1 *Prompt engineering types and critical thinking outcomes*

Prompt Types Used	Reported Outcomes on Critical Thinking / Reasoning	Study
Knowledge construction, Inquiry-based, Self-assessment, Peer teaching	Scaffolded reasoning, inquiry, metacognition, peer explanation; reduced transmissionist "answer-giving." Sample prompt: "You are a teacher who facilitates inquiry-based learning…"	[10] – Teacher-GAIA
Objective-driven, Context-rich, Role-based, Format-specified, Iterative, Example-anchored, Ethical prompts	Promoted dialogue, reflective questioning, deeper analysis; highlighted risk of shallow reasoning if poorly crafted.	[11] – Conceptual
Goal-aligned, Example-based, Lesson planning, Problem-solving, Engagement/debate, short answer question, multiple choice questions, Assessment, essay question prompt	Supported higher-order reasoning (analysis, evaluation, synthesis), authentic assessment, divergent/convergent thinking. Sample prompt: "Craft a complex scenario related to [topic]."	[12] – HE Teaching
Empathise, Define, Contextualise, Design puzzles, Briefing, Debriefing, Prototype, Evaluation prompts	Fostered problem-solving, reflective reasoning, collaborative analysis, metacognition through game-based design. Sample prompt: "Write a debriefing guide with reflection questions…"	[13] – Escape Rooms
Descriptive, Iterative/variation, Synectic triggers, Peer reinterpretation, Ethical reflection	Encouraged conceptual refinement, divergent thinking, metacognition, ethical reflection on AI use.	[14] – Art Education
Open idea generation, Elaboration, Contextualisation, Specification, Risk assessment, Instructional, Assessment design, Self-critique	Enhanced analytical reasoning, problem-solving, reflective thinking, collaborative reasoning; required human critical validation of AI outputs. Sample prompt: "Critique your suggestion for the marine litter field activity…"	[15] – Field Course Design
Analytical, Evaluative, Synthetic, Reflective/metacognitive, Perspective-taking prompts	Measurable gains in analysis (+7.8), evaluation (+7.8), synthesis (+9.8); 83% reported broader perspectives; strong metacognitive development.	[24] – Empirical Study
Close question prompts, Open question	Survey (N=120 teachers): Prompts fostered evaluation (4.3/5),	[22] – Teacher Survey

prompts, Role-playing, Socratic prompts;	exploration (4.1/5), reflection (4.1/5); promoted perspective-taking and collaborative reasoning. Sample prompt: "Provide examples and discuss the mechanisms…"	
Input-output prompting, Zero-/Few-shot, Chain-of-thought (CoT) prompting, Self-consistency prompting, Tree-of-thought (ToT), Role-play or expert prompting, automatic prompt, generating knowledge	CoT/ToT - structured reasoning; Adversarial - critical evaluation of AI; Role-based - applied judgment; scaffolded analysis, evaluation, ethical reasoning. Sample prompt: "Provide me step by step…";	[2] – AI Literacy
Lesson creation, Socratic dialogue (What, Why, How), Adaptive feedback, Scenario-based	Graduate pilot: +80% positive on understanding/motivation; iterative Socratic prompts stimulated critical questioning, reflection, synthesis, and problem-solving. Sample prompt: "What aspect of the context do you find most challenging to understand?"	[23] – Socratic Playground (SPL)
Standard (few-shot) prompting, chain of thought prompting, ablation prompt variations	Arithmetic reasoning, commonsense reasoning, symbolic reasoning,	[28]
Expert identity prompts, instruction prompts	Evaluation and judgment, reflective and contextual reasoning, analytical and explanatory depth	[29]

Research of [10] illustrates that prompt engineering is capable of converting LLM results to switch logically from the detailed answer provision to the facilitation of the reasoning process. Such change is consistent with advanced cognitive functions from Bloom's taxonomy (analysis, evaluation, creation) and with the Paul–Elder critical thinking framework (reasoning, questioning, metacognition). Hence critical thinking, logical reasoning, and self-regulation were enhanced by the use of carefully tailored persona prompts, which elicited these skills to a far lesser extent when generic prompts were used.

Authors of [11] state that prompt engineering forms the core of constructing a "conversational pedagogy" with AI. Although the paper is conceptual (not an empirical study), it draws implications for critical thinking and reasoning in education. Good prompts can stimulate critical thinking, deeper reasoning, and co-creation of knowledge. Bad prompts can lead to the situation whereby students continue to learn only at a surface level or even get false information. Therefore, the development of critical thinking skills is very much dependent on the characteristics of prompts (clarity, context, structure, and purpose).

The study of [12] shows that carefully engineered prompts can be used by educators to: generate content that elicits critical thinking from students, scaffold problem-solving tasks, and redesign assessments to be more authentic, reflective, and reasoning-oriented. In other words, prompt engineering acts as a pedagogical design tool that supports the creation of learning tasks requiring analysis, evaluation, and argumentation which are the building blocks of critical thinking.

Well-engineered prompts transform ChatGPT from a content generator into a pedagogical design partner [13]. By embedding narrative, puzzles, reflection, and evaluation, prompts create immersive escape rooms that demand learners' critical thinking, reasoning, and problem-solving skills.

The study of [14] integrated OpenAI's DALL-E 2 into digital art courses. Students were required to design and refine prompts as part of their creative process. The study demonstrates that prompt engineering, when used iteratively and collaboratively, supports critical reasoning in at least three ways:

Cognitive – analyzing the effect of language on AI outputs (analytical thinking).

Creative – exploring divergent design possibilities (imaginative reasoning).

Ethical/Reflective – evaluating the implications of AI use in art (critical reflection).

Authors of [15] demonstrate that iterative, structured prompts can scaffold the design of field courses that foster data analysis, environmental reasoning, risk assessment, and reflective discussion. At the same time, the need for human oversight and adaptive management highlights critical thinking as both a learning outcome and a methodological safeguard. Authors of [28] established Chain of Thought (CoT) prompting as a foundational method for eliciting higher-order reasoning in large language models.

Authors of [29] demonstrate that expert identity prompting effectively scaffolds LLM reasoning by embedding domain-specific perspectives. This improves analytical depth, explanatory detail, and evaluative precision which are all central to critical thinking. Moreover, the ExpertLLaMA model trained on such data showed that systematic prompt engineering can enhance reasoning capabilities at scale.

This research underlined two different pathways by which prompt engineering has developed, teaching and algorithmic, and illustrates how they are joining together at the same result with the implications of critical thinking. Educational studies emphasize the importance of role-based dialogue and Socratic questioning [10, 11, 23], scenario-based and field applications [12, 15]), and creative integration in arts education [14]. Algorithmic approaches, by contrast, focus on refining reasoning through systematic design, as seen in chain-of-thought, self-consistency, and tree-of-thought methods [28, 16], as well as automated and expert prompts [29]. Despite their different origins, both strands converge on fostering analytical depth, reflective practice, and evaluative precision. Such method raises the reliability, correctness, and the logical framework of the reasoning system in the large language models, thus opening up new prospects of their application in education. Though the two are different from the outset, the paths of both strands meet at common critical thinking outcomes. Despite this review showcasing the capability of prompt engineering to facilitate critical thinking, it is still necessary to have more proof from real cases in college environments. A set of possible future research areas could be controlled classroom experiments, longitudinal studies, disciplinary case studies and mixed-methods approaches combined with quantitative testing and qualitative methods (student interviews, reflective journals, classroom observations) for obtaining students' perceptions of prompts and understanding their reasoning processes.

5 Conclusion

This review demonstrates that prompt engineering spans both educational strategies and algorithmic innovations. New hybrid models are opening the way for linking these two areas with a common goal, i.e. to develop critical thinking skills through analysis, evaluation, synthesis and reflection.

When prompts are aligned with learning objectives, evidence shows quantifiable improvements in analytical reasoning, evaluative judgment, synthesis of ideas, and metacognitive reflection. In addition, the results highlight the need for teacher training, curriculum modification, and the teaching of critical AI literacy to guarantee that the advantages of prompt engineering are brought about and at the same time controlling the risks like simple logic or AI hallucinations.

In conclusion, designing prompt should be considered just as much a teaching skill as a technical skill, which could noticeably change the way users interact with AI, if used properly. This is because it is not only easy to access source of information anymore but an engaged partner that helps the user to learn through critical and reflective thinking.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

[1] Heston, T. F., Khun, C.: Prompt engineering in medical education. International Medical Education, 2, 198–205 (2023). https://doi. org/10.3390/ime2030019

[2] Walter, Y.: Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. Int. J. Educ. Technol. High. Educ. 21, 1–29 (2024). https://doi.org/10.1186/s41239-024-00448-3

[3] Tseng, Y. J., Yadav, G.: ActiveAI: Introducing AI literacy for middle school learners with goal-based scenario learning (arXiv:2309.12337) (2023). https://doi.org/10.48550/arXiv.2309.12337

[4] Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., Kim, H.: Few-shot is enough: Exploring ChatGPT prompt engineering method for automatic question generation in english education. Education and Information Technologies (2023). https://doi.org/10.1007/s10639-023-12249-8

[5] Vargas-Murillo, A. R., Pari-Bedoya, I. N. M. de la A., & Guevara-Soto, F. de J. Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. International Journal of Learning, Teaching and Educational Research, 22(7), 122 – 135 (2023). https://doi.org/10.26803/ijlter.22.7.7

[6] Graefen, B., Fazal, N.: From Chat bots to virtual tutors: An overview of Chat GPT's role in the future of education. Arch. Pharm. Pract. 15(2), 43-52 (2024). https://doi.org/10.51847/touppjedsx
[7] Bai, L., Liu, X., Su, J.: ChatGPT: The cognitive effects on learning and memory. Brain-X, 1(3), 1–9 (2023). https://doi.org/10.1002/brx2.30
[8] Essien, A., Bukoye, O. T., O'Dea, X., Kremantzis, M.: The influence of AI text generators on critical thinking skills in UK business schools. Studies in Higher Education, 49(5), 865–882, (2024). https://doi.org/10.1080/03075079.2024.2316881
[9] Nufus, R. H., Octoviani, R., Eliyana, S., Sari, I. J.: The Implementation of the Teaching Campus Program on Students' 21st Century Skills: A Systematic Review, Jurnal Pendidikan Indonesia Gemilang, 3(2), 278–294 (2023).
[10] Ali, F., Choy, D., Divaharan, S., Tay, H. Y., & Chen, W.: Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering. Learning: Research and Practice, 9(2), 135–214 (2023). https://doi.org/10.1080/23735082.2023.2258886
[11] Bozkurt, A., Sharma, R. C.: Generative AI and prompt engineering: The art of whispering to let the genie out of the algorithmic world. Asian Journal of Distance Education, 18(2), i–vii (2023). https://doi.org/10.5281/zenodo.8174941
[12] Eager, B., Brunton, R.: Prompting higher education towards AI-augmented teaching and learning practice. Journal of University Teaching & Learning Practice, (2023). https://doi.org/10.53761/1.20.5.02
[13] Fotaris, P., Mastoras, T., Lameras, P.: Designing educational escape rooms with generative AI: A framework and ChatGPT prompt engineering guide, 17th European Conference on Games Based Learning, 17(1), 180–189 (2023). https:// doi.org/10.34190/ecgbl.17.1.1870
[14] Hutson, J., Cotroneo, P.: Generative AI tools in art education: Exploring prompt engineering and iterative processes for enhanced creativity. Faculty Scholarship. 477 (2023). https:// digitalcommons.lindenwood.edu/faculty-research-papers/ 477
[15] Tupper, M., Hendy, I. W., Shipway, J. R.: Field courses for dummies: can ChatGPT design a higher education field course? EdArXiv Preprints (2023). https://doi.org/10 35542/osf.io/b65nu
[16] Wang, M., Wang, M., Xu, X., Yang, L., Cai, D., Yin, M.: Unleashing ChatGPT's power: A case study on optimizing information retrieval in flipped classrooms via prompt engineering. IEEE Transactions on Learning Technologies. (2023) https:// doi.org/10.1109/TLT.2023.3324714
[17] Cain, W.: Prompting change: Exploring prompt engineering in large language model AI and its potential to transform education. TechTrends. (2023). https://doi.org/10.1007/s11528-023-00896-0
[18] Hu, L.-K., Lee, D., C.: Prompt Engineering for Critical Thinking and Equity in General Education, American Research Journal of Humanities and Social Sciences, 10(1), 122-129 (2024).
[19] Atlas, S.: ChatGPT for higher education and professional development: A guide to conversational AI. University of Rhode Island Publications (2023). https://digitalcommons.uri.edu/cba_facpubs/548
[20] van den Berg, G., du Plessis, E.: ChatGPT and generative AI: possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. Education Sciences, 13, 1–12 (2023). https://doi.org/10.3390/educsci13100998
[21] Lo, L. S.: The CLEAR path: A framework for enhancing information literacy through prompt engineering. The Journal of Academic Librarianship, 49(4), 102720 (2023). https://doi.org/10.1016/j.acalib.2023.102720

[22] Leung, C. H.: Promoting optimal learning with ChatGPT: A comprehensive exploration of prompt engineering in education. Asian Journal of Contemporary Education, 8(2), 104–114 (2024). https://doi.org/10.55493/5052.v8i2.5101
[23] Zhang, L., Lin, J., Kuang, Z., Xu, S., Yeasin, M., & Hu, X.: SPL: A Socratic Playground for Learning Powered by Large Language Mode (2024). https://doi.org/10.48550/arxiv.2406.13919
[24] Bushell, C.: Investigating the Integration of AI Chat Models to Enhance Critical Thinking Skills through Engineered Prompting. (2025). https://doi.org/10.2139/ssrn.5017061
[25] Fabio, R.A., Plebe, A., Suriano, R.: AI-based chatbot interactions and critical thinking skills: an exploratory study, Curr Psychol 44, 8082–8095 (2025). https://doi.org/10.1007/s12144-024-06795-8
[26] Page, M. J., McKenzie, J. E., Bossuyt, P., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D.: The prisma 2020 statement: An updated guideline for reporting systematic reviews. Medicina Fluminensis, 57(4), 444–465 (2021). https://doi.org/10.21860/medflum2021_264903
[27] Nowell, L. S., Norris, J. M., White, D. E., Moules, N. J.: Thematic analysis: Striving to meet the trustworthiness criteria. International Journal of Qualitative Methods, 16(1), 1–13 (2017). https://doi.org/10.1177/1609406917733847
[28] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., Zhou, D.: Chain-of-thought prompting elicits reasoning in large language models. arXiv: 2201.11903 (2023). https://doi.org/10.48550/arXiv.2201.11903
[29] Xu, B., Yang, A., Lin, J., Wang, Q., Zhou, C., Zhang, Y., Mao, Z.: Expertprompting: Instructing large language models to be distinguished experts. arXiv:2305.14688. (2023). https://doi.org/10.48550/arXiv.2305.14688

Digital Literacy and Competences in the Field of Tourism and Hospitality in English Language Teaching in the Third Millennium

Brankica Bojovich[1][0009-0009-0279-7686]

[1] University Mediterranean, Faculty of Tourism, Montenegro
brankicaboj@gmail.com

Abstract. This study investigates the extent of digital literacy among students specializing in Tourism and Hospitality, highlighting the critical necessity of acquiring advanced digital skills to sustain competitiveness within a digitally-driven, multilingual labor market. The theoretical framework underpinning this research incorporates the utilization of the Internet, websites, and various online tools that facilitate the introduction of innovative methodologies and resources in the specialized context of English language instruction.

To examine whether modern digital tools contribute to the development of digital competences, mixed-methods research was employed. Data were collected from approximately 150 undergraduate students of the Faculty of Tourism at the University of Mediterranean during the academic years 2023/24 and 2024/25. The data collection instruments included online questionnaires (Google Forms), semi-structured interviews, as well as follow-up communication via email, short message service (SMS), phone calls and Viber calls.

The data were analyzed using statistical software (SPSS and Excel) and thematic coding to identify patterns in digital literacy and English language use. The analysis draws on prior research (Bojovich, 2016; Gee & Hayes, 2011; Kwok, 2023) and focuses on three key aspects: (1) students' mastery of the four language skills in English, (2) their digital literacy within the context of Tourism and Hospitality and (3) the digital competences required in this professional field. Particular attention is paid to the practical application of English as a lingua franca in tourism marketing, reflecting current industry needs.

The findings indicate that digital tools play a significant role in shaping students' language acquisition and digital skills. Moreover, the results suggest a paradigmatic shift in teaching practices due to the growing integration of digital technologies. These changes call for an adaptation of pedagogical strategies to better prepare students for the evolving demands of the tourism and hospitality sectors. This study is an empirical investigation that combines both qualitative and quantitative methods to provide a comprehensive understanding of the development of digital literacy among Tourism and Hospitality students.

Keywords: Digital literacy, Digital competence, English for Specific Purposes.

1 Introduction

The tourism and hospitality industry is undergoing rapid digital transformation, significantly influencing how professionals interact with international guests and

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.245B
Part of ISBN: 978-86-89755-37-4

manage day-to-day operations. English, as the global lingua franca, remains the dominant medium of communication, especially in online environments.[1] Consequently, mastering both English language and digital competencies has become essential for tourism professionals who aim to thrive in a technology-driven and globally competitive market.

Kwok [2] in the article "Tech-Savvy Tourism: Trends and Innovations" published in the Journal of Hospitality and Tourism Technology, examines the transformative impact of digital technologies on the tourism and hospitality sector. The study emphasizes the evolving role of technology, from a supportive tool to a central enabler of market competitiveness by facilitating data collection and enabling personalized service offerings. Moreover, the article highlights how technological advancements have shifted power toward consumers, enhancing their ability to make informed decisions and actively participate in the co-creation of value.

In this context, ESP for tourism must cover both language skills and digital literacy. This can be more effectively achieved by referencing models like DigComp, which provide a framework for digital competences. This is in line with the European Digital Competence Framework (DigComp), which outlines essential digital competencies for both personal and professional development. According to DigComp, students should be proficient in the use of digital tools, content creation, and communication. Our study found that the majority of students scored above the basic proficiency level in these areas, although there were variations based on their previous exposure to digital tools and English.

The evolving nature of the tourism sector demands that professionals operate fluently within digital ecosystems, from booking platforms and social media to CRM (Customer Relationship Management) systems[1] and live chat support tools.

This study explores how students of tourism and hospitality develop digital literacy and English language competencies, examining the effectiveness of integrating digital tools into language instruction.

This study also examines linguistic interference between English and Serbian/Montenegrin in digital contexts, with over 60 lexemes identified from communication via Viber, phone, and Zoom.

Additionally, the study contributes to the discourse on the role of digital technology in facilitating lexical borrowing and English language acquisition. It underscores both the benefits and drawbacks of technological influence on language interference, particularly among students, whose growing dependence on online communication is increasingly reflected in their lexical choices and everyday discourse.

The essence of the social and communicative aspect of the Internet lies in the interconnection of participants within an online community, where the central activities include communication, information exchange, and interaction. Participants establish an entire network of social relations through both synchronous and asynchronous communication, thereby creating an online social and public sphere.[3]

---

[1] CRM is a system (software) that companies use to: track customer contacts – including name, phone number, email address, and history of previous communication, manage sales processes – identifying potential customers, tracking who has purchased, and monitoring customer interest, analyze customer behavior – understanding what they searched for, when, and how often, improve customer support – by providing faster solutions since all client information is already stored in the system.

Moreover, student networks are monitored due to the use of anglicized vocabulary.[4] Numerous anglicized examples from the student corpus are the following translated or transcribed into their English equivalents such as Instagram, to hate (hejtovati), to reset (resetovati), to update (apdejtovati), to install (instalirati), and web (veb), mail (mejl), blog (blog), clip (klip), cyber (sajber), videophone (videofon), spam (spem), inbox (inboks), smartphone (smartfon), connection (konekcija), application (aplikacija), implementation (implementacija), attachment (atačment), to attach (atačovati), to download (daunlodovati), to upload (aploudovati), input (input), online (onlajn), to chat (četovati), notification (notifikacija), navigation (navigacija), password (pasvord), camera (kamera), game (gejm), Viber (vajber), Zoom (zum), player (plejer), Bluetooth (blutut), Wi-Fi (vaj faj), channel (kanal), chip (čip), antivirus program (antivirusni program), backup (bek ap), browser (brouzer), to create a program (kreirati program), to delete (dilitovati), desktop (desktop), display (displej), driver (drajver), file (fajl), filter (filter), gadget (gadžet), to generate (generisati), header (heder), folder (folder), host (host), interface (interfejs), to install (instalirati), provider (provajder), link (link), log (log), menu (meni), media (media), mouse (maus), offline (oflajn), online (onlajn), option (opcija), system (sistem), screenshot (skrinšotovati).

Words are, therefore, an instrument through which we acquire knowledge of the language and culture of a foreign country. At the same time, by learning foreign words, the student expands the meanings of words in their native language, thereby gaining a better understanding of the concepts those words express. Thus, knowledge of vocabulary plays an important role in mastering a language.[5]

Digital communication includes selfies, various posts, and explanations covering both general and professional topics. Communication takes place through emails, Zoom meetings, online and Viber social groups, blogs, and video clips. Digitalization as a process is interesting to students, especially to those who have access to new or cutting-edge technological trends.

2 Methods

This empirical study adopts a mixed-methods approach to assess the digital and linguistic competencies of students specializing in tourism and hospitality. The research was conducted at the Faculty of Tourism, University of Mediterranean, during the academic years 2023/24 and 2024/25. The innovation of this study is based on a clear data setup and analysis methods, as well as the participants, data collection instruments, and data analysis, in order to achieve the research results.

Table 1. Data Collection and Analysis Methods

Category	Description
Data Collection Instruments	Online questionnaires (Google Forms), Semi-structured interviews
Data Analysis Tools	Follow-up communication: email, SMS, phone calls, and Viber calls

Analysis Methods	SPSS (Statistical Package for the Social Sciences)[2] - Microsoft Excel
Focus of Analysis	Statistical analysis (quantitative data) - Thematic coding (qualitative data)

Approximately 150 undergraduate students enrolled in tourism and hospitality programs participated in the study.

Data were collected through online questionnaires distributed via Google Forms to gather quantitative information, while semi-structured interviews were conducted to obtain qualitative insights. Follow-up communication was carried out using email, SMS, phone calls, and Viber. Quantitative data were analyzed using SPSS (Statistical Package for the Social Sciences) and Microsoft Excel, with a focus on frequency, percentage, and correlation analysis. SPSS facilitated data entry, management, analysis, and visualization through its user-friendly interface. Qualitative data from the interviews were analyzed using thematic coding, allowing for the identification of recurring patterns and categories related to students' use of English and digital tools.

3 Results and discussion

Table 2. 1. Research Results (n = 150)

Digital Tool	% of Students Using Frequently
Google Docs	87%
Microsoft Teams	68%
Zoom	75%
Canva	52%
ChatGPT / AI Tools	61%

Table 3. Self-Assessment of Digital Competences

Competence Level	% of Students
High	39%
Medium	51%
Low	10%

Use of English in Digital Contexts

- 78% of students reported frequent use of English when using digital tools.
- 65% stated that digital technology helped them improve their understanding of English.
- 22% reported feeling insecure when using English in online communication.

Identified Themes (from Thematic Coding of Interviews)

- Increased Digital Autonomy: Students reported that online tools enabled more effective learning management.

[2] SPSS (Statistical Package for the Social Sciences) is a software package used for statistical data analysis. It is used in various fields such as social sciences, business and many others for processing, analyzing, and interpreting quantitative data

- Language Barriers: A number of students mentioned difficulties using more complex English terms in digital environments.
- Positive Perception of AI Tools: Most participants felt that tools like ChatGPT made research and writing tasks easier and more efficient.

The results of this study highlight the essential role that digital tools and English language proficiency play in the daily tasks of tourism and hospitality students. The integration of technology into their educational process significantly influences their vocabulary, communication strategies, and the development of digital literacy.

From the online questionaries, 61% of students reported that they use digital platforms for learning purposes, such as online courses and educational videos, and 74% stated that they frequently engage in social media communication related to their field of study. Social media platforms like Facebook, Instagram, and YouTube emerged as key tools for gathering information about the latest tourism trends, destination marketing, and industry news. The highest frequency of English usage was noted in online interactions with tourists, where 56% of students indicated that English was their primary language for communication.

In contrast, 52% of respondents acknowledged that they have encountered difficulties in understanding specific tourism-related terms, especially those linked to digital services (e.g., booking engine, virtual tour, customer engagement platform). Students reported that they often rely on online dictionaries or machine translation tools (e.g., Cambridge online dictionary, Trados, Lingue, Google) to overcome these language barriers. This reliance on technology aligns with the study by Coyle et al. [6] which emphasized that students in technical fields frequently face challenges in understanding discipline-specific terminology in English, particularly when it is associated with emerging digital technologies.

When asked to evaluate their own digital competences, 39% of students rated themselves as highly competent, while a majority of 51% identified their skill level as moderate, indicating that most students are still developing their digital competencies. Only 10% considered themselves to have low digital skills. This distribution suggests that the majority of students possess a functional level of digital literacy, with a significant portion exhibiting advanced skills. These findings support the idea that exposure to digital tools contributes to the enhancement of digital competences.

Also, the study highlighted cases of code-switching and translanguaging. i.e. the practice of mixing English with Serbian/Montenegrin in digital communication. For instance, students would often use English technical terms such as update, subscribe or click, while continuing their communication in Serbian/Montenegrin. This phenomenon suggests that digital competence, in the context of tourism, encourages students to blend linguistic elements from both languages. Furthermore, the use of anglicisms seems to be increasing, especially in informal communication on platforms such as Viber and WhatsApp.

The integration of digital communication tools into the curriculum was shown to have a positive impact on students' professional development. They feel more confident in their ability to handle digital marketing tasks (e.g., creating content for social media), use CRM software, and engage in e-commerce transactions. Additionally, many students reported an increase in their interest in digital entrepreneurship, as they saw the practical benefits of using digital tools for personal branding and promoting tourism services online.

3.1 **English Language Use in Digital Environments**

James Paul Gee and Elisabeth Hayes explore how digital media are reshaping not only the way we use language but also how we learn. They argue that digital technologies such as video games, social media, and online communities have created new forms of literacy and learning that differ significantly from traditional schooling.[7]

English plays a key role in students' digital experiences. Seventy-eight percent (78%) of respondents reported frequent use of English when interacting with digital tools and platforms, especially those with no localized language options. Moreover, 65% stated that their exposure to English through digital environments has contributed to improved comprehension and vocabulary.

Despite this, 22% of students reported feeling insecure when using English in online communication, particularly in professional or academic contexts. This points to the need for continued language support in higher education, especially in programs that rely heavily on digital resources and international communication

3.2 **Thematic Insights from Interviews**

Qualitative data from semi-structured interviews and follow-up communication revealed several themes:

- Increased Digital Autonomy: Students noted that modern tools allowed them to manage their learning more independently, organize their schedules, and access diverse learning resources.
- Language Barriers: Some students, particularly those with lower English proficiency, expressed challenges in understanding specialized terminology or navigating platforms available only in English.
- Positive Attitudes Toward AI Tools: Many participants viewed AI technologies, such as ChatGPT, as helpful companions in the learning process, offering assistance with writing, clarification of concepts, and even translation.

Overall, the data suggest that modern digital tools play a significant role in enhancing both digital competences and English language proficiency among students. However, challenges related to language confidence and digital literacy gaps still persist for a minority of users. These findings underline the importance of integrating digital skills and English language support into university curricula, particularly in programs preparing students for the global tourism industry.

The above mentioned findings are categorized into three core areas: (1) students' English language proficiency, (2) their digital literacy in the tourism context, and (3) the intersection of digital tools and language use.

(1) Students' English language proficiency The majority of students demonstrated functional English proficiency, particularly in reading and listening. However, speaking and writing skills varied, especially in formal or professional settings. Common challenges included limited vocabulary for tourism-specific contexts and lack of confidence in spoken interactions.

(2) their digital literacy in the tourism context Students showed high familiarity with basic digital tools such as booking websites (e.g. Booking.com, TripAdvisor), social media platforms (Instagram, Facebook), and communication tools (WhatsApp, email).

However, their usage was often limited to consumer roles rather than professional applications. CRM systems and POS (Point of Sale) software were less familiar, indicating a gap in industry-specific digital training. (3) the intersection of digital tools and language use It refers to the integration of language and digital skills. Students who actively engaged in e-learning platforms (e.g. Coursera, Udemy, Alison) reported improved digital and language skills. Many noted that using English-language tools (e.g. Google Maps, virtual tours, chatbots) helped them acquire vocabulary and functional expressions in a practical context.

4 Concluding remarks

The results confirm the interdependence of digital and linguistic competencies in the tourism and hospitality industry. While students demonstrate moderate proficiency in both domains, the integration of English language instruction with digital tools remains inconsistent.

The current findings call for a pedagogical shift toward more integrated ESP curricula that reflect real-world digital interactions. Role-playing in simulated environments (e.g. hotel reception, virtual customer service) and tasks involving online content creation (e.g. writing reviews, responding to emails) could enhance both language and digital skills.

Challenges such as outdated curricula, lack of access to digital tools, and limited opportunities for real-world practice remain significant. The fear of making mistakes in English, especially in public or digital contexts, hinders active participation. Institutions must adopt blended learning models, incorporate industry software into instruction, and foster confidence-building strategies.

This study highlights the critical role of digital literacy and English language proficiency in preparing future professionals for the tourism and hospitality industry. Students must be equipped not only with academic knowledge but also with practical, industry-relevant skills. Digital tools significantly enhance language acquisition and professional readiness, integrated instruction leads to better learning outcomes and greater student engagement, there is a need for curriculum reform that reflects the digital transformation of tourism, and e-learning platforms offer scalable solutions for ongoing professional development.

Additionally, the study highlights the growing use of anglicized vocabulary in online communication, especially among students in tourism and hospitality, where English terms are often integrated into everyday digital interactions, reflecting both industry trends and the influence of globalized communication norms. Moreover, student communication networks are increasingly monitored due to the frequent use of anglicized vocabulary. Numerous examples from the student corpus illustrate this trend, with English words either directly borrowed or phonetically transcribed into the local language. In the overall analysis, the frequent presence of anglicized vocabulary was observed, indicating that many English-origin terms have become fully integrated into everyday language use among students.[8]

Despite the advancements, there are still challenges in overcoming language barriers, especially concerning industry-specific terminology. Students' increasing reliance on digital tools and English for Specific Purposes reflects the growing importance of digital literacy as a core skill in tourism education.

This study also emphasizes the need for targeted English for Specific Purposes (ESP) courses that combine linguistic knowledge with digital competence to help future professionals navigate the complexities of the modern digital environment. It suggests that further research should investigate how emerging technologies, such as AI tools and virtual assistants, can be integrated into tourism education to further enhance both language learning and digital skill development.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Bojovich, B., Osnovi translotologije, Filološki fakultet Nikšić, Univerzitet Crne Gore. pp. 118-125 (2016). [1]
2. Bojovich, B., Models of translation procedures in the processing of specific cultural terms in Comprehensive Serbian-English dictionary, Srpski jezik, vol.28, Beograd, pp. 371-386 (2023). [8]
3. Coyle, D., Holmes, B., & King, L. Towards the European Digital Competence Framework: A Report of the European Commission's DigComp Framework Working Group. European Commission. (2017). [6]
4. Filipović, J., Vučo, J., & Đurić, Lj. Rano učenje stranih jezika u Srbiji. Inovacije u nastavi, 19(2), Beograd: Univerzitet u Beogradu, Učiteljski fakultet. pp. 113–124 (2006). [5]
5. Gee, J.P. & Hayes, E.R., Language and Learning in the Digital Age. Routledge. (2011). [7]
6. Habermas J., The structural transformation of the public sphere: An inquiry into a category of bourgeois society. The MIT Press, Cambridge, Massachusetts (1991).[3]
7. Hornby, M. S., The Turns of Translation Studies: New Paradigms or Shifting Viewpoints? Springer. (2006). [4]
8. Kwok, L., Tech-Savvy Tourism: Trends and Innovations, Journal of Hospitality and Tourism Technology, Emerald Publishing Limited (2023). [2]

The Impact of Gamification on the Motivation of Online Students

Marina Grubor[1[0009-0000-0130-5197]], Miloš Stojanović[2[0009-0009-3774-6483]], Pavle Mitić[3[0009-0003-3616-8999]], Anja Veličković[4[0009-0008-8192-469X]]

[1,2,3,4] Belgrade Metropolitan University, Serbia
marina.grubor@metropolitan.ac.rs,
milos.stojanovic@metropolitan.ac.rs,
pavle.mitic@metropolitan.ac.rs,
anja.velickovic@metropolitan.ac.rs

Abstract. It's getting harder and harder for schools of the digital age to keep students interested, motivated and engaged, particularly in an online learning environment. The use of game elements not in the context of games, or gamification, provides innovative ways to increase emotional engagement and interactivity in the classroom. The research provides discussions noting the impact of gamification on motivation of e-learners in real situations and its studies applied at Metropolitan University. It also discusses the neuromarketing principles for enhancing instructional design through cognitive and affective triggers. In particular, the potential barriers for real-world use are discussed: moral questionings, decreased intrinsically motivation to act and dependence on technology. The findings of the study suggest that learners can benefit from gamification as it could potentially contribute to raising the quality of education and to the development of crucial competences in 21st century skills, provided that it is properly designed and supported pedagogically.

Keywords: Gamification, Neuromarketing, Motivation.

1 Introduction

With the development of information, communication and learning technologies there is a new "digital age" in which we are experiencing significant changes in education. This is no superficial assimilation of new tools; it is a profound change in the logic of education itself. Age-old models (teacher as the sage on stage who possessed information that they then imparted to students) are being replaced by shared, interactive and synergistic methods of learning. With the rise of to the internet, artificial intelligence, virtual and augmented reality systems, new possibilities are arising that we can utilize to customize education for each specific student and deliver knowledge in novel ways.

In addition, the digital world has drastically changed communicating -- students can communicate with teachers not only in person but also via multiple digital media channels instantly, so that we can track learning progress better and provide them feedback.

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.253G
Part of ISBN: 978-86-89755-37-4

This transition involves not only technological tools but a shift in educational philosophy and educator and learner mindset. In this context, it is important to know how such tools promote the acquisition of 21st-century skills that are considered fundamental in education, which include creativity, critical thinking, collaboration and digital literacy.

As part of this paper, a research design will be presented, to be conducted at Metropolitan University, with the aim of identifying the effects of gamification on students.

2 Online Education and Gamification

What were perceived to be alternative or experimental ways of learning online in the past, are now mainstream practices in many parts of the world. The widespread use of the internet, advances in reliable digital platforms and the current pandemic of COVID-19 have all contributed to the widespread expansion of online learning. In many universities and colleges, online learning has now become the most common form of teaching, radically transforming understandings of - and approaches to - this mode of education. Oyebode- Fakuade says: There was resistance to quality and effectiveness of online education initially but this has largely turned into broad acceptance and major investment in intuitive digital tools, software and teaching techniques.

Distance learning facilitates the achievement of hybrid models of learning that merge the best in digital and face-to-face forms, thus facilitating easier accommodation to different needs and learning styles [1]. Additionally, access to a broad variety of courses and programs globally has opened up education opportunities for people from remote or marginal communities, thus facilitating international democratization of knowledge. But this model also raises significant concerns regarding quality assurance, accreditation, socialization of students, and long-term impact—all of which necessitate ongoing study and ongoing revision of policies and educational practice. However, this setup brings some real headaches, too—like, who's making sure the quality doesn't tank? Who's actually accrediting this stuff? And don't even get me started on whether students actually get to, you know, make friends or learn how to be humans together. Long-term? Jury's out. Honestly, figuring all this out isn't a one-and-done thing. Schools and researchers pretty much have to tweak and rethink stuff every year just to keep up.

Sure, tech lets more people get their hands on knowledge—it's easier to jump in, learn wherever, and there's less gatekeeping. But let's be real: keeping students pumped about their classes online? That's a whole other beast. Motivation's all over the place. You've got all these things crashing together—kids' own ambitions, stuff happening around them, if their friends are there to hype them up, personal goals, how much freedom they have to run the show. It's not one-size-fits-all, that's for sure.While digital environments offer various benefits, they can also lead to a decline in motivation due to excessive distractions, attention fragmentation, and feelings of isolation [4]. Distance learning often demands a high level of self-discipline, organization, and intrinsic motivation, which many students find challenging—particularly in the absence of adequate support or when they experience a sense of disconnection from the academic community.

In addition, information overload can result in cognitive fatigue, reducing the willingness to engage in learning and encouraging procrastination. Therefore, in contemporary education, it is essential to develop strategies that go beyond mere technical access to knowledge and actively foster and sustain learner motivation and engagement. Creating supportive and stimulating learning environments is crucial for promoting continuous learning and personal development.

3 Online Education and Gamification

The current digital ecosystem has definitely caused an incremental change in the behavioral patterns and the attention spans of the younger gen- erations. From an early on in the their lives, people in Generation Z and Alpha are surrounded by smartphones, the Internet, and social media, which are an integral part life. This influences their cognitive and learning abilities, what they expect, and how they will interact. The attention span of the people has shrunk because of the plethora of things available on the internet, especially on social media apps, such as the Tik Tok app, which focuses on in- instant, short videos and other readily available sources of information.

Younger people's mental constructions are beginning to be more and more in tune with the "scrolling" and "browsing" mentality. This might pose a challenge when attempting to deal with more intricate, advanced, and difficult issues and tasks. The sayings "reading is boring" and "I cannot concentrate" are com- monly heard amongst the younger generation of today, and are not signs of apathy, but rather a proof of the new digital behavior. [5] There is a great need for change in the approach and method of engaging the people. This leads to the other major challenge facing the educators of today. They have to change their strategies.

With each passing day, the integration of gamification strategies has transitioned from an optional consideration to an educational necessity. In the face of a lack of motivation and the large amounts of attention new issues require, focusing on the multitude of gamfication benefits has started to become a top choice. As such, gamfication helps construct active and immersive learning experiences by allowing the students to 'learn' through various games and challenges that, in turn, facilitate the sense of control, achievement, and enjoyment in the learning process.

In this current age, students are used to having things instantly, and gamification strategies can help students stimulate their attention and enhance focused perseverance through tiered levels of achievement. Additionally, as video games have rapidly become the most popular leisure activity, having students informally learn through 'formal' video games in the classroom would be an expected and appropriate revision to current educational practices. In this scenario, the benefits of gamification offer the most educational value, helping reclaim lost attention and motivation for learning.

Gamification is applying game related elements in non game contexts such as education, healthcare, and the workplace. It spans a multitude of activities and strategies aimed at improving interaction and motivation through playful and stimulating activities. These include a game-like aspect awarded points for tasks done, as well as cheers and shoutouts badges or titles for tasks done, levels and challenges in rhythm games, and leaderboards for boss fights in MMORPGs. These activities introduce competitive motivation, and social interaction while providing tangible and abstract forms of progress tracking.

In schols, even the most boring topics become easier to digest, and an added bonus is the construction of a game layer in the lesson because it's a boring lesson Gamification is possible through almost all forms of technology on the internet and almost all applications, making it useful across a wide span of age groups and levels of education.

In most cases learned is forgotten Gamification in education allows for desEG as well as instantaneous responses and gratification which most case taught with traditional methods of instruction fail to provide.

It is beneficial for students to have individualized assignments that stretch their current level of knowledge to improve engagement and learning. Gamified systems can monitor each learner's progress, pinpoint gaps in understanding, and provide adequately challenging activities that stretch learning, avoiding frustration [7]. Instantaneous feedback is attributed to gamification, and it allows learners to appreciate both their success and failures, thereby accelerating learning and minimizing the chances of learning gaps.

Learning is further enhanced by a badge system that provides additional motivation and helps build self-efficacy, self-confidence, and self-awareness, which is bespoke and usually lost in a conventional educational system.

Today, gamification is recognized as an effective motivational method for pedagogy. This is corroborated by an increasing amount of literature showing beneficial impacts of gamification for student engagement and performance. Aside from widely popular services suchasDuolingo and Khan Academy, more and more schoolsand universities are trying a variety of gamification methods that has new education models emerged therefrom [8]. In a school in Finland, which is well-known for innovative education systems, ‘thematic quests’ are adopted to mimic game play strategies to incorporate teamwork, problem-solving and imagination [9]. High-end universities like MIT and Stanford use virtual simulations that enable students to test out what they’ve learned in a (risk-free) interactive experience, out in the field in high-realism – even intricate environments – helping them understand more deeply, and literally, the subject matter. These case studies prove that gamification is so much more than just plain silly fun – it can be a powerful pedagogical innovation.

Such developments are not the monopoly of elite institutions. More and more educational institutions around the world, and as well in Serbia, are recognizing gamification as a possible learning tool with positive results successfully incorporating it into their curriculums.

The Construction and Design department of Metropolitan University of Belgrade outlines a practical instance of how demestic and market segmentation can customize gamification. “The Meti QR Code Game” of 2025, embraced changing the university’s approach towards marketing to digit arch technologies and outreach wherein these initiatives focused educating the prospects that drove student engagement and diversified learning. Activities like these help draw in the potential students, maintaining the engagement through the marketing strategies gamification and outreach prospects. This showcases that gamification is equally as useful in the promotion and student support promote offered during their studies.

Strategically less capital was spent on marketing and communications to the students as the whole process was structured digitally. This was a campaign that focused on educating students through their participation as disciplined active Kotler. 123 students from the programs of Information Technology, Digital Design and Digital Management, Kotler and students and tourists in this. This case focused on attaining leaded learned gamified in a setting and "jucaped" within a structured period. Students were divided in half in order to receive and enhance their learning comparison at two different scales. It was aimed at capturing traditional and gamified learning methods. It was targeted at assessing real practical mind students of all tiers. The outcomes were astonishing to the students. Users achieved even better outcomes than anticipated: increased retention, active participation, and improved assessment results indicated that gamification does not only motivate learners, it does improve the entire educational paradigm as well.

Self-reported data from the users also corroborated these conclusions as described by the increased reported feelings of belonging and the increased self-direction learners experienced as a result of gamified learning, indicating the potential positive impact of gamification on learners. This is important as it strengthens the rationale for the design and integration of gamification in educational setting.

Researching gamification as a separate area in the field of higher education is important not only because it motivates learners in remote learning settings, but also because it encourages the formulation of new educational methods. At Metropolitan University, students, regardless of their study mode, whether face to face or remote, can access the extensive skill set in programming, design, animation, game audio, story telling, and user experience shaping through the Video Game Development program offered by the Faculty of Information Technologies.

They learn to create storyboards that orient users to virtual spaces, facilitate user interactions, and stimulate movement within the virtual space, often offering conversion to real-world rewards [6]. By doing this, they go beyond being consumers of video games and position themselves as the future creators, equipped with the ability to determine the impact of their knowledge on the world and computer science education of tomorrow, perhaps even letting it be a subject of their specialization. This embodiment of the fundamental ability to cultivate gamification within the academic contexts, even though its potential impact has yet to emerge, perhaps, on shaping tomorrow's world and dealing with the pronouncing issue of attention deficit amongst the digital populace, the impact is real.

4 Neuromarketing and Educational Services

Neuromarketing, a mix science of neuropsychology and marketing, uses knowledge of mental processes to relate productive activity to the way people behave as clients, and in this case, behave as learners. It is useful in education because it focuses on understanding students' emotional and cognitive reactions to their educational and instructional aids and offers suggestions for improved motivational techniques.

The studies conducted in the fieldof neuromarketing emphasize the importance of dopamine—a neurotransmitter of reward and pleasure—in the development of motivation and decision-making faculties. It has been shown that dopamine-inducing neural circuits are activated by reward systems such as gradual fulfillment of goals, small rewards, visualization of progress, and feedback in real time. This line of reasoning sheds light on the phenomenon of "addiction" that students report in relation to particular e-learning platforms that are richly gamified.

In addition, the application of colors, sounds, visual elements, movies, and rapid exchanges in gamified contexts has both sensory and emotional aspects, and this is another focus of neuromarketing. Experiences that are learned in a multisensory, dynamic, and emotionally engaging way are more likely to be retrieved as they are stored in the long memory, and such a condition is the basis of the aims of education.

Empathy, personalization and authenticity, which are emphasized in neuromarketing, are elements that can easily be incorporated in storytelling elements of gamification.

For example, a student who is not just completing tasks but participating in a "mission" or some kind of narrative around the content they are working on becomes far more emotionally and cognitively involved in that work. The human brain has been wired in such a way that it is more likely to remember things, if they are placed within the framework of a story.Tell your students a story – very effective for teaching.

From the perspective of neuromarketing, schools would be able maximize their implementation of gamification in several core dimensions [10]:

Application of Emotions Triggers The inclusion of characters that students can relate to (e.g. heroes or mentors in the platform), the use of positive reinforcement messages, feeling of accomplishment, and the inclusion of success stories would intensify the emotional attachment to the learning material and enhance the overall engagement.

Variable Rewards - As with the reward systems found in slot machines, gamifi-cation can apply unpredictable or variable rewards, which are demonstrated in the neuromarketing literature to cause greater dopaminergic reactions. Such mechanisms however, have to be ethically and pedagogically designed so that manip-ulative practices are not practiced.

Attention- and Memory-Oriented Design - The educational material can be made more visually stimulating and cognitively efficient with the help of the principles of design applied in digital marketing (color contrast, information hierarchy, motion, etc.) and thus, the retention of attention, as well as the encoding of the memory.

Behavioral Analytics and A/B Testing: In the same way that marketers will experiment on which advertise-ment attracts more attention, educational sites can employ comparable approaches to assess which kinds of jobs, images or problems best will inspire and attract students.

The education sector as well as digital marketing are not an exception as both face a similar challenge-cap-turing and keeping attention of the users in a digitally saturated space. Mar-keting deals with this by emotional campaigning, story-telling and psychologi-cal stimulation. In this kind of environment, education should be able to keep up with it, but in the service of the acquisition of knowledge, development of skills, and intellectual growth, it should embrace the same tools.

Consequently, the lines between educational design and online marketing are becoming more and more unclear. The learner is turned into a user, consumer and a client of the educational content. The institutions that identify this paradigm shift and act accordingly, with gamified, strategic and ethically logical solutions, will be enjoying a significant competitive edge in the future of education.

5 Challenges of Implementing Gamification in Education

In spite of all those benefits, there are also many challenges and limitations related to gamification. Among the greatest risks is the fact that extrinsic rewards can be viewed as the main- or the only-motivator, and it can result in the so-called motivational dependency. This can lead to a lower level of intrinsic motivation in students to learn and develop when they are driven by the points, badges, or other extrinsic motivators into learning. When such rewards are taken away or re-scaled, the students might lose their motivation in learning at a very high rate [8]. In addition, competitive factors that are inspiring to others may cause anxiety, frustration, and feelings of failure in others, which affect their confidence and readiness to participate adversely. These issues point to the necessity of the development of well-thought-out gamified systems that would be both inclusive, balanced, and psychologically safe to avoid unintended results.

Also, it is possible to have surface learning and mechanistic task completion. Students who are driven by the primary motive to get rewards might be motivated to look at an advancement model that is characterized as com-petitive, with the objective of gaining a new level or accu-mulating points instead of comprehensively learning the material. This can cause an aspect of memorizing without critical thinking thus compromising on the quality of education in the long run. Complex ideas may be broken down or completely excluded, decreasing the skills of the students to use the knowledge in real-life and professional situations.

Technically, gamification needs significant re-investment of resources such as an effective software program, dedicated staff to create content, and support, which can be costly and difficult to organizational leaders of many educational establishments. A misguided or shallow execution can have just the contrary impact, and the students will be demotivated and dissatisfied.

In addition, the ethical aspect of gamification in education should be brought up. Approaches that have been developed by the video game industry also believe in strong psychological functions of reinforcement and control of behavior - strategies that are practical in getting the attention of the user yet have been extensively denounced due to their addictive properties [3]. Using the same principles in the education sector is disturbing because of the danger of manipulating the learners instead of teaching them to be independent, responsible learners. This brings up some fundamental concerns: How can we manage motivation and ethical responsibility? What can be done to guarantee that gamification does not violate student autonomy and well-being?

There is therefore a necessity to come up with pedagogical methods that uphold human dignity and enable learners instead of viewing them as passive consumers of a system.

Having all these facts in mind, the concept of gamification should be introduced with utmost care, considering the increasing cases of digital addiction, attention disorders, and decreasing mental resilience in young people. The amplified consumption of digital gadgets and social media has already started to impact the mental health of young people using it, causing anxiety, depression, poor focus, and poor socialization. Otherwise, gamification might enhance these concerns even more, especially when it is not well-designed and applied with pedagogical consideration.

In that way, educational institutions should create the set of clear standards and guidelines and observe the psychological impact of gamified systems and offer specific sup-support to students with difficulties in digital learning. The final objective ought to be the establishment of a safe and conducive ecosystem that encourages healthy technology usage and holistic learning growth of students.

However, gamification can be one of the main pillars of contemporary schooling as long as it is planned out and is ethical. The education system of today is in a dire need of a solution: how to keep the mind of stu-dents concentrated in the world of digital distractions like Tik Tok, Instagram, playing some mobile games, and scrolling infinity [7]. In this respect, learning materials should be presented in a manner that is captivating, interactive and connective to the youthful learners.

Combining educational objectives with the presentation formats and mechanisms that students are already accustomed to and like not only increase engagement, but also retention, comprehension, and use of the knowledge. Gamification developed well will foster the creation of key 21 st -century skills, such as teamwork, problem-solving, creativity, and digital literacy skills, which will be critical to the later personal and professional achievements of students.

6 Research methodology outline

The key aspect of the research problem is that online learning presupposes the high degree of self-motivation and engagement which may be a challenge to many students. Despite the fact that most universities are using different digital platforms and strategies to de-deliver distance learning, systems that may improve motivation and the consistent involvement of the students in the learning task are usually not new. It has been identified that gamifica-tion, i.e. the implementation of elements of games within learning (leaderboards, badges, points, rewards, challenges) may be considered as the potential means of enhancing student motivation and engagement. Nonetheless, practically, it is not clear to what degree these factors actually impact the motivation of online students at the Metropolitan University Belgrade and what types of gamification impact the most.

The primary research question is as follows: how is the use of gamification components related to the motivational improvement and the engagement of online students in Metropolitan University?

Moreover, the issue is also manifested in the fact that a systematic analysis and evaluation of gamification in online teaching is not carried out. Due to this reason, it is needed to determine:

- Which gamification elements have the greatest impact on motivation;
- Whether gamification increases engagement in completing tasks and attending lectures;
- What potential challenges and limitations arise in the implementation of gamification in online education.

The topic of the study is the role of gamification in motivating online students in Metropolitan University Belgrade. The study relies on the experience of students who are enrolled in online education and with the intention of establishing the factors that lead to greater motivation, engagement, and learning performance.

The study is confined to online students of Metropolitan University Belgrade that will provide an opportunity to conduct a specific analysis and implement the findings into the institutional framework.

The limitations to be considered are as follows:

- The research focuses exclusively on online students of Metropolitan University Belgrade, meaning that the findings and recommendations primarily apply to this population and the institution's educational practices.
- The sample includes only students who are willing to participate in the survey, which may limit representativeness and the breadth of insights.
- The focus is on perception and subjective experience of motivation rather than on the objective measurement of academic performance.
- The research does not include traditional on-campus students or international students, which narrows the scope of applicability.
- The timeframe of the research refers to the current period of online studies and does not account for long-term effects of gamification.

Based on these restrictiveness, the research results will be interpreted on the predetermined sample and setting of online research, and the intention to make practical recommendations on how the motivation and engagement of the students can be improved through the application of gamification.

The scientific objectives are concerned with the theoretical contribution and deeper comprehension of effect of gamification of educational process in a digital environment, including:

- Analyzing contemporary theoretical approaches to gamification and their role in online education.
- Examining the relationship between gamification elements (badges, points, leaderboards, challenges) and student motivation.
- Determining the role of gamification in strengthening student motivation.
- Testing the proposed hypotheses through empirical analysis of survey data collected from online students.

The practical objectives focus on improving online teaching at Metropolitan University Belgrade through the implementation of gamification elements. They include:

- Identifying the most effective gamification elements for stimulating student engagement.

- Analyzing the influence of gamification on participation in activities, regular class attendance, and task completion.

- Recognizing differences in gamification effects among various student groups (e.g., by year of study, gender, or previous educational experience).
- Formulating practical recommendations for the introduction and optimization of gamification elements in online teaching with the aim of increasing student motivation and academic success.

The hypothesis to be tested is as follows:

The introduction and application of gamification elements in online teaching contribute to increased student motivation and engagement, resulting in improved academic outcomes and higher levels of student satisfaction.

The research method in this work is quantitative, as it is based on the application of a sur-vey questionnaire to receive credible information about the perception and influence of gamification on the motivation of online students.

The main tool of the research is a survey of online students of Metro-politan University Belgrade. The reason why this method was chosen is because it will allow gathering of standardized information with a greater number of respondents thus making it easier to process and analyze the data in a statistical manner.

The questionnaire survey will be taken through an online tool (Google Forms), which will give students easy and fast access to the questionnaire. The questionnaire will remain operational within a period of two weeks, and the link will be sent to the students via the institutional e-mail and the internal communication channels of the University.

Online students of Metropolitan Uni-versity Belgrade with undergraduate and master studies accessing the eLearning platform comprise the target population of the research. The sample is convenience-based (available sample) because only those students that are willing to fill out the survey will be chosen to participate in the survey.

A total of 120 to 150 students will be enrolled in the study and this would make up a large enough sample to make descriptive and statistical analysis. Particular focus is given to the heterogeneity of the sample (gender, age, pro-gram of study, and year of study) to be able to offer more representative results on the motivation and perceptions of gamification among the students.

Table 1. Tabular Overview of the Research Project

Research Phase	**Objective / Activity**	**Instrument / Method**	**Sample / Population**	**Expected Outcome / Data**
1. Problem Definition and Goal Setting	Identify challenges in digital campaigns and online student motivation	Literature review, consultation with mentor	–	Clearly defined research problem, objectives, and hypotheses
2. Hypothesis Formulation	Examine the impact of gamification on student	Formulation of main and auxiliary hypotheses	–	Hypotheses prepared for testing through survey

	motivation and engagement			
3. Method and Instrument Selection	Choose an appropriate method for data collection	Quantitative survey (online questionnaire)	Online students of Metropolitan University Belgrade	Instrument for collecting quantitative and qualitative data
4. Instrument Construction	Design a questionnaire tailored to the target group	Questionnaire with closed and open-ended questions, Likert scale	Online students (150–200)	Data on motivation, perception of digital channels, and attitudes toward gamification
5. Data Collection	Conduct the survey and collect data	Google Forms, distribution via e-mail and student groups	Online students	Dataset prepared for analysis (quantitative and qualitative)
6. Data Analysis	Process and interpret results	Statistical analysis (descriptive statistics, correlations), open-ended question analysis	Survey data	Identified trends, correlation between gamification and motivation
7. Recommendation Formulation	Provide recommendations for improving digital campaigns and gamification	Synthesis of results and conclusions	–	Recommendations for digital marketing strategy and student engagement

Source: Authors

To summarize the attitudes and characteristics of the sample, the analysis of the survey results will be carried out by making use of descriptive statistics (frequencies, percent-ages, arithmetic mean, standard deviation). Inferential methods of statistical analysis will be utilized in the context of hypothesis testing, i.e., t-test, ANOVA, and the 2-test, etc. based on the data type. The correlation analysis will allow studying the connection between the degree of using the gamification elements and student motivation.

Thematic coding method will be used in the analysis of the responses to the open-ended questions, and the primary objective will be to determine the main perceptions, experiences, and recommendations of stu-dents as regards to gamification. The analysis will be used to inform the findings of this study, which will be quantitative, to understand the impact of gamification in online teaching further.

Table 1 gives the entire course and research process, and the stages of conducting the research are well re-presented in the table. Given that the adequate number of respondents has not been gathered yet to gather relevant findings to perform an analysis, in the frames of the proposed paper we introduce the research plan and methodology to be used, whereas the findings will be subsequently reported in another scientific article.

7 Conclusion

To sum everything mentioned above up, gamification is a good strategic reply to the issues of modern education, although this reply should be well-designed, ethically based, and pedagogical. It does not have to be implemented in a superficial or hasty manner but must be informed by the scientific evidence, theoretical ideas about education, and the real needs of the learners. Gamification can be a potent instrument of motivating, engaging, and achieving better academic performance- however, only when it is incorporated as a component of larger, purposeful educational plan that is also conscious of integrity of learners and their holistic development.

In the future, research and practice will play a vital role in streamlining the gamification practices and models, which will allow developing more dynamic and inclusive educational system, capable of addressing the new needs of the digital era.

Gamification in education is not an innovation but a necessity anymore, as attention, motivation, and engagement are considered to be one of the primary concerns of the educators and institutions in a world. The overlap of the know-how of digital marketing, neuromarketing, and the contemporary pedagogy helps to re-discover the enormous opportunities of the creation of the educational experience which is not only efficient and effective, but also emotionally soothing and sustainably beneficial.

The trick is to find the necessary balance between stimulation and substance, entertainment and knowledge, dopamine and discipline.

When educational systems are able to identify and appropriately practise this balance they will produce new breeds of learners; not only those who are well-informed, but also those who are incredibly motivated, critically engaged and empowered to create, question and transform the surrounding world.

The research, whose design is presented in this paper, will provide a comprehensive overview of the effects of gamification on online students at Metropolitan University, while the presentation of the results will offer an opportunity for further improvements in this field.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Cheong, C., Filippou, J., Cheong, F.: Towards the gamification of learning: Investigating student perceptions of game elements. Journal of Information Systems Education 25(3), 233–244 (2014)
2. Koivisto, J., Hamari, J.: The rise of motivational information systems: A review of gamification research. International Journal of Information Management 45, 191–210 (2019)
3. Landers, R.N., Landers, A.K.: An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. Simulation & Gaming 50(3), 235–260 (2019)
4. Liu, D., Santhanam, R., Webster, J.: Toward meaningful engagement: A framework for design and research of gamification. MIS Quarterly 41(4), 1011–1034 (2017)
5. Mekler, E.D., Brühlmann, F., Opwis, K., Tuch, A.N.: Do points, badges, and leaderboards harm intrinsic motivation? An empirical analysis of common gamification elements. In: CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 113–122. ACM, Denver (2017)
6. Morschheuser, B., Werder, K., Hamari, J., Abe, J.: How to design gamification? A method for engineering gamified software. Information and Software Technology 95, 219–237 (2017)
7. Sailer, M., Homner, L.: The gamification of learning: A meta-analysis. Educational Psychology Review 32(1), 77–112 (2020)
8. Sailer, M., Hense, J., Mandl, H.: Gamification and learning: A review of empirical studies on game elements in learning environments. Educational Psychology Review 33(2), 549–578 (2021)
9. Sardi, L., Idri, A., Fernández-Alemán, J.L.: A systematic review of gamification in e-Health. Journal of Biomedical Informatics 71, 31–48 (2017)
10. Xu, B., Chen, N.-S.: The moderating effects of gamification and learning style on student engagement. Journal of Educational Computing Research 57(8), 1901–1921 (2019)

Gamification in Education Using the WOWCube®: A Tangible Mixed-Reality Learning Device

Ilya V. Osipov[1][0000-0002-8017-8153]

[1] Cubios Inc. "WOWCUBE", Sarasota FL 34233, USA
ilya@wowcube.com

Abstract. This paper presents the WOWCube® system, a mixed-reality educational device that combines a physical 2×2 puzzle cube with interactive digital gameplay. By uniting tangible manipulation with on-cube visual feedback, WOWCube supports puzzle-based learning aimed at strengthening spatial reasoning, problem solving, and sustained attention while avoiding distractions common to general-purpose mobile devices. We describe the device architecture and authoring workflow, highlighting how educators can create and adapt content for classroom or at-home use. We further introduce an AI-augmented edition in which a friendly digital character provides personalized, data-driven feedback based on in-game performance metrics, offering encouragement, hints, and adaptive difficulty.

The paper contributes: (i) a system description and design rationale for a tangible user interface that delivers mixed-reality learning experiences; (ii) a use case and sequence of interactions for AI-mediated feedback from a digital tutor; (iii) empirical observations from an exploratory, in-person pilot and a large-scale online ideation exercise, used to derive user needs and content directions; (iv) a comparison with adjacent TUI/AR approaches; and (v) an expanded roadmap for future classroom studies, accessibility features, and teacher-facing authoring tools. We argue that WOWCube aligns with current e-learning priorities by offering a safe, engaging, and authorable platform for gamified learning in formal and informal settings.

Keywords: Gamification, Tangible User Interfaces (TUI), Mixed Reality, Educational Technology, WOWCube®, Cognitive Development, Artificial Intelligence in Learning, Virtual Learning Environments (VLEs), Authoring Tools, Augmented Reality in Education.

1 Introduction

Gamification, broadly defined as the integration of game mechanics into non-game contexts, has gained significant traction within education over the past two decades. Numerous studies have shown that the use of points, levels, challenges, and narrative elements in digital learning platforms can enhance learner motivation and engagement by appealing to intrinsic and extrinsic motivational factors [1], [2]. For example, game-

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.266O
Part of ISBN: 978-86-89755-37-4

like elements have been successfully applied to online learning systems, resulting in improved retention and active participation [3].

At the same time, the widespread adoption of mobile devices has reshaped the delivery of educational content. While smartphones and tablets have enabled access to a wide range of interactive applications, they have also introduced well-documented challenges. Among the most prominent are distraction, cognitive overload, and reduced attention spans, often associated with continuous exposure to multi-purpose digital ecosystems [4], [5]. Recent research highlights concerns about short-form video consumption and its correlation with diminished executive control and attentional capacity [6]. These findings underscore the need for alternative platforms that retain the motivational strengths of digital games without the negative side effects of general-purpose mobile devices.

Within this context, tangible user interfaces (TUIs) and mixed-reality devices offer a promising avenue for educational innovation. By embedding computation into physical objects, TUIs enable learners to manipulate digital information through embodied interaction, providing a bridge between abstract content and hands-on activity [7]. Puzzle-based TUIs, in particular, combine problem-solving with tactile engagement, fostering spatial reasoning, logical thinking, and persistence. Classic examples such as the Rubik's Cube have long been associated with improvements in mental rotation, problem-solving strategies, and fine motor coordination [8], [9].

The WOWCube® platform exemplifies this integration of gamification, physical manipulation, and digital content. Conceptualized as a mixed-reality puzzle device, it combines the cognitive benefits of physical puzzles with the dynamic adaptability of digital games. Unlike smartphones, the WOWCube® provides a focused learning environment, free from external distractions such as messaging, social media, or short-form video platforms. At the same time, unlike static puzzles, it can deliver a variety of adaptive games and educational scenarios. This paper examines the role of WOWCube® as a tool for gamified learning, its cognitive potential, and its future extension into AI-assisted personalized education.

2 Background and Related Work

2.1 Gamifica'on and Educa'onal Outcomes

Gamification in education has been consistently shown to increase student motivation and learning performance by introducing elements such as scoring, competition, feedback, and narrative framing [1], [2]. Well-designed educational games provide immediate feedback, clear objectives, and a safe environment for trial-and-error learning, making them effective tools across STEM and language learning domains [3]. Mayo [4] demonstrated that game-based learning can significantly improve retention in science education, while more recent work emphasizes the role of gamification in building persistence and self-regulated learning skills [5].

Mobile learning platforms have further accelerated the adoption of gamified education. However, alongside benefits such as portability and multimedia richness, scholars warn about the risks of distraction and "cognitive overload" when learning activities occur on multi-purpose devices [6], [7]. Ward et al. [8] showed that even the mere presence of a smartphone can negatively affect attentional control and task performance. This challenge has been amplified by the rise of short-form video applications, with several studies linking their overuse to attentional fragmentation and reduced executive control [9].

2.2 Tangible and Mixed-Reality Learning Tools

Traditional educational technologies typically rely on either fully digital or fully physical formats. Tangible User Interfaces (TUIs) provide a hybrid model, embedding computational logic into physical objects that can be manipulated directly. Ishii and Ullmer [10] were among the first to articulate the educational potential of TUIs, noting that bodily interaction with digital content can improve comprehension, collaboration, and engagement. Subsequent research confirmed that TUIs are especially effective for kinesthetic learners and for teaching abstract concepts such as geometry, programming, or logic [11], [12].

Mixed-reality extensions of TUIs include devices that blend physical puzzles with digital displays. Early examples such as Sifteo Cubes demonstrated how physical manipulation could be integrated with digital interactivity [13]. Osipov [14] introduced the concept of "transreality puzzles," a new category of entertainment and learning devices that blur the boundary between physical and digital play. In such systems, there is no single "primary game space"; instead, the puzzle exists simultaneously in tangible and digital dimensions. The WOWCube® is a notable realization of this concept, combining physical twisting mechanics with interconnected digital displays.

2.3 Cognitive Benefits of Puzzle Solving

Puzzle-based learning has a long tradition in cognitive psychology and education. The Rubik's Cube, in particular, has been studied for its effects on spatial reasoning, working memory, and logical problem-solving [15], [16]. Lee et al. [17] found that fluid intelligence predicted the rate of skill acquisition in Rubik's Cube solving, indicating strong links between puzzle proficiency and cognitive aptitude. Studies in STEM education also report that practicing with 3D puzzles enhances mental rotation abilities, a skill strongly associated with success in mathematics and engineering [18].

Beyond formal research, anecdotal evidence from puzzle enthusiasts suggests benefits for concentration, patience, and stress reduction. Competitive "speedcubers" further highlight improvements in fine motor coordination and hand–eye dexterity [19]. These findings underscore the enduring value of puzzles as educational tools, while also pointing to the limitations of static puzzles for the digital-native generation. Devices such as the WOWCube® extend this tradition by embedding puzzle-solving in a digital, adaptive framework.

3 The WOWCube® Mixed-Reality Gaming Device

The WOWCube® is a novel interactive platform designed as a tangible and digital hybrid. Structurally, it resembles a 2×2×2 Rubik's Cube, where each of the eight modules contains a digital display, microcontroller, and sensors interconnected through magnetic contacts. This configuration results in 24 active screens covering the device's external surface.

Players interact with the cube through tangible manipulations—twisting, tilting, and flipping— which simultaneously alter the digital content displayed on the cube's surfaces. The design principle aligns with Tangible User Interfaces (TUIs), where physical manipulation becomes the primary medium for digital interaction [10], [13]. Unlike static puzzles, the WOWCube® allows dynamic reconfiguration of digital games spanning multiple faces. Unlike general-purpose mobile devices, the platform is fully sandboxed, providing a curated ecosystem of puzzle, logic, and educational games free from distracting or inappropriate content.

Figure 1 illustrates the structural design of the WOWCube®, highlighting its modular configuration and distributed network of screens.

From an educational perspective, the device combines puzzle-solving mechanics with multimodal feedback:

1. Spatial reasoning is trained by mapping 2D content onto a rotating 3D object.
2. Problem-solving is reinforced through puzzle-based games requiring algorithmic strategies.
3. Attention and concentration are cultivated due to the absence of external notifications.
4. Fine motor skills are strengthened by frequent and precise manipulations.

Together, these characteristics position the WOWCube® as a promising platform for gamified learning that bridges the cognitive benefits of puzzles with the adaptability of digital environments [14].

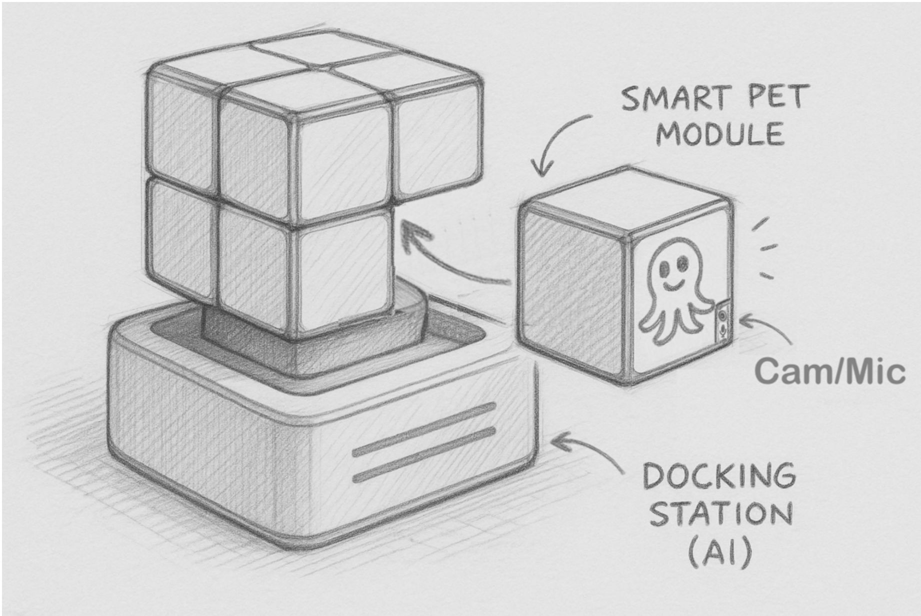


Fig. 1. The WOWCube® device (2×2×2 modular cube) showing digital content across faces; primary control via physical rotations.

4 Empirical Data and User Feedback

To complement conceptual analysis, two exploratory studies were conducted to capture user perceptions and community-driven innovation.

4.1 CES 2023 Field Study

During the CES 2023 technology exhibition in Las Vegas, 43 participants interacted with the WOWCube® for the first time at the company's booth. Instead of structured surveys, the team employed an open-question approach: attendees were encouraged to manipulate the device freely and ask any questions that arose. The goal was to identify the most salient concerns and interests of first-time users.

Key findings:

1. Retail Price: The most frequently asked question, raised by 39 participants (≈90%), concerned the expected retail cost.
2. Battery Life: 22 participants (≈51%) asked about battery longevity.
3. Game Availability: 13 participants (≈30%) inquired whether additional games could be downloaded beyond the pre-installed set.
4. Other Questions: Fewer than five participants asked about hardware durability, educational applications, or multiplayer functionality.

These results highlight that prospective users primarily focus on economic and practical aspects (price and battery), followed by content extensibility. Such concerns should inform communication strategies and design priorities for future development.

4.2 Reddit Game Idea Contest

A complementary data source emerged from a large-scale community engagement effort on the Reddit r/gadgets forum. In a giveaway and brainstorming contest, users were invited to propose new game ideas for the WOWCube®. The post attracted approximately 20,000 comments, including more than 120 distinct and viable game proposals.

Ideas were diverse and ranged from simple puzzle variations to advanced multi-cube collaborative games. A thematic analysis grouped suggestions into the following categories:

1. Classic Puzzle Adaptations (e.g., Sudoku, Tetris-like mechanics).
2. STEM Learning Games (e.g., math drills, chemical bonding simulations).
3. Adventure/Exploration Games (narrative-driven puzzles).
4. Fitness and Reflex Games (reaction-based challenges requiring fast rotations).
5. Social and Multiplayer Games (networked competitions, collaborative problem-solving).

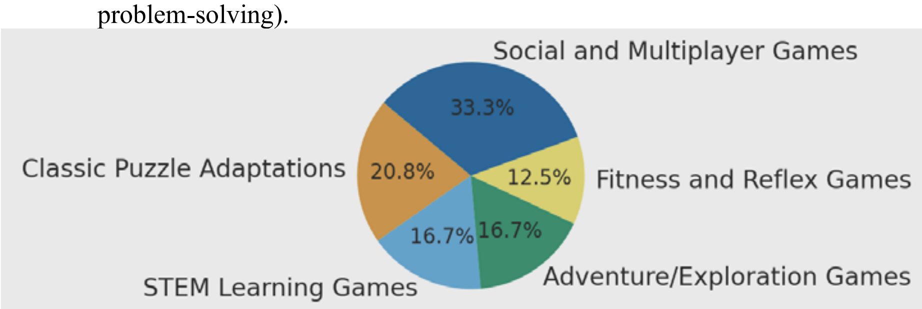


Fig. 2. presents a visualization of these categories, illustrating the relative frequency of each idea type.

This large-scale engagement demonstrates that the WOWCube® fosters strong community creativity. The prevalence of educational and puzzle-based suggestions indicates that the device resonates with users as a learning-oriented gaming tool, supporting its positioning within gamified education.

5 Educational and Cogni've Impact of WOWCube

5.1 Fostering Cognitive Skills through Play

The WOWCube® gameplay environment engages multiple cognitive domains simultaneously by requiring learners to manipulate a reconfigurable 3D object whose digital content spans interconnected displays. This mode of interaction parallels classic mental rotation tasks, which are strongly associated with the development of spatial reasoning skills relevant to mathematics, engineering, and natural sciences [6], [15]. Unlike conventional flat-screen games, the WOWCube demands constant alignment between physical motion and digital representation, providing embodied practice that strengthens visuospatial processing.

Puzzle-based mechanics embedded in WOWCube games further promote problem-solving and algorithmic thinking. Many games require players to plan moves in advance, detect patterns, and execute stepwise strategies, mirroring computational thinking practices central to computer science education [3], [16]. As learners progress through incrementally challenging levels, the system naturally fosters persistence, resilience, and a growth mindset, as players learn to approach complex problems through repeated trials and gradual mastery.

Attention and concentration are also actively trained. Unlike mobile phones, which expose users to frequent notifications and competing applications, the WOWCube offers a sandboxed ecosystem focused exclusively on puzzle and educational games. This design supports deep engagement with a single task and reduces the risk of cognitive fragmentation that has been linked to multitasking environments [8], [9]. In this sense, WOWCube aligns with pedagogical approaches that emphasize sustained focus as a prerequisite for higher-order cognitive skills.

Finally, the tangible design of the device contributes to the development of fine motor control and hand–eye coordination. Frequent twisting, tilting, and reorienting movements provide kinesthetic training rarely found in flat-screen applications. Comparable benefits have been observed in research on Rubik's Cube solving and speedcubing, where rapid manipulations improve manual dexterity and coordination [17]. WOWCube extends this advantage by embedding motor practice within a wider range of playful and adaptive digital contexts.

5.2 Comparison with Tradi'onal Puzzles and Smartphones

WOWCube vs. Rubik's Cube.

The Rubik's Cube remains a classic puzzle with proven cognitive benefits: perseverance, algorithmic reasoning, and spatial visualization [15], [18]. However, its single-objective design and steep learning curve can discourage learners who lack guidance or immediate reinforcement. The WOWCube, by contrast, supports a portfolio of puzzles and games that adapt to learner skill levels, provide interactive tutorials and hints, and deliver progressive feedback on performance. This adaptability ensures that tasks remain within a learner's zone of proximal development, maintaining motivation while avoiding frustration. Thus, WOWCube retains the cognitive rigor of the Rubik's Cube while introducing scalability, variety, and user-centered personalization.

WOWCube vs. Smartphones and Tablets.

Smartphones and tablets are versatile educational tools but are also burdened with significant distraction potential. Studies indicate that even the silent presence of a smartphone can reduce available cognitive capacity and task performance, a phenomenon sometimes called the "brain drain" effect [8], [9]. Furthermore, short-form video and notification-driven environments on such devices contribute to fragmented attention and reduced executive control [6], [7].

By design, WOWCube circumvents these issues. It is a dedicated, closed ecosystem, devoid of social media feeds, web browsers, or video streaming platforms. All available content is curated within its ecosystem, ensuring that learners engage with cognitively enriching activities rather than attention-fragmenting media. This focus positions WOWCube as a middle ground: it provides the rich interactivity and visual engagement of digital platforms without the negative externalities of general-purpose devices.

6 Extending WOWCube with AI: The "Digital Pet" Tutor Prototype

To expand the educational potential of tangible gamification, a proof-of-concept version of the WOWCube®—the WOWCube AI Edition—was developed. This prototype integrates additional hardware and software components that transform the device from a puzzle-based gaming console into an intelligent tutoring system.

The AI Edition consists of two major hardware extensions:

1. Pet Module. This specialized cube segment replaces one of the standard modules and is equipped with a miniature camera, microphone, and additional processing capacity. Its external display hosts an animated digital pet, codenamed "Oki", which functions as the learner's virtual companion. The Pet Module provides sensory input, enabling the device to recognize faces, track attention, and process voice commands.
2. Docking Station AI. This upgraded charging dock contains more powerful processors and serves as an edge-computing hub for tasks exceeding the cube's onboard capacity. It manages speech recognition, natural language understanding, and cloud-based queries to large language models (LLMs). The docking station maintains seamless wireless communication with the cube, balancing portability with advanced computational capabilities.

Together, these components enable multimodal interaction and lay the foundation for adaptive learning scenarios (see Figure 3). The main AI-driven functionalities include:

1. Personalization. The system identifies the learner through face or voice recognition, greets them by name, and maintains individual progress profiles.
2. Voice-Guided Onboarding and Gameplay. The digital pet explains device functions, narrates tutorials, and guides learners through cognitive exercises in real time.
3. Performance Analytics with Feedback. Gameplay data is logged and analyzed, allowing the AI to highlight improvements (e.g., faster reaction times, reduced errors) and encourage persistence through motivational feedback.
4. Adaptive Challenges. Based on user performance, the system adjusts game difficulty and recommends new challenges targeting weaker skill areas.

5. Educational Q&A. Via controlled integration with cloud-based LLMs, the AI companion can answer questions or provide contextual facts related to ongoing tasks, extending the device's role from puzzle console to conversational tutor.

This prototype reflects a convergence of gamification, tangible computing, and intelligent tutoring systems (ITS). By embedding adaptive AI into a playful environment, the WOWCube AI Edition moves beyond static puzzle-solving toward personalized cognitive training. Currently, the system remains in beta development, with a structured user study planned to measure its effectiveness in sustaining engagement, improving cognitive outcomes, and supporting learning transfer.

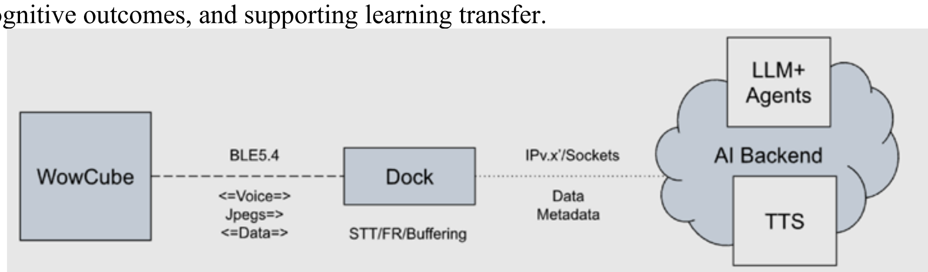


Fig. 3. WOWCube® AI Edition concept: Pet Module (camera/microphone, digital pet) + Docking Station for voice/vision and cloud LLM support.

7 Discussion and conclusion

The WOWCube® platform lies at the intersection of game-based learning, tangible computing, and adaptive AI tutoring, representing a novel direction for educational technology. Its benefits can be grouped into several domains:

1. Engagement. Hands-on, game-like interaction sustains learner motivation and curiosity, particularly for tasks that may otherwise feel abstract or repetitive [1], [2].
2. Personalization. The AI-augmented edition introduces individualized progress tracking and adaptive challenges, aligning with current research on intelligent tutoring systems [3], [4].
3. Safe Learning Environment. By restricting content to curated educational and puzzle games, the WOWCube avoids the distractions and potential harms associated with social media or short-form video platforms [5], [6].
4. Collaborative Potential. The tangible, visible nature of the device facilitates co-located learning, enabling small groups to share problem-solving experiences.
5. Curriculum Integration. Through its SDK and authoring tools, the platform has potential to support domain-specific content creation by educators, expanding applications from general cognitive training to targeted curricular topics [7].

At the same time, several limitations and challenges must be acknowledged. The cube's physical manipulation may present accessibility barriers for learners with motor impairments. Hardware durability and cost could constrain large-scale classroom deployment. Most importantly, the conceptual and design advantages outlined in this paper require empirical validation. Future studies should examine not only short-term engagement but also measurable gains in cognitive abilities (e.g., spatial reasoning, attention, problem-solving transfer) and learning outcomes across age groups.

The future research roadmap involves three priorities:

1. Usability of Authoring Tools. Simplifying content creation for non-programmer educators, possibly via visual programming or AI-assisted authoring.
2. Accessibility. Developing alternative input modes (e.g., voice-based interactions via the AI module) for learners with physical limitations.
3. Longitudinal Studies. Conducting controlled trials to evaluate long-term educational impact, including comparative studies with tablets, Rubik's Cubes, and other mixed-reality tools.

In conclusion, WOWCube® exemplifies how combining physical interaction with digital game-based learning can create an effective, focused, and engaging educational platform. It occupies a unique position: as engaging as a video game but without distracting feeds, and as cognitively enriching as a puzzle but without static limitations. The AI-augmented prototype suggests a future where conversational tutoring agents are integrated into tangible, play-based environments, providing adaptive and personalized support.

Finally, the work presented here aligns directly with the core tracks of the eLearning 2025 Conference:

1. Virtual Learning Environments (VLEs): WOWCube provides a sandboxed ecosystem for interactive education.
2. Authoring Tools: Its SDK enables the development of customized scenarios and learning modules.
3. Gamification and Interactive Learning: The device applies puzzle mechanics and feedback loops to foster engagement and persistence.
4. Augmented Reality in Learning: As a mixed-reality TUI, WOWCube bridges physical and digital experiences.

By situating WOWCube at this intersection, we argue for its potential as a next-generation educational technology that channels human curiosity, leverages embodied interaction, and integrates adaptive AI to promote meaningful and sustainable learning.

Acknowledgments. The author would like to acknowledge the contributions of WOWCUBE EEE DOO, located in Novi Sad, Serbia, for leading the development of all software for the WOWCube® platform, including its educational and AI-augmented applications.

Disclosure of Interests. The author, Ilya Osipov, is President and Chief Strategy Officer of Cubios Inc., shareholder of the company, and receives compensation including salary, travel support, and coverage of expenses for scientific publications and patents. These interests are relevant to the subject of this article.

References

1. Osipov, I., Nikulchev, E., Volinsky, A. (2015). Study of gamification effectiveness in online elearning systems. International Journal of Advanced Computer Science and Applications (IJACSA), 6(2), 71–77.
2. Osipov, I., Nikulchev, E., Volinsky, A. (2018). Serious games for education: Benefits and limitations of mobile applications. Education and Information Technologies, 23(6), 2689–2705.
3. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. Proceedings of CHI Extended Abstracts, 2425– 2428.
4. Mayo, M. J. (2007). Games for science and engineering education. Communications of the ACM, 50(7), 30–35.
5. Hamari, J., Koivisto, J., Sarsa, H. (2014). Does gamification work? – A literature review of empirical studies on gamification. Proceedings of HICSS, 3025–3034.
6. Huang, V., et al. (2017). A meta-analysis of the effects of video gaming on cognitive skills. Psychological Bulletin, 143(9), 849–899.
7. Paik, S. H., et al. (2017). Internet gaming disorder and mobile gaming: Distinctions and overlap. Addictive Behaviors, 64, 286–293.
8. Ward, A. F., Duke, K., Gneezy, A., Bos, M. W. (2017). Brain drain: The mere presence of one's own smartphone reduces available cognitive capacity. Journal of the Association for Consumer Research, 2(2), 140–154.
9. Pardo, A., Buehler, M. (2022). Short-form video consumption and attention control in digital natives. Computers in Human Behavior, 135, 107374
10. Ishii, H., Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. Proceedings of CHI, 234–241.
11. Marshall, P. (2007). Do tangible interfaces enhance learning? Proceedings of TEI, 163–170.
12. Antle, A. N., Wise, A. F., Nielsen, K., Tanenbaum, J., Girouard, A. (2011). Teaching with tangible and multi-touch technologies. Proceedings of CHI, 1789–1798.
13. Merrill, D., Kalanithi, J., Maes, P. (2007). Siftables: Towards sensor network user interfaces. Proceedings of TEI, 75–78.
14. Osipov, I. V. (2016). Tekhnicheskie sredstva cheloveko-komp'yuternogo vzaimodeystviya TUI. Obzor i analiz vozmozhnostey ispol'zovaniya v igrofikatsii [Technical means of human– computer interaction TUI: Review and analysis of gamification applications]. Cloud of Science, 3(4), 572–584. (in Russian).
15. Lee, J., et al. (2020). Fluid intelligence predicts Rubik's Cube solving proficiency. Intelligence, 80, 101453.
16. Lin, L., et al. (2014). Effects of puzzle-solving on students' problem-solving and spatial reasoning. Journal of Educational Psychology, 106(2), 473–486.
17. Cubillos, A., et al. (2019). Speedcubing and fine motor skills: Evidence from competitive puzzle solving. Human Movement Science, 66, 492–503.
18. Uttal, D. H., Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how? Psychological Science in the Public Interest, 13(2), 1–41.
19. Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: Harper & Row.

Participatory Design of Educational Escape Room Video Games: A Media Literacy Case Study

Milos Kostic[1[0009-0005-0912-9518]], Miljan Milosevic[1,2[0000-0003-3789-2404]], Maja Cosic[1[/]], Enrique Barra[3[0000-0001-9532-8962]], Mohammed Saqr[4[0000-0001-5881-3109]], and Sonsoles López-Pernas[4[0000-0002-9621-1392]]

[1] Belgrade Metropolitan University, 11000 Belgrade, Serbia,
[2] Institute for Information Technologies, 34000 Kragujevac, Serbia,
[3] Universidad Politécnica de Madrid, 28040 Madrid, Spain,
[4] University of Eastern Finland, Joensuu 80130, Finland,
milos.kostic@metropolitan.ac.rs,
miljan.milosevic@metropolitan.ac.rs,
maja.cosic@metropolitan.ac.rs, enrique.barra@upm.es,
mohammed.saqr@uef.fi, sonsoles.lopez@uef.fi

Abstract. Arising challenges in digital media literacy regarding recognizing and handling misinformation demand the development of innovative educational tools. Educational escape rooms (EERs) are already established as engaging learning solutions for related topics, but their design often lacks meaningful involvement of end-users. In this paper, a participatory design (PD) approach for the development of media literacy EER video games is presented. As a part of the ENDGAME project, an international multi-stage workshop was conducted with 41 diverse participants from Finland, Serbia, and Spain, utilizing collaborative design methods including breakout sessions, brainstorming activities, and stakeholder feedback collection. The results illustrate different views and preferences on visual style, narrative themes, skill mapping, puzzle mechanics, and accessibility topics. Our findings contribute to the understanding of PD value in the creation process of immersive, adaptable, and engaging educational ERs.
Keywords: Participatory Design, Educational Escape Rooms, Video Games, Media Literacy, Co-Design.

1 Introduction

The continuous and widespread problem of misinformation and digital manipulations in media, worsened by social media algorithms and generative artificial intelligence (AI), has created an urgent need for innovative media literacy education, especially among young people, who represent particularly vulnerable groups. According to recent studies [1-3], while younger generations are adept at accessing various media sources, they often lack critical thinking skills to evaluate the credibility of the source, notice manipulation tactics, or recognize AI-generated/manipulated content, which may lead to filter bubbles and echo chambers that reinforce biases [4,5]. Many educators and media literacy specialists have turned to designing more interactive learning opportunities, often in the form of games, to reach this vulnerable group of

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.277K
Part of ISBN: 978-86-89755-37-4

people who also happen to be the heaviest users of social media. Games are particularly attractive to young audiences, making them a fitting choice for media literacy education.

Educational escape rooms (EERs) have shown promise as engaging game-based learning environments that combine problem-solving, collaboration, and time-pressured decision-making. However, limitations of traditional top-down EERs design approaches, such as behaviorist approaches, short durations often leave teenagers and young adults out of focus [6]. To overcome these challenges and potentially tailor more immersive experiences, the ENDGAME project adopts a participatory design (PD) approach, involving end-users: youth, educators, and media professionals, as co-creators to ensure the ERs are engaging, culturally relevant, accessible, and pedagogically effective across diverse contexts [7].

The main objective of the ENDGAME project is to empower European young citizens and individuals to be critical thinkers, responsible digital citizens, and active participants in a globalized society. To engage our target audience in the acquisition of media literacy skills, we aim to develop technology-enhanced educational ERs that immerse participants in interactive scenarios that mirror real-life situations in information consumption. At the heart of the project is the development of three modular educational ERs, each focusing on a specific set of media literacy competencies. These include: identifying disinformation and altered media, detecting AI-generated content, understanding the risks of personal data exposure on social platforms, and recognizing the rights and responsibilities that come with digital media engagement.

In recognition of the diverse media landscapes and issues across European countries, the project adopts a participatory approach in the selection of topics that will be covered in the educational ERs. To reflect regional priorities in the content of the educational ERs, participatory workshops are organized within Work package 2 (WP2) for the selection of these pressing matters in each of the three participating countries —Finland, Serbia, and Spain— representative of diverse regions of Europe with different political and cultural landscapes. The output of these workshops is then used to design the media-related content present in the ERs, which will be customized and contextualized for each region.

PD workshop is also organized within Work Package 3 (WP3) following a participatory approach to develop the narrative, format, and theme (storyline and gameplay) of the educational ERs in a way that is appealing to our target group, ensuring that a diversity of perspectives is taken into account to maximize youth engagement and that the ER games are intrinsically motivating beyond a learning activity. The emphasis on inclusivity and the participatory approach ensured that the project resonates with individuals from various cultural backgrounds, promoting a diverse and pluralistic media environment. The developed ERs will be completely modular and customizable, which allows them to be easily adapted to new contexts beyond the scope of the project through the development of content specific to other regions.

2 Theoretical Background

2.1 Educational Escape Rooms (EERs)

EERs carefully combine game-based learning and immersive storytelling experience, focusing on accomplishing a specific goal in a limited amount of time [8]. EERs enhance learner motivation, develop modern-day skills, and improve knowledge acquisition through playful, interactive experiences, as research indicates [9-12]. Frameworks like EscapED [13] have emerged to guide the creation of EERs in higher education settings, while design thinking principles have been applied to create learner-centered frameworks for EER development. EERs require careful matching of actual game goals and introduced educational objectives in order to provide a seamless user experience, which represents a very unique and specific challenge, not present in the recreational escape room design process.

Even though EERs are still considered a novel educational tool, many researchers have investigated this topic. Early research by [14] and [15] explored EER teaching potential and design considerations. Later studies focused on implementation and curriculum integration, though theoretical foundations remain limited [16], while [17] concluded EERs are generally effective across different education levels. Domain-specific reviews in STEM [18], medical education [19], nursing [20], as well as media literacy specifically [21] show positive outcomes as their research finds that ERs are enjoyable, contribute to knowledge gain, and increase motivation.

2.2 Participatory Design in Educational Game Development

PD originates from Scandinavian workplace democracy movements [22, 23] and it evolved over time into a comprehensive methodology for involving end-users in design processes. Further research [24] expanded PD toward collaborative creation, generative toolkits, and infrastructuring. PD has been applied in educational contexts to co-design digital tools, games, and curricula through involvement of end-users in iterative cycles of brainstorming, prototyping, and evaluation, ensuring designs align with users' needs and contexts [25]. So far, PD has had limited influence in serious game development, despite its popularity in interaction design, which opens up opportunities for methodological innovation.

Analyzed experimental results [26] suggest that user involvement in game design may contribute to effectiveness by creating a better fit with user preferences, making it a particularly suitable method for teaching students and children. However, systematic reviews reveal diverse and inconsistent participatory methods in educational game design, with unclear outcomes regarding optimal approaches and a lack of details on how the process is being evaluated. For example, PD methods analyzed in [26] showed significant variation. These ranged from involving students in pre- and post-tests of game interfaces to determine usability effectiveness, to poster-making activities using materials and questionnaires as multisensory design techniques. Other approaches included daily camp activities for creating and testing hands-on games as learning tools, and workshops focused on concept development, personas, and storyboards.

2.3 Media Literacy Education Through Games

Ever-growing challenges of information verification, determining source credibility, and noticing digital manipulation have gradually turned media literacy education toward game-based approaches to increase engagement among the most vulnerable groups. Research [5] suggests that, for developing misinformation identification skills, games proved to be more effective than conventional, often static content, in both online and offline delivery modes. The time-sensitive nature of escape rooms particularly suits media literacy education by simulating the fast-paced information environment young people navigate through daily [21]. This creates opportunities for experiential learning about detecting different forms of misinformation under pressure.

The Dutch Media Literacy Competency Model [27] identifies 8 key skills: operating devices and software, exploring applications, finding information (including detecting misinformation), creating with media, connecting through media, discussing media, understanding media (recognizing biases and business models), and reflecting on media usage. The framework also covers topics of health, education, employment, and identity. Ongoing media literacy training initiatives are often fragmented and short-term [21,28], with calls for more evidence-based, scalable approaches asserting gamification.

3 Methods

The participatory design process in WP3 aimed to ensure that the educational ERs being developed in the scope of the ENDGAME project (https://endgameproject.github.io) are engaging, contextually grounded, inclusive, and reflective of regional disinformation challenges. Building on the findings from WP2 (Identification and creation of contextualized media literacy scenarios), the co-design process focused on translating country-specific disinformation narratives into compelling and pedagogically relevant scenarios tailored for game-based learning environments. The goal of the participatory workshop, where participants contributed to shaping both the narrative structure and puzzle mechanics of the ERs, was to co-create learning experiences that go beyond instructional objectives, immersive storylines, visually dynamic environments, and meaningful challenges that resonate with the everyday media experiences of young people.

A core principle of this approach was the meaningful inclusion of diverse perspectives, particularly those often underrepresented in digital design and media education. Rather than treating participants as end-users, the process positioned them as co-creators. The workshops were thus grounded in the belief that participatory design is not a peripheral consultation tool, but a creative and empowering methodology through which learners directly influence the form, tone, and function of the educational media they engage with.

3.1 Participants and Recruitment

An international PD webinar was held with stakeholders from all three partner countries to collaboratively shape the high-level narrative arc, structure, and educational strategy of the ERs. To ensure representational diversity in ER design, each national team invited participants from various age groups, gender identities, socio-economic backgrounds, and digital media familiarity levels. Stakeholders included: secondary and university students, schoolteachers and digital literacy educators, media literacy trainers and youth workers, civic society representatives and media professionals, individuals from underrepresented groups. Out of 58 people registered for the workshop, 41 participated, ensuring diversity in gender, ethnicity, ability, and region.

3.2 Workshop structure and tools

Conducted as a 100-minute online event via the Zoom platform, the workshop combined presentations, interactive quizzes, breakout room discussions, and digital collaborative tools (AhaSlides and Google Forms) structured as presented in Table 1.

Table 1. Participatory design workshop structure.

Sessions	Duration	Description
Welcome and Introduction	10 min	Project overview, ER example ("The Hoax Factory") demonstrating puzzle-skill links
Presentation of Insights from WP2	10 min	"Real-Fake News" quiz (5 timed questions), Insights from WP2 (Recurring themes across regions, How insights will potentially inform escape room narratives)
Brainstorming Narrative Ideas (Breakout rooms)	25 min	"Truth and a Lie" game for rapport. Word clouds on game visuals, themes and settings.
Puzzle and Skill Mapping (Breakout rooms)	20 min	Puzzle and skill mapping example, Discussions regarding puzzle types, challenged skills and techniques for avoiding revealing solutions.
Inclusion and Accessibility Discussion	10 min	Addressing stereotypes and accessibility
Wrap-Up and Next Steps	20 min	Survey capturing additional ideas

“Welcome and Introduction” session started with a brief introduction of facilitators and explanation of the workshop purpose, as well as a brief overview of the following session activities, followed by a brief overview of the ENDGAME project and the goals of the ERs. After a short mention of project partners, the Educational ERs concept was explained and demonstrated using an example of a digital educational ER [29], which had the purpose of visually describing the presented concept.

“Presentation of Insights from WP2” began with the first ice breaker – a simple 5-question “real-fake” quiz based on different news articles and social media posts. Participants were asked to guess if the post in question was true or false, and the results were presented to participants after each question. Each question had a time limit to simulate a potential ER challenge.

The purpose of the quiz was to set the base for the rest of the workshop and provide participants with hands-on experience. The quiz was implemented and conducted through the *AhaSlides* platform. After a short discussion about the quiz results and participants' impressions, a concise summary and highlight of recurring themes of disinformation narratives gathered through national workshops in Finland, Serbia, and Spain was presented to the participants, along with a few examples of how these insights will potentially inform ER narratives.

Each breakout room facilitator started "Brainstorming Narrative Ideas" session with a simple "truth and a lie" game (Fig. 1) as a form of another more personal ice breaker, where participants could volunteer, share their name and role, and present two simple statements where only one statement is true, and other participants must guess which one. The facilitator presented the first pair of personal statements to start the game, after which other participants continued the game. The session utilized AhaSlides's Word Cloud feature (Fig. 2) for two key brainstorming activities: narrative choice selection and real-life setting preferences. After each word cloud generation session, facilitators introduced previously developed team concepts to gather participant opinions and feedback.

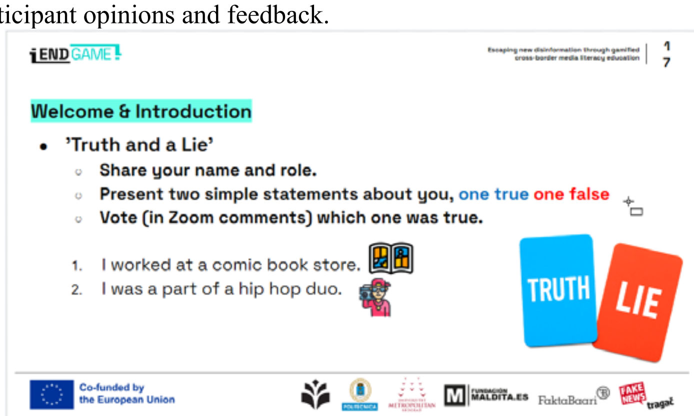


Fig. 1. Presented "Truth and a Lie" game rules as a part of the "Brainstorming Narrative Ideas" session.

The goal of the "Puzzle and Skill Mapping" breakout session was to link media literacy skills (e.g., identifying manipulated images, evaluating sources) to potential puzzle types. Facilitators presented one valid puzzle with mapped skills for easier understanding and guided groups to suggest specific puzzle ideas tied to the narrative.

Fig. 2. Narrative choice - Word cloud.

"Inclusion and Accessibility" session raised some important questions regarding how to keep content free of stereotypes, how to improve accessibility of the ER, will the game include experiences relevant to marginalized communities, etc.

In the "Wrap-Up and Next Steps" session, participants were asked to fill out the Google Forms form (25 out of 41 participants completed the survey), and the workshop was concluded by outlining the following steps that should be performed, such as development of room prototypes, piloting and testing, and follow-up invitation for those interested in contributing to storyboarding or testing.

3.3 Data collection and analysis

Data was collected through a multi-method approach to capture comprehensive participant feedback throughout the workshop. Real-time polling was conducted using *AhaSlides* to gather immediate responses during interactive activities. Facilitators documented discussions and outcomes from breakout room sessions, while word cloud generation captured brainstorming outputs on game visuals, themes, and settings. Post-workshop structured questionnaires were administered via Google Forms to collect participant feedback on workshop components and their experiences. Additionally, qualitative data was gathered through open discussions, allowing participants to provide unstructured insights and suggestions. This data collection strategy ensured comprehensive documentation of participant input across different phases of the workshop.

Analysis combined quantitative preference data with thematic analysis of qualitative discussions, identifying patterns.

4 Results

The findings from breakout sessions showed strong alignment with wrap-up survey results, confirming the consistency of participant preferences across different collection methods. Importantly, the interactive nature of breakout discussions allowed gathering more subtle details and contextual explanations that represent additional value beyond the quantitative survey data.

4.1 Visual and Thematic Preferences

Participants demonstrated clear preferences for illustrative art styles (Fig. 3), with 47.1% votes favoring this approach compared to pixel art (38.2%) or real-life photography (14.7%). This preference was consistent across all breakout rooms and was reinforced by participant concerns about AI-generated art, with explicit requests for human-created visual content.

Analysis of thematic preferences revealed three primary areas of interest (Fig. 4). Political misinformation and election interference emerged as the most preferred theme, accounting for 32% of responses. Health and science misinformation followed closely at 24% of responses, while AI-generated content represented 20% of participant preferences. Participants emphasized the timeliness and global relevance of these themes, particularly noting their impact on emotions, politics, and public trust.

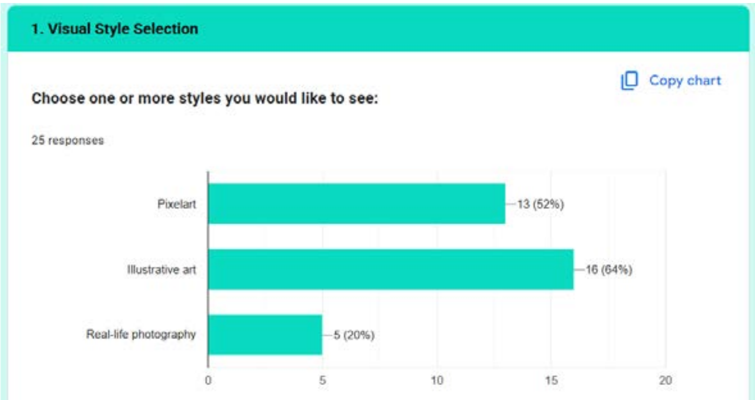


Fig. 3. Visual style selection responses in *Google Forms* (Multiple choice was allowed).

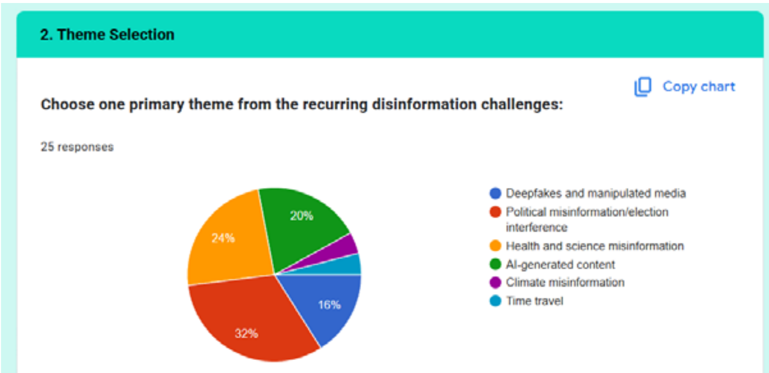


Fig. 4. Theme selection responses in *Google Forms*.

4.2 Setting and Character Preferences

Real-world settings demonstrated clear hierarchical preferences among participants (Fig. 5). Newsroom or media organization settings ranked highest at 24.4%, followed by social media company headquarters at 19.5%. Government offices received 12.2% preference, while university or classroom environments attracted 9.8% of participant interest.

Character role preferences (Fig. 6) showed a tie between journalists/reporters and students/researchers, each receiving 26.7% of responses. Fact-checkers garnered 12.1% preference, while ordinary citizens received 10.3% of participant support. These preferences reflect participants' desire for relatable, investigative scenarios rather than abstract or fantastical contexts.

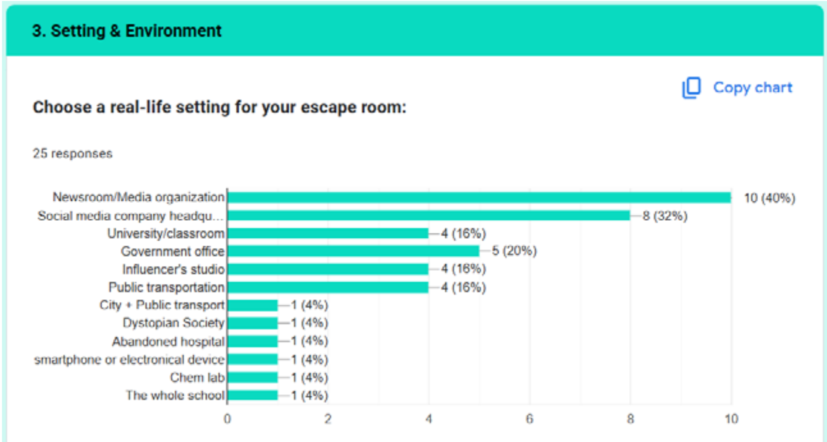


Fig. 5. Setting and Environment selection responses in *Google Forms* (Multiple choice was allowed).

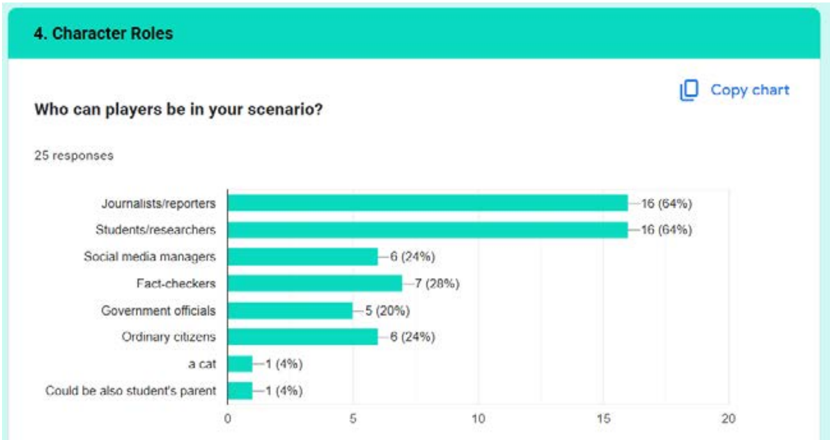


Fig. 6. Character roles selection responses in Google form (Multiple choice was allowed).

4.3 Game Objectives and Puzzle Design

Central game objectives revealed participants' focus on actionable misinformation encounter (Fig. 7). Finding the source of disinformation emerged as the primary objective preference at 30.4%, followed closely by stopping viral fake news from spreading at 26.1%. Protecting individuals victimized by disinformation was 15.2% and restoring trust in legitimate media gathered 13% of participant selections.

Puzzle type preferences (Fig. 8) emphasized several key elements. Participants favored visual analysis, source verification, pattern recognition activities, and cross-referencing tasks, and collaborative problem-solving approaches. Additional preferences included progressive difficulty with multiple solution paths and integration with narrative elements rather than standalone challenges.

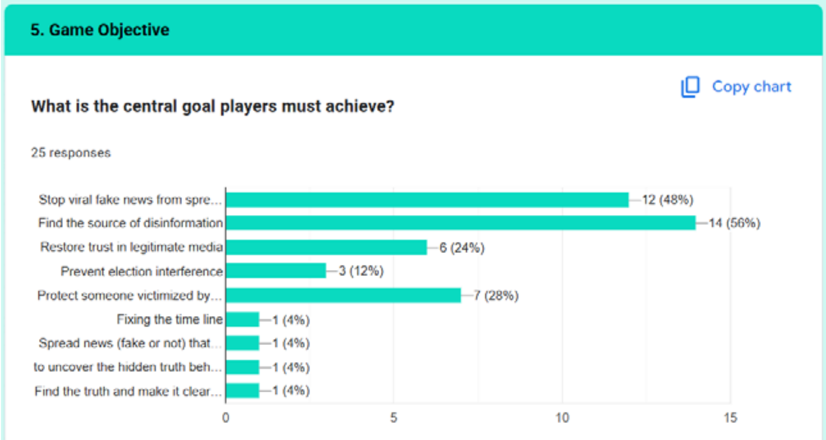


Fig. 7. Game objective selection responses in Google form (Multiple choice was allowed).

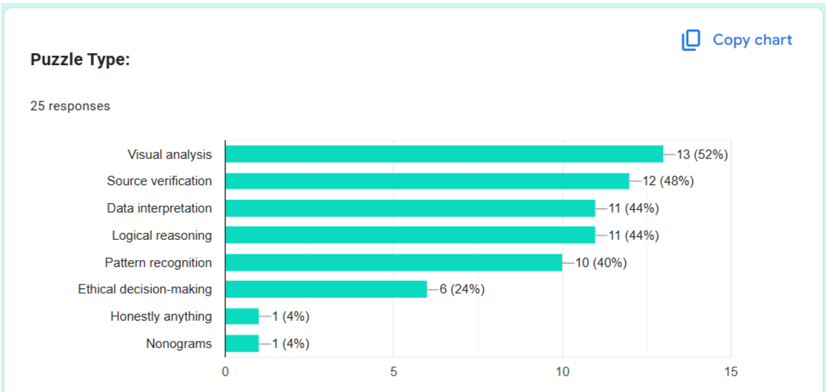


Fig. 8. Puzzle type selection responses in Google form (Multiple choice was allowed).

4.4 Accessibility and Inclusion Insights

Participants provided comprehensive recommendations for inclusive design implementation. Visual accessibility suggestions included high contrast themes for low-vision users and text alternatives for visual and audio content. Technical accessibility considerations encompassed multiple information channels (visual, textual, audio), simple language avoiding jargon, and accommodation for slower internet connections and older devices. Cultural inclusivity recommendations emphasized diverse character representation while avoiding stereotypical portrayals.

5 Discussion

Workshop results confirm the value of PD in effective involvement of end-users in pre-planning grounded, immersive, and accessible educational ER experiences. While maintaining participant attention and engagement during the process, valuable data has been gathered through a combination of real-time polling, brainstorming, and structured feedback questionnaires. Data will be used in the creation of 3 different ER games, and implementation and evaluation of these games will help to further confirm the effectiveness of PD in the development process.

Participant feedback, gathered through surveys and open discussions, emphasized several recurring themes. The PD process was perceived as fun, collaborative, and informative. Many participants, including youth, reported learning new perspectives on disinformation throughout the session. A strong preference emerged for realistic scenarios that felt grounded in everyday digital experiences and avoided abstract or moralizing tones. Neutral missions framed around investigation, teamwork, or digital problem-solving were preferred over those suggesting the pursuit of "truth."

Visual inclusion and accessibility were consistently mentioned, including suggestions on color contrast, text readability, gender balance in character design, and overall interface clarity. Both individual and group gameplay modes were seen as valuable for maximizing accessibility and engagement.

The online workshop format proved to be very effective for the inclusion of participants from different states, but revealed limitations such as limited time for deeper discussion of complex topics, trouble building rapport in a virtual environment, and the challenge of managing different views on a subject within a given time. Next to that, all participants were self-selected and may not ideally represent a broader population that may interact with the created video games.

6 Conclusion

This paper demonstrates the potential and feasibility of incorporating PD approaches as an important building block of educational ER video games. Structured online collaboration with 41 stakeholders across three countries allowed identifying end-user preferences for narrative themes, visual styles, and puzzle mechanics that balance engagement with educational objectives.

The study contributes to both methodological insights about conducting international PD workshops and real findings about stakeholder preferences for media literacy education games. The created guidelines provide a valuable starting point for similar project workshops while also highlighting the importance of user-oriented design and co-creation in educational technology development. Combined with the findings from our recent literature review about EERs for media literacy [21], the results of the PD workshops will serve as the departure point for designing the escape rooms of the ENDGAME project.

As misinformation continues to challenge democratic discourse and social cohesion, educational tools must be improved and adapted with and for the young people they aim to serve. PD offers a path toward more relevant, engaging, and effective educational games that respect learner agency while advancing critical learning objectives.

Acknowledgments. This work has been co-funded by the European Commission through its Creative Europe Programme (CREA) funding programme as part of the project “Escaping New Disinformation through GAmified cross-border Media literacy Education” (ENDGAME) (ID: 101185763).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article..

References

1. Danry, V., Pataranutaporn, P., Groh, M., Epstein, Z., Maes, P.: Deceptive AI systems that give explanations are more convincing than honest AI systems and can amplify belief in misinformation (2024)
2. Kertysova, K.: Artificial intelligence and disinformation: How ai changes the way disinformation is produced, disseminated, and can be countered. Security and Human Rights 29(1-4), 55–81 (2018)
3. Skipper, Y.: 'I'd Probably Scroll by': An Exploration of Young People's Views on Spotting and Stopping Misinformation. Children & Society 39, 940–949 (2025)
4. Nee, R. C.: Youthquakes in a post-truth era: Exploring social media news use and information verification actions among global teens and young adults. Journalism & Mass Communication Educator (2019)
5. Lu, C., Hu, B., Bao, M.-M., Bi, C., Ju, X.-D.: Can media literacy interventions improve fake news discernment? A meta-analysis. Social Science Research Network (2023)
6. Contreras-Espinosa, R. S., Eguia-Gomez, J. L.: Evaluating video games as tools for education on fake news and misinformation. Computers 12(9), 188 (2023)
7. Khaled, A. E., Gulikers, J. T. M., Tobi, H., Biemans, H. J. A., Oonk, C., Mulder, M.: Exploring the Validity and Robustness of a Competency Self-Report Instrument for Vocational and Higher Competence-Based Education. Journal of Psychoeducational Assessment 32(5), 429–440 (2014)
8. Nicholson, S.: Peeking behind the locked door: A survey of escape room facilities. https://ischool.syr.edu/wp-content/uploads/2015/05/erfacwhite.pdf (2015)
9. Pan, R., Lo, H., Neustaedter, C.: Collaboration, awareness, and communication in real-life escape rooms. Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17. the 2017 Conference, Edinburgh, United Kingdom (2017)
10. Warmelink, H., Mayer, I., Weber, J., Heijligers, B., Haggis, M., Peters, E., Louwerse, M.: AMELIO: Evaluating the Team-building Potential of a Mixed Reality Escape Room

Game. CHI PLAY '17: The annual symposium on Computer-Human Interaction in Play, Amsterdam, The Netherlands (2017)
11. Williams, P.: Using escape room-like puzzles to teach undergraduate students effective and efficient group process skills. 2018 IEEE Integrated STEM Education Conference (ISEC), Princeton, NJ, USA (2018)
12. Wu, C., Wagenschutz, H., Hein, J.: Promoting leadership and teamwork development through Escape Rooms. Medical Education 52(5), 561–562 (2018)
13. Clarke, S., Peel, D., Arnab, S., Morini, L., Keegan, H., Wood, O.: EscapED: A Framework for Creating Educational Escape Rooms and Interactive Games to For Higher/Further Education. International Journal of Serious Games (2017)
14. Fotaris, P., Mastoras, T.: Escape rooms for learning: A systematic review. In: Lars Elbaek, Gunver Majgaard, Andrea Valente, Saifuddin Khalid (eds.) Proceedings of the 13th European Conference on Games Based Learning (ECGBL 2019) (2019)
15. Veldkamp, A., van de Grint, L., Knippels, M.-C. P. J., van Joolingen, W. R.: Escape education: A systematic review on escape rooms in education. Educational Research Review 31, 100364 (2020)
16. Taraldsen, L. H., Haara, F. O., Lysne, M. S., Jensen, P. R., Jenssen, E. S.: A review on use of escape rooms in education -- touching the void. Education Inquiry 13(2), 169–184 (2022)
17. López-Pernas, S.: Educational Escape Rooms Are Effective Learning Activities Across Educational Levels and Contexts: A Meta-analysis. IEEE Transactions on Learning Technologies (2023)
18. Lathwesen, C., Belova, N.: Escape Rooms in STEM Teaching and Learning---Prospective Field or Declining Trend? A Literature Review. Education Sciences 11(6), 308 (2021)
19. Quek, L. H., Tan, A. J. Q., Sim, M. J. J., Ignacio, J., Harder, N., Lamb, A., Chua, W. L., Lau, S. T., Liaw, S. Y.: Educational escape rooms for healthcare students: A systematic review. Nurse Education Today 132, 106004 (2024)
20. Reinkemeyer, E. A., Chrisman, M., Patel, S. E.: Escape rooms in nursing education: An integrative review of their use, outcomes, and barriers to implementation. Nurse Education Today 119(105571), 105571 (2022)
21. López Pernas, S., Barra, E., Kostić, M., Milošević, M., Gordillo, A., & Saqr, M. Escape Rooms for Media Literacy: A Systematic Literature Review. Social Science Research Network. (2025).
22. Schuler, D., Namioka, A.: Participatory Design: Principles and Practices. 1st edn. CRC Press (1993)
23. Bodker, S.: Creating Conditions for Participation: Conflicts and Resources in Systems Development. Human--Computer Interaction 11(3), 215–236 (1996)
24. Sanders, E. B. N., Stappers, P. J.: Co-creation and the new landscapes of design. CoDesign 4(1), 5–18 (2008)
25. Rilla, K., Asimina, V.: Bridging serious games and participatory design. International Journal of Child-Computer Interaction 2(2), 93–100 (2014)
26. Ismail, R., Ibrahim, R., S. Yaacob: Participatory Design Method to Unfold Educational Game Design Issues: A Systematic Review of Trends and Outcome. 5th International Conference on Information Management (ICIM), Cambridge, UK, pp. 134–138 (2019)
27. Netwerk M.: The Dutch Media Literacy Competency Model 2021. https://netwerkmediawijsheid.nl/wp-content/uploads/2021/11/The-Dutch-Media-Literacy-Competency-Model-2021.pdf (2021)
28. Dumitru, E.-A., Ivan, L., Loos, E.: A generational approach to fight fake news: In search of effective media literacy training and interventions. In: Human Aspects of IT for the Aged Population. Design, Interaction and Technology Acceptance pp. 291–310 (2022)

29. Universidad Politécnica de Madrid, Maldita.es: The Hoax Factory - Educational Escape Room. https://escaperoom.maldita.es, last accessed 2025/08/22

Probabilistic Reasoning, Statistical Attitudes, and AI's Role in Analyzing Students' Explanations

Marija Kaplar[1][0000-0002-0920-8276] and Aleksandra Stevanović [2][0000-0001-5408-608X]

[1] Faculty of Technical Sciences, University of Novi Sad, Serbia,
[2] Faculty of Information Technology, Belgrade Metropolitan University, Serbia,
marija.kaplar@uns.ac.rs,
aleksandra.stevanovic@metropolitan.ac.rs

Abstract. In contemporary society, the ability to understand and interpret data has become a fundamental skill. Regardless of STEM or non-STEM fields, all individuals is expected the minimum level of stochastic literacy. This study examined Non-STEM students' base-rate reasoning, their attitudes toward statistics, and the potential of artificial intelligence (AI) to support the analysis of students' explanations. There participated 105 students, from the University of Novi Sad (non-STEM: Law, Economics) during regular classes. They solved a base-rate task, provided written explanations, and completed a attitude questionnaire. Responses were coded into six categories reflecting correctness and explanation type. Overall, 41% of students answered the task correctly, with no significant differences by faculty or gender. However, almost half of the participants chose the same wrong option. Students of the Faculty of Economics were more likely to provide explanations (70%) compared to law students (25%). When considering only responses with explanations, a significant gender difference was found, with male students more likely to provide correct probability-based reasoning, while female students more often relied on equiprobability explanations. Attitude measures showed generally positive orientations although self-efficacy was weaker. AI-based (ChatGPT) coding of explanations gives results comparable to human classification and shows its potential for identifying misconceptions, albeit with certain limitations. Findings emphasize the need to strengthen probabilistic reasoning and statistical self-efficacy among Non-STEM students, while also pointing to AI's potential as a research and pedagogical tool.

Keywords: probabilistic reasoning, base-rate reasoning, attitudes toward statistics, non-STEM students, AI-assisted coding of explanations

1 Introduction

In contemporary society, possessing a basic level of stochastic literacy, encompassing knowledge of probability and statistics, is essential not only for professional purposes but also for personal decision-making (Kapler et al, 2021; 16]. The ability to correctly interpret and evaluate data is increasingly important across diverse fields, extending

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.291K
Part of ISBN: 978-86-89755-37-4

beyond traditionally STEM-oriented professions. Even for students in non-STEM disciplines, foundational competencies in probability and statistics are crucial for informed reasoning and critical evaluation of information [16, 19].

The present study examined the extent to which non-STEM students are able to solve a task that can be regarded as a measure of basic stochastic literacy at the elementary level. In addition to selecting an answer, students were asked to provide written explanations in order to gain deeper insight into their reasoning processes. This task is also widely recognized as a diagnostic measure of the base rate neglect misconception (also referred to as base rate bias or insensitivity to prior probabilities) occurs when a base set or some other important information is neglected. Instead of considering the base rates, conclusions are typically drawn from the descriptive details presented in a task, even though such information is insufficient for making an accurate judgment. A common misconception is that if the description of a case appears to fit one category more closely, then the likelihood of belonging to that category must be higher (Kapler et al., 2021; 14, 15]. When a description does not clearly correspond to any category, individuals are often prone to the equiprobability bias**,** the belief that all possible outcomes are equally likely. In these cases, the base rate is again ignored, leading to erroneous probabilistic reasoning [10].

Beyond assessing students' susceptibility to base rate neglect, the study also explored their attitudes toward statistics, including the perceived importance of statistics, self-assessed knowledge, trust in data and media, and motivation to improve their statistical competence. Finally, in the third part of the study, we investigated the potential of artificial intelligence (ChatGPT) to support the analysis of students' explanations and to evaluate its possible role as a tool for identifying and addressing misconceptions related to probabilistic reasoning. Based on these considerations, the present study addresses three research questions concerning students' base-rate reasoning, their attitudes toward statistics, and the potential role of AI in analyzing students' explanations.

1.1 Research Questions

RQ1. To what extent do non-STEM students correctly solve a probability task, and how prevalent are base-rate neglect and the equiprobability bias in their written explanations?

RQ2. What levels and patterns characterize students' attitudes toward statistics across the four subscales: Value/Utility, Self-efficacy, Skepticism, and Motivation to Learn?

RQ3. Can an AI-based tool (ChatGPT) reliably code students' explanations against the human coding scheme?

All analyses were conducted with group comparisons by *faculty* (Law vs. Economics) and *gender***.** Because non-STEM programs are heterogeneous, meaningful differences may exist within this population; in our study, Law and Students of Economy follow substantially different curricula, so faculty-based contrasts were explicitly examined.

2 Literature Review

Decision-making in both daily practices and broader social and political contexts often relies on data analysis, highlighting the importance of data interpretation for fostering informed and active citizenship [5, 17, 20]. In many domains such as for example public health and climate change, misinterpretations of data can lead to severe and far-reaching consequences [8, 23]. Across many professional fields, the competence to accurately analyze and interpret data is regarded as essential, and professional associations, governmental bodies, and scholarly work alike stress the importance of cultivating these competencies [1, 7]. In this regard, familiarity with the fundamental concepts of probability and statistics is necessary for citizens, regardless of their professional background to participate actively in society and make sound decisions.

Although critical data literacy is fostered at all stages of education, higher education students, as future leaders and decision-makers, are expected to develop stronger knowledge and greater confidence in applying these skills. Previous research has explored misconceptions related both to probability and to statistics. Some of the results of studies conducted in fields such as banking, investment, auditing, management, insurance, and healthcare suggest that even highly educated adults often face difficulties when interpreting information and making judgments under uncertainty [3, 4, 11, 15].

Research by Khazanov and Prado [18] indicated that nearly four out of five university students enrolled in disciplines such as accounting, business, liberal arts, and mental health exhibited difficulties with probabilistic reasoning, particularly showing susceptibility to the equiprobability bias. Similarly, Hirsch and O'Donnell [12] reported that approximately three-quarters of students, predominantly from psychology but also across other fields, struggled with probability misconceptions. In such cases, individuals often resort to intuitive strategies or heuristics, which, while common, frequently lead to systematic errors in reasoning [13].

In the Serbian context, base rate neglect has been examined among STEM students, specifically engineering students, where around 35% were found to be prone to this type of misconception [17]. The study by Kaplar et al. [17] offers a useful point of reference for the present work because it employed the same base-rate task under equivalent testing conditions and procedures, allowing direct comparability of findings. Research on Serbian Non-STEM students to date has focused primarily on their interpretation of visually presented data, most notably the accurate reading of the arithmetic mean and standard deviation (Luzanin et al., 2022). In a cross-cultural design comparing Serbian STEM and Non-STEM students with STEM and Non-STEM students from Kent State University (USA), Luzanin et al. (2022) reported that Serbian Non-STEM students performed worse in interpreting the arithmetic mean than both Serbian STEM students and their STEM and Non-STEM peers from the U.S. Taken together, these studies motivated an examination of how Non-STEM students tackle probability-based tasks that index elements of stochastic literacy.

The overarching aim of such investigations is to identify potential weaknesses in general education and to highlight curricular areas where targeted improvement is warranted. This is particularly important given that previous research has shown that, even when students achieve high performance on probability tests, they often lack a deeper understanding of fundamental concepts, especially when making judgments about uncertain events in real-life contexts [9, 24]. At the same time, numerous studies highlight that carefully designed courses in probability and statistics can play an essential role in addressing these difficulties [6, 10, 11, 21]. These studies recommend emphasizing students' prior experiences, incorporating experimental and frequency-based approaches, and fostering intuitive understanding of probability concepts. Such strategies have been shown to reduce the likelihood of misconceptions and underscore the importance of instructional methods in shaping statistical reasoning.

Students' explanations provide valuable insights into their reasoning, yet analyzing large volumes of written responses is time-consuming, which is why artificial intelligence is increasingly being considered as a supportive tool in this process. Previous studies show that large language models can make a significant contribution to data analysis, but their role remains primarily supportive. McClure et al. [22] emphasize that the accuracy of AI coding is often comparable to, and in some cases higher than, human performance in deductive tasks, especially when the instructions are clearly defined. However, researchers' experiences also highlight certain limitations: Yan et al. [26] point out that AI can accelerate analysis, but its lack of reliability and deeper contextual understanding requires human validation. Williams [25] further warn about ethical and methodological challenges, such as hallucinations and the absence of theoretical grounding in interpretation. The common conclusion across these studies is that AI can enhance the speed and scope of data processing, but the human researcher remains essential to ensure accuracy and meaningful analysis.

3 Methodology

3.1 Sample

The sample consisted of undergraduate students from the University of Novi Sad (Table 1). Data were collected during regular class sessions. All students from the Faculty of Economics reported having passed at least one exam in probability and statistics at the time of testing. Participation in the study was voluntary, and all participants were informed about the aims of the research and provided their consent prior to participation.

Table 1. Sample distribution by gender and faculty

Faculty	Female	Male	Total
Faculty of Economics – Subotica	28	9	37
Law Faculty – Novi Sad	39	29	68
Total	67	38	105

3.2 Procedure

The study was conducted using a structured test consisting of several components. First, students provided demographic information. They were then asked to solve a probability task designed to assess base rate neglect and elements of statistical reasoning**.** In order to gain deeper insight into students' reasoning processes, participants were instructed to provide written explanations for their answers. The test further included a questionnaire on students' attitudes toward statistics, measured through Likert-scale items. Finally, students' open-ended responses were analyzed using an AI-based tool to explore the extent to which artificial intelligence can support the identification and understanding of misconceptions related to base rate neglect.

Statistical literacy through base rate neglect

As part of the study, participating non-STEM students were presented with a probability task (Fig. 1) that required them to reason about base rates in order to reach the correct solution. This task can be considered a component of statistical literacy, as it involves applying basic probabilistic reasoning to everyday contexts. At the same time, it is commonly used in the literature on cognitive biases as a diagnostic measure of the base rate neglect misconception, which occurs when individuals give insufficient weight to prior probabilities in favor of more salient but less relevant information.

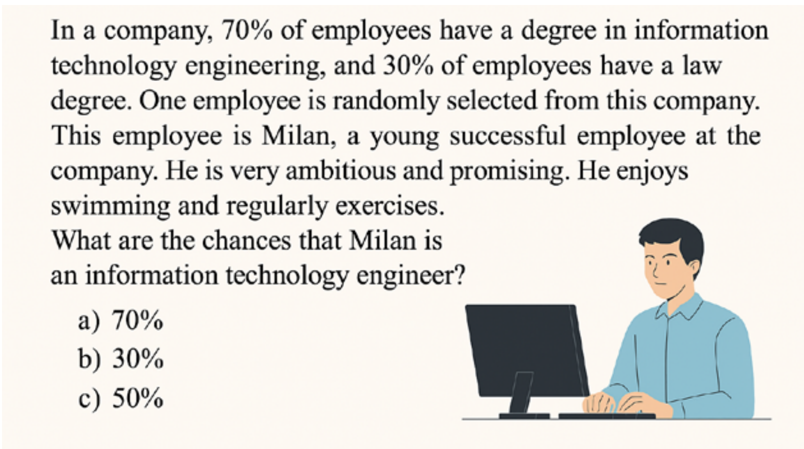


Fig. 1. Example of a probability problem illustrating base rate neglect and elements of statistical literacy *(Figure generated with an AI-based tool)*

In the present study, the task was administered before the attitude questionnaire, allowing us to examine both the correctness of students' responses and the types of justifications they provided. Focusing on a non-STEM student population provided an opportunity to explore statistical reasoning and misconceptions in a group with typically less formal training in mathematics and statistics, but whose professional competence increasingly requires data-informed decision-making.

Coding of Students' Explanations

Students' responses to the base rate task were categorized using a coding scheme that distinguished between accuracy of the answer and the presence and type of explanation. The first digit indicates whether the answer itself was correct (1) or incorrect (0). The second digit specifies the type of explanation:

- **0** – no explanation provided, regardless of correctness of the answer.
- **1** – explanation consistent with a key misconception (equiprobability bias: reasoning that two outcomes must be equally likely simply because they are two).
- **2** – alternative explanation that is incorrect or not based on probabilistic/statistical principles.
- **1 (with correct answer)** – explanation based on probability/statistical principles.
- **2 (with correct answer)** – explanation accompanying a correct response but not grounded in probability/statistical reasoning.

Accordingly, the six possible categories were:

- **00** – Incorrect answer without explanation.
- **01** – Incorrect answer with equiprobability bias explanation.
- **02** – Incorrect answer with alternative incorrect explanation.
- **10** – Correct answer without explanation.
- **11** – Correct answer with probability/statistics-based explanation.
- **12** – Correct answer with explanation not based on probability/statistics.

This scheme allowed us to differentiate between accuracy, presence of reasoning, and the type of reasoning students employed, thereby providing deeper insight into the nature of misconceptions such as base rate neglect and equiprobability bias.

Attitudes toward statistics

Attitudes were assessed with 9 items on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Guided by content and reliability screening, we formed four subscales and computed scale scores as the mean of available items (minimum half of items present). Item 2 was reverse-coded (6 − response) so that higher scores reflect greater skepticism. Higher scores indicate more of the construct. The items used in this questionnaire were previously validated and applied in two studies [16, 19]. The subscales shown in Table 2 were tested.

Table 2. Attitude toward statistics subscales with items and reliability coefficients

Subscale	Items (abbrev.)	Cronbach's α
Value/Utility (VU)	STATS_IMPORTANCE — "It is important to possess basic knowledge of statistics, regardless of profession." CRITICAL_EVALUATION — "The statistical knowledge I possess helps me critically evaluate research findings." PROFESSION_IMPORTANCE — "Statistics is essential for the profession I have chosen." PERSONAL_USAGE — "I use my knowledge of statistics to better interpret data."	0.72
Self-efficacy (SE)	STATS_EDUCATION — "During my education, I studied enough statistics." KNOWLEDGE_SATISFACTION — "My knowledge of statistics is satisfactory."	0.71
Skepticism (SK)	MEDIA_TRUST_R — "I trust research reported in the media." *(reverse-coded)* MISUSE — "Statistics are often misused."	0.46
Motivation to Learn	IMPROVE_STATS — "I would like to improve my knowledge of statistics.	—

Note: MEDIA is reverse-coded (Item 2; 6 − response).

AI-Supported analysis of student reasoning in base rate tasks

In the third part of the study, we focused on students' written explanations accompanying their answers to the base rate neglect task. These justifications were subjected to qualitative coding in order to categorize types of reasoning. The coding scheme distinguished between correct probability-based reasoning and several types of misconceptions (e.g., equiprobability bias, alternative incorrect explanations, or responses without justification).

To complement the manual categorization, students' responses were also processed using an AI-based tool, with the aim of examining whether artificial intelligence can reliably identify reasoning patterns and misconceptions. This dual approach allowed for both human-coded and AI-supported perspectives, providing a richer understanding of students' reasoning processes and exploring the potential of AI to support educational research in addressing misconceptions such as base rate neglect.

4 Results and discussion

4.1 Statistical literacy through base rate neglect

Distribution of responses to the base rate task by faculty is shown in Figure 2. Overall, 41% of students provided the correct solution to the base rate task, with 43% of students of the Faculty of Economics and 40% of law students responding accurately (Table 3). Across both faculties, approximately half of the students (around 50%) selected the 50% response option. This pattern suggests that nearly every second student was inclined toward the equiprobability bias, a misconception reflecting the belief that when two outcomes are possible, they must be equally likely [10].

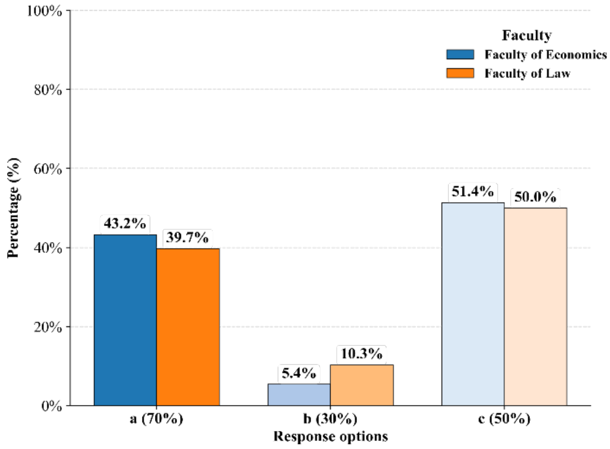


Fig. 2. Distribution of responses to the base rate task by faculty

A chi-square test revealed no statistically significant differences in accuracy between the two faculties, $\chi^2 = 0.12$, $p = 0.724$. Likewise, gender differences were not significant, $\chi^2 = 0.15$, $p = .698$. These findings are comparable to the results reported by Kaplar et al. [17], where approximately 65% of engineering students correctly solved the same problem. Taken together, the results suggest that while misconceptions are present among both STEM and non-STEM students, they tend to be somewhat more pronounced among non-STEM students.

Table 3. Correct and incorrect responses by faculty and gender with chi-square results

Category	Group	Correct	Incorrect	Total	χ^2	p
Faculty	Economics	16	21	37	0.12	0.724
	Law	27	41	68		
Gender	Female	26	41	67	0.15	0.698
	Male	17	21	38		

It might have been expected that students of the Faculty of Economics would perform better on this type of task, given that all of them had completed at least one course in probability and statistics. However, this was not the case. This finding aligns with previous research suggesting that misconceptions may persist regardless of formal education or the number of completed courses [9, 24]. The presence of such misconceptions is often rooted in incorrect intuitive reasoning rather than a simple lack of knowledge. Consequently, addressing these issues requires targeted instructional strategies and carefully designed practice integrated throughout the educational process [6, 10, 11, 21].

4.2 Analysis of students' explanations

Providing explanations in problem-solving tasks can serve as an indicator of how deeply students engage with the task and the extent of their genuine involvement. Prior research [16] has shown that students who offer explanations tend to achieve higher test scores. Providing explanations also was discussed by Attali et al. [2] who argue that the ability to articulate reasoning reflects a willingness to thoughtfully engage with the task, whereas a correct response alone may sometimes result from mere guessing. In the current study only about 40% of non-STEM students provided an explanation for their response to the base rate task (Table 4), which is lower compared to STEM students in earlier studies, where approximately 56% offered justifications (Kapler et al., 2021).

In the present study, the largest proportion of students (40%) fell into category 00, which represents incorrect answers without any explanation (Table 4). Of the total sample, 21% were able to select the correct response and provide a correct justification, whereas in a comparable study with STEM students this proportion was higher, at 32% [16]. Approximately 20% of participants were classified in category 01, where the explanations clearly reflected base rate neglect and the associated equiprobability bias. Illustrative responses from this category included:

"The chances are 50% because the worker could be either an IT engineer or a lawyer."
"There are two options, so the probability must be 50–50%."
"Since the worker was chosen randomly, the chance is equal for both professions."

According to previous literature, such explanations provide strong evidence for the presence of base rate neglect, specifically the equiprobability bias [10]. One possible reason for the persistence of this misconception may even be prior education, since many introductory courses in probability and statistics often rely on examples with equally likely outcomes, such as coin tosses or dice rolls [6, 10, 11, 21]. Taken together, these findings suggest that nearly half of the participants in our study may potentially hold the equiprobability bias, with approximately 20% providing explanations that explicitly confirm this misconception.

Table 4. Distribution of student' answers and explanations by faculty and by gender (cells show n and %; includes 00 & 10 and aggregates).

Type	Group	00	01	02	10	11	12	Without explanation (00+10)	With explana-tion (01+02+ 11+12)	N (total)

Faculty	Economics	8 (21.6%)	12 (32.4%)	1 (2.7%)	3 (8.1%)	12 (32.4%)	1 (2.7%)	11 (29.7%)	26 (70.3%)	37
	Law	33 (48.5%)	7 (10.3%)	1 (1.5%)	18 (26.5%)	9 (13.2%)	0 (0.0%)	51 (75.0%)	17 (25.0%)	68
Gender	Female	21 (31.3%)	18 (26.9%)	2 (3.0%)	12 (17.9%)	13 (19.4%)	1 (1.5%)	33 (49.3%)	34 (50.7%)	67
	Male	20 (52.6%)	1 (2.6%)	0 (0.0%)	9 (23.7%)	8 (21.1%)	0 (0.0%)	29 (76.3%)	9 (23.7%)	38

Note. Percentages are row-wise within each group over all responses. 'Without explanation' = 00+10; 'With explanation' = 01+02+11+12.

Additionally, students from the Faculty of Law were more often classified in categories 00 and 10 (answers without explanation), while students of the Faculty of Economics more frequently provided justifications. A chi-square test confirmed a significant faculty difference ($\chi^2 = 20.31$, $p < .001$) about 70% of students of economics gave an explanation, compared to only one quarter of law students. This shaped the category distribution, with students of economics equally divided between 11 (correct with explanation) and 01 (incorrect with equiprobability bias) at 32.4% each, whereas law students most often fell into 00 (incorrect without explanation) (Fig. 3).

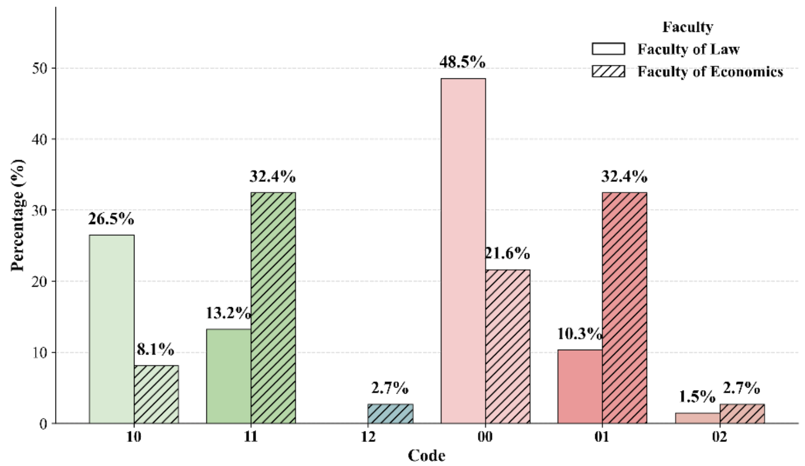


Fig. 3. Distribution of students' explanations across response categories by faculty

Gender differences were also significant, $\chi^2 = 7.34$, $p = .0068$, where roughly half of female students provided an explanation, compared to less than one quarter of males (Figure 4). When only responses with explanations were considered, a significant gender difference emerged (Figure 5). Female students were more frequently classified in category 01 (incorrect answer with equiprobability bias explanation), whereas male students were more often placed in category 11 (correct answer with correct explanation). Male students more frequently provided correct answers with explanations, and this difference was confirmed as statistically significant ($\chi^2 = 4.43$, $p = .035$) when responses with explanations were considered.

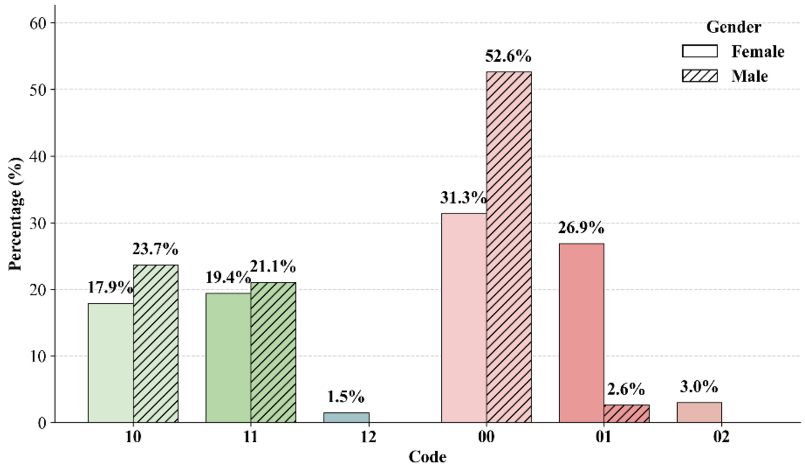


Fig. 4. Distribution of students' explanations across response categories by gender

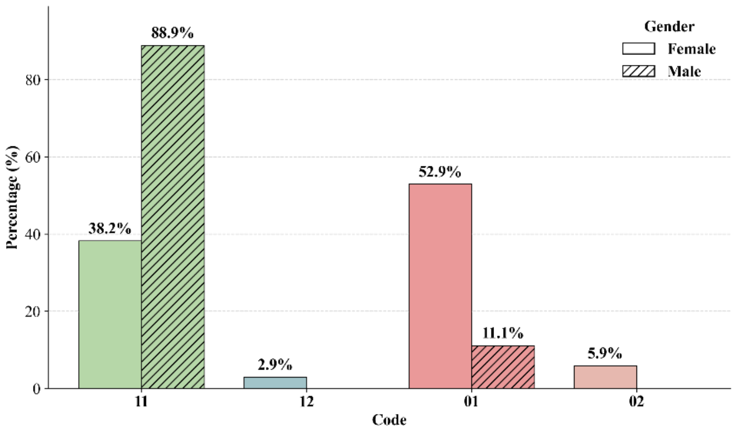


Fig. 5. Distribution of students' explanations across response categories by gender (excluding categories 00 and 10).

These findings indicate that, although no statistically significant differences emerged in the overall choice of options (a, b, or c), a closer analysis of students' explanations reveals meaningful distinctions. Students of economics were more inclined to provide justifications for their answers and, compared to their peers from the Faculty of Law, significantly more often selected the correct option accompanied by a valid explanation (code 11). In terms of gender, female students more frequently provided explanations overall, but male students were more likely to give correct probability-based explanations.

Sections 4.1 and 4.2 together provide the answer to RQ1: about 40% of non-STEM students solved the base-rate task correctly, while nearly half showed the equiprobability bias, with faculty and gender differing mainly in how students explained their answers.

4.3 Attitudes toward statistics

Table 5 presents the descriptive statistics for all items across the four subscales of attitudes toward statistics. Overall, the mean values suggest moderately positive attitudes, with the highest means observed for items reflecting the value and utility of statistics, as well as students' motivation to improve their knowledge and concerns about statistical misuse. Items related to self-efficacy yielded somewhat lower means, indicating that many students do not fully perceive their current knowledge as satisfactory.

Table 5. Descriptive overview of students' responses to attitude items toward statistics

Subscale	Item (abbr.)	N	Mean	SD	Mode	Mean (bar)
Value/Utility						
	STATS_IMPORTANCE	105	3,94	0,88	4	3,94
	CRITICAL_EVALUATION	105	3,26	0,97	3	3,26
	PROFESSION_IMPORTANCE	105	3,1	1,13	3	3,1
	PERSONAL_USAGE	105	3,18	1	3	3,18
Self-efficacy						
	STATISTIC_EDUCATION	105	2,94	1,28	2	2,94
	KNOWLEDGE_SATISFACTION	105	3,02	0,99	3	3,02
Skepticism						
	MEDIA	105	2,48	0,89	3	2,48
	MISUSAGE	105	3,62	0,91	4	3,62
Motivation						
	MOTIVATION	105	3,6	1,14	3	3,6

When compared to STEM students, no substantial difference was observed for value and utility of statistics (Mean = 3.99; Mode = 4), where responses were nearly identical [16]. Regarding students' motivation to improve their statistical knowledge, non-STEM students expressed somewhat lower interest than their STEM peers (Mean = 3.99; Mode = 5), although they still demonstrated a moderate willingness to enhance their skills. Finally, with respect to the belief that statistics are often misused, both STEM (Mean = 3.92; Mode = 5) and non-STEM students expressed similarly high levels of agreement, suggesting a shared perception that misuse of statistics is widespread.

When considering the mode values, additional nuances emerge. For several items (e.g., Importance of statistics and Misusage of statistics), the mode was 4, indicating that the most frequent response leaned toward agreement, even when the overall mean was somewhat lower. By contrast, items such as Statistical education had a mode of 2, showing that the most common response clustered around disagreement.

Interestingly, when comparing STEM and non-STEM students, only minor differences were observed for items related to Statistical education (STEM_Education_Mean = 2.12; Education_Mode = 1) and Media trust (STEM_Media_Mean = 2.36; Media_Mode = 3) [16]. Taken together, these findings indicate that both groups similarly rate their statistical knowledge as insufficient and report limited exposure to statistics during their education. At the same time, both STEM and non-STEM students demonstrated low trust in research reported by the media, suggesting awareness of potential misuse of statistics, while still recognizing the broader importance of statistical knowledge.

Table 6. Comparison of Law and Economics Students' Attitudes Toward Statistics

Item	Subscale	Mean- Faculty of Low	Mean- Faculty of Economic	U	p-value
PROFESSION_IMPORTANCE	Value/Utility	2,66	3,92	487	<0.001
STATISTIC_EDUCATION	Self-efficacy	2,54	3,68	614	<0.001
KNOWLEDGE_SATISFACTION	Self-efficacy	2,85	3,32	901,5	0,0127
MEDIA	Skepticism	2,29	2,81	886	0,0077
MISUSAGE	Skepticism	3,79	3,3	1595,5	0,0167
IMPR	Motivation	3,46	3,86	974,5	0,0491

Mann–Whitney U tests were conducted for nine items to examine potential differences between students of the Faculty of Law in Novi Sad and the Faculty of Economics in Subotica. As shown in Table 6, statistically significant differences emerged for six of these items. Students of the Faculty of Economics reported higher scores on Profession Importance, Statistic Education, Knowledge Satisfaction, Media trust, and Motivation to improve, whereas students of the Faculty of Law expressed stronger agreement with the statement that statistics is often misused. For the remaining three items, no significant differences were observed.

When comparing responses by gender, Mann–Whitney U tests revealed no statistically significant differences across any of the items. In summary, the findings suggest that students from the Faculty of Economics generally demonstrate more positive attitudes toward the value, personal relevance, and acquisition of statistical knowledge, whereas law students are more inclined to perceive statistics as frequently misused. Importantly, no gender-based differences were detected, indicating that these attitudes are more strongly shaped by academic context than by gender.

Findings for RQ2 show that students' attitudes toward statistics vary across the four subscales, with value and utility rated relatively high, while skepticism and low self-efficacy remain notable concerns.

4.4 AI in Analyzing Student Reasoning

To examine the consistency between human and AI coding, Cohen's kappa was calculated, yielding a value of $\kappa = 0.85$. This value indicates a high level of agreement, while the percentage agreement was 88.6%. The findings suggest a stable correspondence between the two coding approaches, although some discrepancies remain that require further analysis and consideration.

Two systematic patterns of disagreement emerged. In nine cases, the researcher coded the responses as 01 – equiprobability bias, while the AI coded them as 02 – alternative incorrect explanation. These were answers implying equal likelihood, such as "it either is or isn't" or "everyone has an equal chance." In three cases, the researcher coded the responses as 11 – probability/statistics-based explanation, while the AI coded them as 12 – correct but non-statistical explanation. These were verbal justifications such as "since there are more engineers, the chance is higher," which represent statistical reasoning but were not recognized as such by the AI unless accompanied by explicit numerical references. Importantly, there were no disagreements in categories without explanations (00 and 10), where coding fully aligned.

Based on these findings, revised prompts were introduced. The new version (AI coding v2) was designed to recognize verbal probabilistic explanations as 11 and to capture equiprobability bias even when not explicitly expressed as "50%." This broader interpretation, however, reduced agreement with the researcher's coding to about 40 percent (63 disagreements out of 105). The reason was again systematic: the AI more often classified ambiguous responses as equiprobability (01 instead of 02) and treated qualitative references to base rates as statistical (11 instead of 12). There were also no errors in answers without explanations (00 and 10), where the coding was consistent.

Overall, the comparison shows that the initial version of AI coding produced high alignment with human coders, with disagreements limited to clear patterns. The revised version broadened the recognition rules, but at the cost of significantly reducing agreement. These findings suggest that AI can be reliable in identifying straightforward cases, but that adjustments in how verbal reasoning is interpreted can substantially alter the results. Finding the right balance between inclusiveness and precision remains a key methodological challenge.

Results for RQ3 indicate that ChatGPT can code clear cases consistently with human raters, but agreement declines when explanations are more complex, suggesting AI has supportive but not substitutive potential.

5 Conclusion

This study gives a better understanding of how non-STEM students think about statistics. It looked at how well they can solve a basic probability problem, how they feel about statistics, and how artificial intelligence (AI) can help analyze their thinking. The results showed that less than half of the students solved the base-rate problem correctly, and many made a common mistake called the equiprobability bias — the belief that all outcomes are equally likely, even when they are not.

When responses were considered collectively, no differences emerged across faculties in the choice of answers. However, a closer examination revealed important contrasts. Students of the Faculty of Economics were more likely to provide written explanations, whereas students of the Faculty of Law more often submitted answers without elaboration. Female students were statistically more likely to provide explanations than their male peers. Among responses with explanations, further gender differences were evident: male students more frequently employed correct probability-based reasoning, while female students more often relied on equiprobability reasoning.

In general, students expressed positive attitudes toward statistics, particularly regarding its value and usefulness. However, many reported low confidence in their own knowledge and indicated that they had not studied statistics sufficiently during their prior education. On the other hand, they showed moderate to high motivation to improve their statistical knowledge. Students were also aware of the frequent misuse of data and, to a considerable extent, expressed a lack of trust in research findings reported in the media. These findings point to an opportunity for developing educational strategies that can enhance statistical knowledge among non-STEM students.

Finally, this study found that AI tools like ChatGPT can help in identifying misconceptions such as base rate neglect, but their accuracy depends on how clearly students express their reasoning and how well the prompts are written. Human oversight is still necessary to ensure reliable interpretation. These findings suggest that AI can support, but not replace, human coders in detecting reasoning errors.

The study is limited by the limited sample size and the descriptive methodology, which restrict the generalizability of the findings. Future research should include larger and more diverse groups of students and apply deeper statistical or mixed-method analyses to strengthen conclusions.

References

1. AlMuraie, E., Algarni, N. A., & Alahmad, N. Sh. (2021). Upper-secondary school science teachers' perceptions of the integrating mechanisms and importance of stem education. *Journal of Baltic Science Education*, 20(4), 546-557.https://doi.org/10.33225/jbse/21.20.546
2. Attali, Y., Laitusis, C., & Stone, E. (2016). Differences in reaction to immediate feedback and opportunity to revise answers for multiple-choice and open-ended questions. *Educational and Psychological Measurement*, 76(5), 787–802. https://doi.org/10.1177/0013164415612548.
3. Bílek, J., Nedoma, J., & Jirásek, M. (2018). Representativeness heuristics: A literature review of its impacts on the quality of decision-making. *Scientific papers of the University of Pardubice*. Series D, Faculty of Economics & Administration, 25(43), 29–38 Retrieved from https://www.scimagojr.com/journalsearch.php?q=21100390414&tip=sid&clean=0.
4. Bramwell, R., West, H., & Salmon, P. (2006). Health professionals' and service users' interpretation of screening test results: Experimental study. *British Medical Journal*, 333(7562), 284–286. https://doi.org/10.1136/bmj.38884.663102.AE.
5. Carvalho, C., & Solomon, Y. (2012). Supporting statistical literacy: What do culturally relevant/realistic tasks show us about the nature of pupil engagement with statistics? *International Journal of Educational Research*, 55, 57–65. https://doi.org/10.1016/j.ijer.2012.06.006
6. Delmas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28–58.
7. Dönmez, I., Idin, S., & Gürbüz, S. (2022). Determining lower-secondary students' STEM motivation: A profile from Turkey. Journal of Baltic Science Education, 21(1), 38. https://doi.org/10.33225/jbse/22.21.38
8. Faghmous, J. H., & Kumar, V. (2014). A big data guide to understanding climate change: The case for theory-guided data science. *Big Data*, 2(3), 155-163. http://doi.org/10.1089/big.2014.0026
9. Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372–396. https://doi.org/10.1111/j.1751-5823.2007.00029.x.
10. Gauvrit, N., & Morsanyi, K. (2014). The equiprobability bias from a mathematical and psychological perspective. Advances in Cognitive Psychology, 10(4), 119–130. https://doi.org/10.5709/acp-0163-9.
11. Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S.(2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, Supplement, 8(2), 53–96. https://doi.org/10.1111/j.1539-6053.2008.00033.x.
12. Hirsch, L. S., & O'Donnell, A. M. (2001). Representativeness in statistical reasoning:Identifying and assessing misconceptions. *Journal of Statistics Education*, 9(2). https://doi.org/10.1080/10691898.2001.11910655.
13. Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristic and biases: The psychology of intuitive judgment*, (pp. 49–81). New York: Cambridge University Press.

14. Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4), 237–251. https://doi.org/10.1037/h0034747.
15. Kang, M., & Park, M. J. (2019). Employees' judgment and decision making in the banking industry: The perspective of heuristics and biases. *International Journal of Bank Marketing*, 37(1), 382–400. https://doi.org/10.1108/IJBM-04-2 018-0111.
16. Kaplar, M. (2022). *Recognizing misconceptions in working with data as a basis to enhance mathematical literacy* (Publication No. 29111073) [Doctoral dissertation, University of Novi Sad]. ProQuest Dissertations and Theses Global.
17. Kaplar, M., Lužanin, Z. & Verbić, S. (2021). Evidence of probability misconception in engineering students —why even an inaccurate explanation is better than no explanation. *International Journal of STEM Education*, 8, Article 18. https://doi.org/10.1186/s40594-021-00279-y
18. Khazanov, L., & Prado, L. (2010). Correcting students' misconceptions about probability in an introductory college statistics course. *ALM Internationa Journal*, 5(1), 23–35.
19. Lužanin, Z., Kaplar, M., & Hrubik-Vulanovic, T. (2022). Undergraduate STEM and non-STEM students' interpretation of mean in an infographic. *Journal of Baltic Science Education*, 21(4), 638-650. https://doi.org/10.33225/jbse/22.21.638
20. Maass, K., Doorman, M., Jonker, V., & Wijers, M. (2019). Promoting active citizenship in mathematics teaching. *ZDM*, 51(6), 991–1003. https://doi.org/10.1007/s11858-019-01048-6
21. Masel, J., Humphrey, P. T., Blackburn, B., & Levine, J. A. (2015). Evidence-based medicine as a tool for undergraduate probability and statistics education. CBE Life Sciences Education, 14(4), 1–10. https://doi.org/10.1187/cbe.15-04-0079.
22. McClure, J., Smyslova, D., Hall, A., & Jiang, S. (2024). Deductive Coding's Role in AI vs. Human Performance. In *Proceedings of the 17th International Conference on Educational Data Mining* (pp. 809-813).
23. Park, S., Bekemeier, B., Flaxman, A. D. (2021). Understanding data use and preference of data visualization for public health professionals: A qualitative study. *Public Health Nursing*, 8(4), 531-541. https://doi.org/10.1111/phn.12863
24. Smyrnaiou, Z., Georgakopoulou, E., & Sotiriou, S. (2020). Promoting a mixed- design model of scientific creativity through digital storytelling—the CCQ model for creativity. *International Journal of STEM Education*, 7(1), 1–22. https://doi.org/10.1186/s40594-020-00223-6.
25. Williams, R. T. (2024). Paradigm shifts: exploring AI's influence on qualitative inquiry and analysis. *Frontiers in Research Metrics and Analytics*, *9*, 1331589.
26. Yan, L., Echeverria, V., Fernandez-Nieto, G. M., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & Martinez-Maldonado, R. (2024, May). Human-ai collaboration in thematic analysis using chatgpt: A user study and design recommendations. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-7).

Remote Project Management in the Metaverse Using BIM and Extended Reality (XR)

Vishak Dudhee[1,2][0000-0002-8155-9513], Prasanna Bandara[3][0009-0004-8978-3888], Charalampos Psarros[4][0009-0004-9564-2643], Azhar Nalakath[1][0009-0001-4895-3746], and Vladimir Vukovic[1,5][0000-0003-0702-2475]

[1] V-LAB LTD, The Beacon, Esplanade Avenue, Redcar, TS10 3AA, UK
[2] University of Exeter, Rennes Drive, Exeter, EX4 4PU, UK
[3] Tricore Technical Services, Villa Jubilant, 34 Falcon Ct, Stockton-on-Tees, TS18 3TX, UK
[4] Teesside University, Borough Road, Middlesbrough, TS1 3BX, UK
[5] Belgrade Metropolitan University, Tadeuša Košćuška 63, Beograd, 11158, Serbia
vishak@v-lab.uk, p.dinesh@tricore-ts.com, c.psarros@tees.ac.uk, azhar@v-lab.uk, vlad@v-lab.uk

Abstract. The integration of Building Information Modelling (BIM) with Extended Reality (XR) technologies has shown strong potential to advance project planning and visualisation in the construction sector. In the aftermath of COVID-19, the shift to remote and hybrid work has accelerated the need for collaborative tools that extend beyond traditional communication platforms. While digital practices are emerging, the potential role of the Metaverse in supporting remote construction project management remains underexplored. This paper looks at the feasibility of managing construction projects and facilitating stakeholder collaboration within a Metaverse environment through the integration of BIM and XR. A case study using a university building demonstrates how immersive environments can enhance spatial understanding, communication, and cross-disciplinary coordination. The findings highlight improved stakeholder engagement, reduced reliance on physical travel, and alignment with sustainability objectives such as net-zero emissions. Challenges were also identified which includes device performance, user comfort, and varying levels of VR proficiency. The study provides evidence that BIM-XR integration in the Metaverse offers a promising pathway for digital transformation in construction project management, while pointing to areas for further refinement and research.

Keywords: Building Information Modelling (BIM); Extended Reality (XR); Metaverse; Virtual Reality (VR); Construction Project Management; Remote Collaboration

1 Introduction

Globally, the construction industry employs approximately 7% of the working-age population, making it one of the largest sectors of the world economy [1]. Despite its scale, the industry has traditionally relied on incremental innovation, with approaches such as off-site construction and 3D printing still at an exploratory stage. At present, off-site

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.308D
Part of ISBN: 978-86-89755-37-4

construction represents only around 2% of the global market [2]. Conventional on-site construction, whereby structures are assembled sequentially at their permanent location, continues to dominate industry practice. Furthermore, compared with other major sectors, construction remains among the least digitised [3].

The COVID-19 pandemic was an accelerator for the adoption of digital technologies and remote work practices, which led to the successful transition of many office-based roles within the construction industry to remote and hybrid arrangements. The post-pandemic period has also normalised flexible working practices, enabling employees to organise their schedules around non-traditional hours and alternative locations. This shift has fostered greater managerial trust in employees' ability to deliver tasks without direct physical supervision [4]. However, research indicates that remote work can undermine collaboration, with reliance on emails and videoconferencing often replacing real-time, in-person interactions. This substitution has been shown to hamper communication and, in turn, project coordination. Consequently, there is a growing demand for more effective modes of remote collaboration.

At this critical juncture in the sector's digital transformation, immersive technologies present a significant opportunity to enhance collaboration and efficiency. The Metaverse, when combined with BIM and XR, provides a platform where geographically dispersed teams can interact with models, simulate construction environments, and engage in real-time coordination. Against this backdrop, this paper investigates the integration of BIM and XR within a Metaverse environment to support remote project management and stakeholder collaboration. The study proposes a Remote Collaboration Framework and evaluates its application through a case study of the Teesside University's Student Life Building, assessing both the benefits and limitations of immersive collaboration for the construction industry.

2 Background

2.1 Remote Project Communication and Management

Effective communication is widely recognised as one of the most critical competencies required by project managers [5]. In the context of remote project delivery, this requirement becomes even more pronounced, as project managers must demonstrate not only advanced interpersonal skills but also the ability to adaptively and competently use a variety of digital communication platforms [6]. Furthermore, managing remote or "virtual" teams requires sensitivity to the communication preferences and working styles of dispersed team members. When managing multinational teams, project managers must also remain highly perceptive of cultural differences and familiar with the preferred communication channels across different contexts [7].

The implementation of video conferencing and other synchronous communication methods requires project managers to handle time zone variations and scheduling requirements [8]. The most commonly used email method for communication proves to be less effective at building trust and clarity between team members [9]. Project managers need to create well-organised written communication that delivers precise information to overcome these communication barriers. Active listening serves as a critical factor for achieving correct information exchange and team member relationship development [10]. Team members should use one-to-one communication to resolve personal issues and create trust while enhancing team unity. Research demonstrates that proper communication methods create high-performing project teams.

In international projects, the complexity of stakeholder management is increased by the differences in nationality, culture, and professional norms. Project managers must first develop cultural awareness and demonstrate empathy, cultural competence, and expertise to lead diverse teams effectively [11,12]. Failing to develop these skills will lead to cultural gaps, resulting in conflicts and misunderstandings, ultimately reducing project performance. Gap analysis has been suggested as a practical tool for evaluating the difference in culture and to identify the possible sources of tension that may arise among stakeholders [13]. Establishment of shared team culture is an essential factor as it provides a framework for acknowledging and respecting the diverse perspectives in the workplace. Engaging with the local and regional partners is also another strategy that can be used for understating the cultural gap and addressing the risks in collaboration [14].

Negotiation skills have an essential role in closing the cultural divides, resolving conflicts and promoting trust within multicultural project environments [15]. Effective project managers must integrate insights from economic and sociological factors of culture, as business and project management practices vary from place to place. This will help to develop more adaptive and inclusive approaches to communication and stakeholder management, leading to enhanced project outcomes in international and remote contexts.

2.2 Building Information Modelling (BIM)

Building Information Modelling (BIM) is a collaborative digital process for generating, storing, and managing multidisciplinary information across the entire life cycle of a construction project [16]. It involves developing a coordinated digital representation of the built asset, typically comprising information-rich 3D models linked with structured data such as product specifications, execution details, and handover information. BIM maturity is defined through a series of levels, from Level 0 to Level 3 and beyond [17].

Level 0 represents no collaboration and relies solely on 2D CAD drafting for production information, distributed via paper or electronic prints [18]. Level 1 introduces limited 3D CAD for concept work alongside 2D drafting for statutory and production purposes. Standards are governed by BS 1192:2007, and information is shared electronically through a common data environment (CDE), often managed by the contractor. Level 2 requires collaborative working with project-specific information exchange, coordinated across systems and participants. Interoperability is ensured through open formats such as IFC (Industry Foundation Class) and COBie (Construction Operations Building Information Exchange). This approach has been mandated by the UK government as the minimum requirement for public-sector projects [19]. Level 3 is characterised as a fully integrated approach in which all project data are embedded within common information models. Unlike the federated approach of Level 2, Level 3 emphasises the use of internationally recognised data standards to structure information consistently across disciplines. Crucially, this integration is conceived from a lifecycle perspective, ensuring that information flows seamlessly from design and construction through to operation and maintenance. Such an approach reflects the ambition to deliver a unified and standardised digital environment for the construction sector [20].

BIM Level 3 centres on cloud-based, extended collaboration [21]. Even prior to the COVID-19 pandemic, project teams were typically distributed across multiple organisations, including architects, engineers, contractors, and specialist subcontractors [22]. Cloud-based BIM offers significant benefits in such contexts by enabling access to complex models via a web browser without requiring high-specification hardware. Models are processed into standardised formats that can be merged from various file types (e.g., Revit RVT, IFC, Bentley DGN, BCF) using platforms such as 3D Repo, creating a single source of truth for all stakeholders [23]. This approach reduces reliance on locally installed software, simplifies IT requirements, and ensures that all users access the most up-to-date model version with revision histories preserved.

Cloud-based BIM has demonstrated significant improvements in collaboration and project management. It provides the ability for stakeholders to access models at any time and work collaboratively in real time by annotating concerns directly within and addressing them either in real time or asynchronously. This will help minimise delays that are often found during email correspondence and promote greater transparency among stakeholders [24]. Project coordinators can continuously monitor design progress, rapidly identify errors such as misaligned elements, and ensure corrections are made before construction commences. Cloud-based BIM provides a scalable and efficient digital framework that enhances collaboration across the Architecture, Engineering, and Construction (AEC) industry by improving accessibility, interoperability, and real-time communication.

2.3 Extended Reality (XR)

Extended Reality (XR) refers to a broad spectrum of immersive technologies which bring together the physical and digital environments. This includes Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). XR has become a transformative tool for visualising designs by offering an immersive and interactive experience that enhances the design, planning and stakeholder communication process [25]. XR supports the design review, assessment of alternatives, early identification of potential issues and informed decision-making for stakeholders in the AEC sector by allowing them to explore the digital models before construction [26].

XR enables the simulation of buildings and designs at scale and within their context. The immersive walkthrough feature allows everyone involved to visualise the spatial relationships and better understand the proportions and aesthetics in a realistic manner, giving them a better understanding of the design intent. The immersive visualisation experience provides more accurate feedback with a collaborative approach and helps reduce the risk of costly design revision later. XR also allows multiple users to be in the same virtual model in real time, wherever they are physically. Such shared environments encourage clearer communication, collective problem-solving, and joint decision-making [27]. For clients and end-users, XR provides a tangible way of engaging with proposed designs, which has been shown to improve satisfaction and alignment with project outcomes.

XR also has significant potential within Metaverse (an interconnected virtual environments that combine different digital worlds and augmented experiences) beyond the individual building projects. Urban-scale planning is supported within this XR framework by featuring modelling and exploration of entire cities. It also allows users to interact with architecture/buildings in innovative ways, with the integration of virtual buildings and infrastructure into a shared digital environment.

XR represents a paradigm shift in building model visualisation by providing immersive, interactive, and collaborative experiences. Its ability to improve design comprehension, support effective communication, and enhance decision-making establishes XR as a critical enabler of digital transformation in the AEC sector. As XR technologies continue to mature and become more widely accessible, their integration with the Metaverse is expected to influence the future of architectural practice, urban planning, and stakeholder engagement [28].

3 Remote Collaboration Framework

3.1 Framework Architecture

The Remote Collaboration Framework paves the path for stakeholders like project managers, architects, engineers, and contractors to collaborate in real time virtually, regardless of wherever they are physically in the world by connecting them to a shared digital workspace through Extend Reality (XR) technologies including Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR).

The digital workspaces are accessed using XR devices like a VR headset. This allows stakeholders to participate in design reviews and manage projects without being physically present on site. This reduces non-essential travel, supports flexible collaboration, and increases efficiency. At the core of the framework is a Federated BIM Model that integrates architectural, structural, and mechanical, electrical, and plumbing (MEP) components. The model minimises miscommunication and errors by ensuring that all participants work with the same source of information.

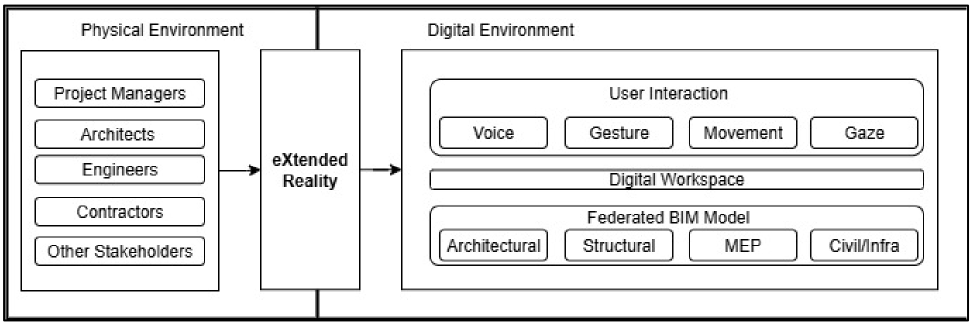


Fig. 1. Collaboration Framework Architecture

The framework (Figure 1) supports multiple user interaction methods, including voice, gesture, movement, and gaze tracking. These provide a collaborative experience which feels more natural compared to other techniques like emails and online meetings. Stakeholders can also conduct a walkthrough of the BIM models and suggest any design changes with annotations. The virtual site inspections also replicate the dynamics of an in-person meeting.

The digital workspace is not just a visualisation tool but a fully interactive environment where users can manipulate models, identify and resolve design conflicts, and maintain coordination across disciplines. Project managers can monitor progress while architects and engineers refine designs collaboratively, and contractors can avoid delays associated with outdated information with the real-time engagement supported by shared access to up-to-date data. Communication clarity and efficient workflow can be achieved with the integration of BIM with Metaverse, offering a practical means of managing construction projects remotely.

3.2 Metaverse Collaborative Platform

The case study was done using Horizon Workroom, which is a VR collaboration platform developed by Meta that provides an immersive workspace environment. The platform can be accessed through VR headsets such as the Oculus Quest 2. The platform also has customised avatars for users to represent themselves in the VR world. Users can then use the avatars to interact with virtual environments and collaborate with others in real-time.

The platform comes packed with a wide range of communication and collaboration tools. The spatial audio feature gives realism to the voice conversation between users. There is also a screen-sharing feature that allows users to showcase their work for discussion and feedback. The virtual whiteboard supports brainstorming and visual collaboration. Files can also be imported directly into the workspace. Options are available to integrate other productivity tools like the Oculus Browser and Zoom to streamline collaboration between users further. Hence, the platform provides web access for the users without having to leave the virtual world. Non-verbal communication is supported by using hand signs, body language and expressive gestures which provides a stronger sense of presence and interaction among the team members.

3.3 Compatible Devices

For visualising BIM models, the case study utilised desktop and laptop computers with sufficient processing power, alongside VR headsets such as the Oculus Quest 2. AR devices, including Microsoft HoloLens, were used to overlay BIM models within real-world settings, enabling users to experience designs in context. Mobile devices with VR headsets were also employed to access the Metaverse model from multiple locations. However, the quality of visualisation depends on factors such as model size, complexity, internet speed, and system hardware capacity.

4 Case Study

4.1 Project BIM Model

The Student Life Building at Teesside University's Middlesbrough Campus (opened in 2020) was designed to support a wide range of learning styles and attendance patterns. This was achieved through a technologically enabled environment. The building offers flexible spaces for collaborative learning alongside an information zone, consulting rooms, and a café. This study used a replicated BIM model of this building. The building has a complex spatial organisation and integration of multipurpose spaces which made it a suitable candidate for this study for evaluating the use of BIM in XR environments.

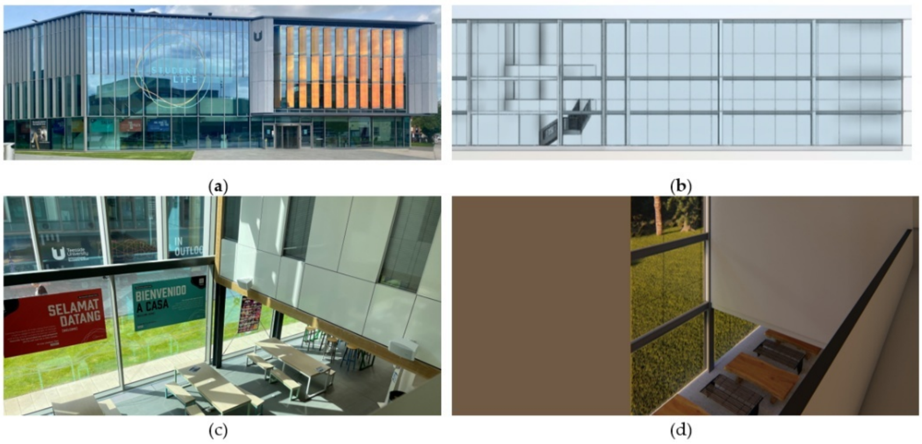


Fig. 2. (a) A photo of the Student Life Building; (b) The developed BIM model (Exterior); (c) Interior photo of the Student Life Building Atrium; (d) The developed BIM Model (Interior)

Architecturally, the building features a glass façade with dichroic glass fins (Figure 2a and 2c) and interconnected interior volumes shaped by timber-clad pods containing 26 consultation and study rooms. These pods are arranged between double-height volumes, with waiting areas doubling as quiet study spaces. The spatial configuration balances openness and connectivity with enclosed areas for individual or group study.

As the original BIM model of the building was not available, a replica was developed in Autodesk Revit using available architectural references, dimensional data, and photographic material (Figure 2b and 2d). The model was then imported into Unreal Engine and processed to generate realistic rendering and textures. Structural elements and materiel properties like the reflective characteristics of the dichroic glass fins and the natural finish of the timber-clad pods were incorporated to mirror the real-world building closely. The spatial layout, including circulation routes and study areas, was modelled to ensure fidelity to the original design. Even though the replica does not fully substitute the original BIM model, its accuracy was validated against available architectural documentation to provide a reliable digital representation. For exploring the flexibility another 3D model of a cricket stadium for the case study. In addition to the Student Life Building, a 3D model of a cricket stadium was also used to further test the applicability of BIM in XR environments.

The BIM model had to be optimised with reduced triangle sizes in tools such as Blender for integration with XR platforms to be accessed through standalone VR and AR devices. In 3D graphics, triangulation is used to construct all digital models, and reducing their number makes the BIM model lighter and more efficient for real-time rendering on standalone VR and AR devices. This allowed stakeholders to conduct immersive walkthroughs, assess spatial relationships, and test alternative design configurations remotely. The model provided valuable insights into spatial dynamics and stakeholder engagement by enabling exploration of the building at scale within the collaborative virtual workspace. Importantly, it demonstrated the potential of combining BIM with XR technologies in the Metaverse to support remote project collaboration, enhance communication, and inform decision-making in construction and architectural practice.

.

4.2 Metaverse Collaboration

Meta Horizon Workrooms is a VR platform designed to support immersive teamwork, which was used to implement the collaborative environment for this case study. Different devices like VR headsets, laptops and tablet computers were used by stakeholders such as Project Manager, BIM Coordinator, Architect and Engineer to connect to the shared virtual workspace. This framework, where users can get connected using different kinds of devices, provides accessibility and flexibility for users to engage with the model according to their role and available technology.

Figure 3 shows the overview of stakeholder interaction within the Metaverse workspace used for this case study. The BIM Coordinator and Project Manager accessed the federated BIM model in immersive 3D via VR headsets, providing them with spatial analysis and in-context discussion. The Architect, working on a laptop, contributed design insights and integrated updates through the desktop interface. The Engineer, accessing the environment via a tablet, was able to participate in mobile reviews and provide feedback during live sessions.

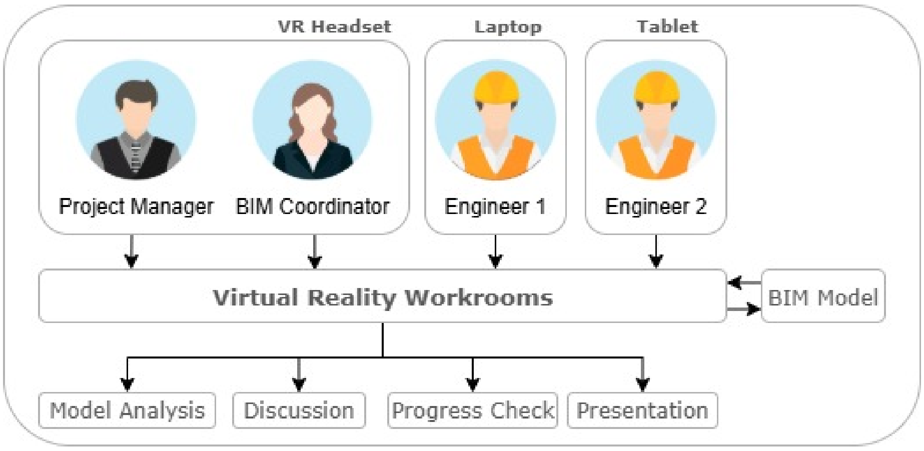


Fig. 3. Case study collaboration approach

Different collaborative activities, such as model analysis, progress reviews, design coordination, and presentations in the immersive world, were done using Horizon Workrooms. These were supported by the platform's core features like spatial audio, avatar, screen sharing and the virtual whiteboard. The combination of all these features resulted in a realistic, immersive and interactive experience that closely replicated the dynamics of an in-person meeting without the physical or geographical constraints.

Participants can annotate models to provide reviews in the virtual workspace. They can also use intuitive gestural input along with verbal feedback and engage in detailed coordination sessions. Being able to communicate with each other in the virtual world using intuitive inputs has improved communication efficiency and helped in reducing delays associated with asynchronous exchanges. This also limited the need for physical travel to an extent, demonstrating clear benefits of the framework for both sustainability and productivity. The results highlight the potential of Metaverse-based collaboration to support digital transformation in the construction industry.

4.3 Results

The case study demonstrated the practical benefits and limitations of integrating BIM models with XR technologies in a Metaverse-based collaborative platform. A federated BIM model of the Student Life Building was successfully uploaded into Meta Horizon Workrooms, enabling participants to interact with the design in real time. Stakeholders joined the environment through different devices, including Meta Quest 2 headsets, desktop computers, and tablets. This setup created a mixed participation scenario where some users experienced full immersion through VR headsets, while others engaged through standard screen-based interfaces.

Participants reported enhanced spatial awareness and design comprehension when reviewing the BIM model in virtual reality compared with traditional 2D plans or desktop-based 3D models. Being able to visualise the model at true scale in the immersive environment during the walkthrough helps improve clarity and detect misaligned elements intuitively. The ability to combine real-time voice and gesture input, along with the avatar-based presence of the user, provided a sense of natural communication. This helped to improve the effectiveness of coordination meetings.

Figures 4(a) and 4(b) show the initial interaction with the uploaded BIM model. Participants accessed the virtual workspace, viewed the building design in 3D, and began collaborative discussions. The shared environment replicated the dynamics of a physical meeting room, while overcoming geographical constraints by connecting participants across different locations.

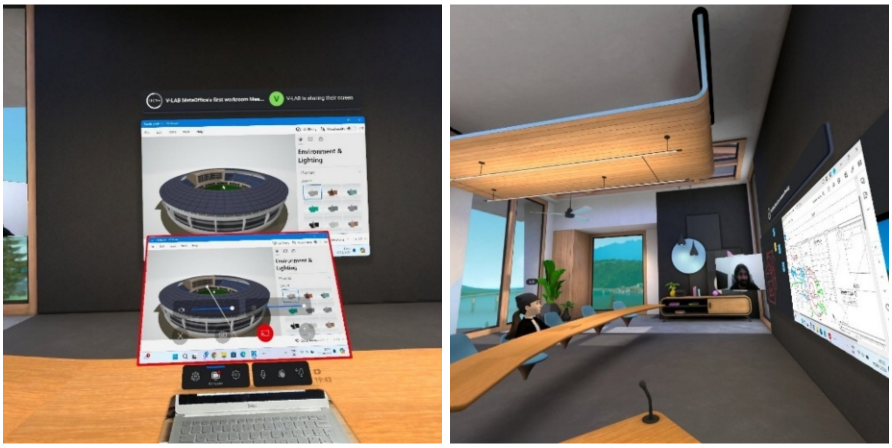


Fig. 4. Stakeholders accessing the uploaded BIM model within Horizon Workrooms. (a) Initial view of the default building model in the shared virtual workspace. (b) Multi-device participation, showing interaction between VR headset users and desktop participants.

Figures 5(a) and 5(b) highlight the use of the platform for design coordination. In Figure 5(a), a building design is displayed on the shared screen, while participants discuss its features and raise coordination issues. In Figure 5(b), an avatar representing one of the team members explains the design directly within the environment, suggesting modifications and highlighting areas of concern. This interaction demonstrated how avatar presence, combined with spatial audio, can make virtual design reviews more engaging and productive.

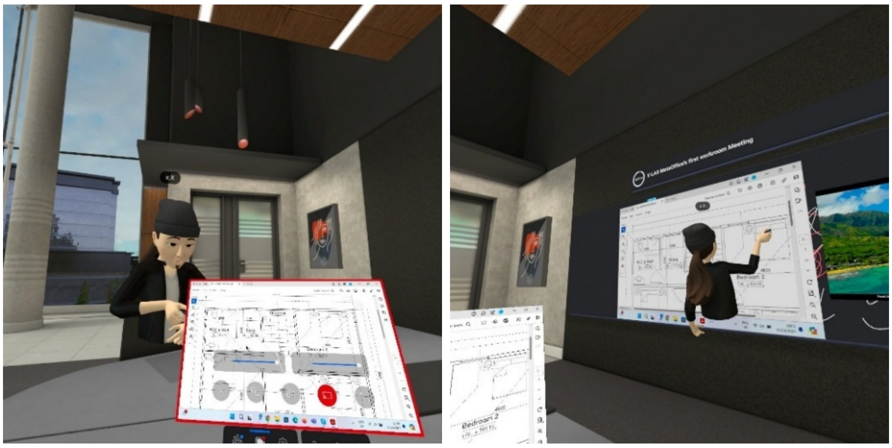


Fig. 5. Collaborative design coordination in the Metaverse environment. (a) Shared screen display of the BIM model during discussion. (b) Avatar-based explanation of design features with real-time feedback and modification proposals.

Figures 6(a) and 6(b) illustrate deeper collaboration and issue management. Figure 6(a) shows suggestions for plan modifications overlaid directly on the model, providing clarity on how changes would be implemented and which sections were affected. Figure 6(b) captures a broader group discussion in which participants debated multiple design items, combining both VR-immersed users and screen-based participants. These sessions confirmed that hybrid participation is feasible, although headset users reported greater immersion and focus compared with those joining via desktop.

Fig. 6. Model modification and interdisciplinary discussion. (a) Overlay of suggested design changes within the shared BIM model. (b) Group interaction and debate on multiple design items across VR and desktop interfaces.

Participants highlighted increased engagement and inclusivity during collaborative sessions. The avatar-based presence helped maintain meeting etiquette, while shared model viewing supported more focused discussions. However, several challenges were noted. Performance varied across devices, with some latency occurring during the rendering of complex geometries and lower-bandwidth connections. Extended VR sessions also caused discomfort for some users, including dizziness and loss of balance. Additionally, varying levels of familiarity with VR controls created a learning curve, indicating the need for onboarding and training.

Even though several limitations were found, the integration of BIM and XR within Horizon Workrooms was positively received. Early stage planning, stakeholder engagement and interdisciplinary coordination were improved with XR as demonstrated by the hands-on sessions. The framework shows strong potential for improving productivity and sustainability in construction project delivery, as evidenced by the reduced reliance on physical travel and real-time design evaluation with virtual collaboration.

5 Discussions

This study investigated the integration of Building Information Modelling (BIM) and Extended Reality (XR) within a Metaverse environment, with a focus on remote project management, stakeholder collaboration, and immersive visualisation. The study's findings highlight the potential and limitations of adopting XR technology in construction project workflows. There was an accelerated shift towards the digital distributed working models that stressed the need for collaborative tools which go beyond the traditional tools like email, online meeting and file sharing when COVID-19 hit the world. The results from this case study put forward the Metaverse-based collaboration platform as a viable solution for addressing such needs. By enabling stakeholders to interact with federated BIM models at scale, supported by real-time communication features such as spatial audio, avatars, and virtual whiteboards, the framework enhanced spatial awareness, strengthened engagement, and improved the efficiency of design coordination.

The use of avatars contributed positively to participant presence and meeting etiquette. While avatar realism in terms of facial expression remains limited, gestures and spatial positioning offered sufficient cues to replicate aspects of non-verbal communication. This indicates that even with current technological constraints, avatar-based collaboration can approximate the dynamics of physical meetings. Future advances in photorealistic avatars and facial tracking could further improve inclusivity and authenticity in virtual collaboration. Communication quality was generally strong, particularly due to the use of spatial audio, which made conversations more natural and intuitive. However, participants noted that multilingual collaboration remains a challenge. The potential integration of real-time translation or subtitles represents an important area for further research, as it could increase accessibility for multinational teams and reduce language barriers in global projects.

With respect to model interaction, the integration of BIM into the XR environment enhanced design comprehension and facilitated intuitive problem-solving. Participants reported a higher ability to detect conflicts such as clearance issues and misalignments, compared with reviewing 2D drawings or desktop-based 3D models. This validates the growing body of research that positions XR-enhanced BIM as a powerful tool for collaborative design review. Nevertheless, limitations in visual fidelity were observed. Certain material textures and reflective surfaces were not rendered with full accuracy (e.g. transparency), which reduced realism and could impact decision-making in design contexts where material finish is critical. Improvements in rendering algorithms and device performance are therefore necessary to achieve fully immersive visual quality. Device performance varied, with latency observed in complex models or under lower bandwidth conditions. Technical and ergonomic challenges were also noted. Some users experienced discomfort during prolonged VR sessions, highlighting the importance of ergonomic headset design and user adaptation periods. Additionally, differences in VR proficiency across participants affected collaborative efficiency, reinforcing the need for onboarding and training sessions.

The findings from the study confirm that XR-enabled collaboration has transformative potential for construction project management. It is also observed that there is a greater impact in early-stage planning and multi-stakeholder engagement. While technical limitations remain, the study provides evidence that immersive environments can improve communication, reduce travel, and contribute to broader sustainability objectives by supporting remote collaboration.

6 Conclusion

This paper presented a Remote Collaboration Framework that integrates Building Information Modelling (BIM) and Extended Reality (XR) technologies within a Metaverse environment. Through a case study using the Student Life Building and the Meta Horizon Workrooms platform, the study demonstrated how immersive environments can enhance spatial awareness, strengthen communication, and support cross-disciplinary collaboration in construction projects. The research results showed that XR-based Metaverse systems replicate the dynamics of face-to-face meetings while limiting physical travel needs, which supports environmental sustainability and digital transformation initiatives. The study revealed multiple challenges such as device performance issues, rendering quality problems, user comfort during long VR sessions and differences in user technical abilities. These challenges highlight the importance of onboarding, training, and continued refinement of XR platforms to ensure effective adoption.

The case study demonstrated BIM-XR integration through Metaverse technology which shows great potential for future remote project management systems. The framework serves as an essential tool to increase productivity while helping construction projects meet sustainability goals through better team communication and coordination.

Future research needs to focus on developing virtual environment realism and making XR technology accessible to all users. The complete realisation of virtual environments for efficient project delivery requires additional research into intelligent collaboration tools.

Acknowledgments. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure of Interests. The authors declare no competing interests. The use of Meta Horizon Workrooms, Oculus Quest 2, Microsoft HoloLens, and other software or hardware in this study was solely for research and demonstration purposes. The authors did not receive financial support, sponsorship, or promotional incentives from the respective companies.

References

1. Coskun, S., Washington, C. and Erdogmus, E. (2024) "A Study on Causes of Gender Gap in Construction Management: High School Students' Knowledge and Perceptions across Genders," Buildings, 14(7), p. 2164. Available at: https://doi.org/10.3390/buildings14072164.
2. National Association of Home Builders (2025) The Offsite Construction Market Share Flattens Nationally in 2024, 28 August. Available at: https://www.nahb.org/blog/2025/08/the-offsite-construction-market-share-flattens-nationally-in-2024
3. Abioye, S.O. et al. (2021) "Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges," Journal of Building Engineering, 44, p. 103299. Available at: https://doi.org/10.1016/j.jobe.2021.103299.
4. McPhail, R. et al. (2023) "Post-COVID remote working and its impact on people, productivity, and the planet: an exploratory scoping review," The International Journal of Human Resource Management, 35(1), pp. 154–182. Available at: https://doi.org/10.1080/09585192.2023.2221385.
5. Henderson, L.S. (2008) "The Impact of Project Managers' Communication Competencies: Validation and Extension of a Research Model for Virtuality, Satisfaction, and Productivity on Project Teams," Project Management Journal, 39(2), pp. 48–59. Available at: https://doi.org/10.1002/pmj.20044.
6. Akhorshaideh, A.H.O. et al. (2024) "The impact of IT tools on project management efficiency in the public sector: The mediating role of team communication," Journal of Project Management, 9(4), pp. 345–352. Available at: https://doi.org/10.5267/j.jpm.2024.8.002.
7. Pullan, P. and Prokopi, E. (2016) Leading virtual project teams: dos and don'ts. Newtown Square, PA: Project Management Institute.
8. Morley, S., Cormican, K. and Folan, P. (2015) "An Analysis of Virtual Team Characteristics: A Model for Virtual Project Managers," Journal of technology management & innovation, 10(1), pp. 188–203. Available at: https://doi.org/10.4067/s0718-27242015000100014.
9. Morgan, L., Paucar-Caceres, A. and Wright, G. (2014) "Leading Effective Global Virtual Teams: The Consequences of Methods of Communication," Systemic Practice and Action Research, 27(6), pp. 607–624. Available at: https://doi.org/10.1007/s11213-014-9315-2.
10. Goldstein, S. and Naglieri, J.A. (eds.) (2014) Handbook of Executive Functioning. Springer New York. Available at: https://doi.org/10.1007/978-1-4614-8106-5.

11. Obikunle, O. (2002). Dealing with cultural diversity in project management: a dilemma in communication. Paper presented at Project Management Institute Annual Seminars & Symposium, San Antonio, TX. Newtown Square, PA: Project Management Institute.
12. Huang, J. (2016) “The Challenge of Multicultural Management in Global Projects,” Procedia - Social and Behavioral Sciences, 226, pp. 75–81. Available at: https://doi.org/10.1016/j.sbspro.2016.06.164.
13. Ranf, D.E. (2010) “Cultural Differences In Project Management,” Annales Universitatis Apulensis Series Oeconomica, 2(12), pp. 657–662. Available at: https://doi.org/10.29302/oeconomica.2010.12.2.18.
14. Isern, G. (2015) “Intercultural Project Management for IT: Issues and Challenges,” Journal of Intercultural Management, 7(3), pp. 53–67. Available at: https://doi.org/10.1515/joim-2015-0021.
15. Lückmann, P. (2015) “Towards Identifying Success Factors for Cross-cultural Project Customer Engagement: A Literature Review,” Procedia Computer Science, 64, pp. 324–333. Available at: https://doi.org/10.1016/j.procs.2015.08.496.
16. Wang, T. and Chen, H.-M. (2023) “Integration of building information modelling and project management in construction project life cycle,” Automation in Construction, 150, p. 104832. Available at: https://doi.org/10.1016/j.autcon.2023.104832.
17. Brewer, R. (1988) “Building information,” Batiment International, Building Research and Practice, 16(6), pp. 346–351. Available at: https://doi.org/10.1080/01823328808726925.
18. Succar, B. (2009) “Building information modelling framework: A research and delivery foundation for industry stakeholders,” Automation in Construction, 18(3), pp. 357–375. Available at: https://doi.org/10.1016/j.autcon.2008.10.003.
19. Eadie, R. et al. (2013) “BIM implementation throughout the UK construction project lifecycle: An analysis,” Automation in Construction, 36, pp. 145–151. Available at: https://doi.org/10.1016/j.autcon.2013.09.001.
20. Lidelöw, S., Engström, S. and Samuelson, O. (2023) “The promise of BIM? Searching for realized benefits in the Nordic architecture, engineering, construction, and operation industries,” Journal of Building Engineering, 76, p. 107067. Available at: https://doi.org/10.1016/j.jobe.2023.107067.
21. Goulding, J. S., Rahimian, F. P., & Wang, X. (2014). Virtual reality-based cloud BIM platform for integrated AEC projects. Journal of Information Technology in Construction, 19, 308-325.
22. Fellows, R. and Liu, A.M.M. (2012) “Managing organizational interfaces in engineering construction projects: addressing fragmentation and boundary issues across multiple interfaces,” Construction Management and Economics, 30(8), pp. 653–671. Available at: https://doi.org/10.1080/01446193.2012.668199.
23. Goulding JS, Rahimian FP, Wang X (2014). Virtual reality-based cloud BIM platform for integrated AEC projects, ITcon Vol. 19, Special issue BIM Cloud-Based Technology in the AEC Sector: Present Status and Future Trends, pg. 308-325, https://www.itcon.org/2014/18
24. Nguyen, Tuan Anh, Nguyen, Tu Anh and Tran, T.V. (2024) “Building Information Modelling (BIM) for Construction Project Schedule Management: A Review,” Engineering, Technology & Applied Science Research, 14(2), pp. 13133–13142. Available at: https://doi.org/10.48084/etasr.6834.
25. Dudhee, V., Bandara, P.D., Vukovic, V., Nalakath, A., Bass, P., Bass, E.: Artificial Intelligence-Enhanced Virtual Reality for Health and Safety Training in Construction (2024)
26. Dudhee, V. and Vukovic, V. (2021) “Integration of Building Information Modelling and Augmented Reality for Building Energy Systems Visualisation,” Springer Proceedings in

Energy. Springer International Publishing. Available at: https://doi.org/10.1007/978-3-030-63916-7_11.
27. Dudhee, V. and Vukovic, V. (2021) "Building information model visualisation in augmented reality," Smart and Sustainable Built Environment, 12(4), pp. 919–934. Available at: https://doi.org/10.1108/sasbe-02-2021-0021.
28. Dudhee, V, Vukovic, V & Rahimian, F 2022, Building Sensory Information in Augmented Reality: A Conceptual Framework. in A Roxin (ed.), Nouvelles perspectives du BIM: Le jumeau numérique de la construction intelligente (smart building) . 1 edn, Eyrolles, pp. 109-122. <https://izibook.eyrolles.com/produit/9610979/9782212356199/nouvelles-perspectives-du-bim>

Immersive Learning in Higher Education: Integrating Artificial Intelligence and Virtual Reality

Miljan Stevanović[1][0009-0007-8912-5625], Dragana Nikolić-Ristić[3][0000-0003-0066-8644], Emilija Kisić[2][0000-0003-3059-2353], Milica Mladenović[3][0000-0003-3210-0316], Vladimir Vuković[2][0000-0003-0702-2475], and Petar Pejić[2][0000-0003-4155-8038]

[1] Faculty of Digital Arts, Belgrade Metropolitan University, Tadeuša Košćuška 63, 11000 Belgrade, Serbia
[2] Faculty of Information Technology, Belgrade Metropolitan University, Tadeuša Košćuška 63, 11000 Belgrade, Serbia
[3] Faculty of Management, Belgrade Metropolitan University, Tadeuša Košćuška 63, 11000 Belgrade, Serbia
miljan.stevanovic@metropolitan.ac.rs, dragana.nikolic@metropolitan.ac.rs, emilija.kisic@metropolitan.ac.rs, milica.mladenovic@metropolitan.ac.rs, vladimir.vukovic@metropolitan.ac.rs, petar.pejic@metropolitan.ac.rs

Abstract. In this paper we propose an innovative framework for improving immersive learning in robotics education that combines virtual reality (VR), extended reality (XR), and artificial intelligence (AI). The main focus is to apply robotics in areas such as manufacturing, logistics, and automation, and to explore how immersive tools can make the learning process more effective. Using VR headsets, learners have interactive simulations that reflect real situations, like checking a production line or managing robotic packaging. These practice-based experiences can help them to gain skills which are highly demanded in modern industry environments. Also, these simulations can show where mistakes in workflow might occur, but without the risks of real production. The aim of the proposed approach is to reduce the global shortage of skilled workers in robotics by offering training that is easier to access and closer to real-world practice. This method allows repetition of procedures at low cost, gives quick feedback, and helps students manage complex information more easily which differs from traditional teaching methods. The framework is flexible and can be tested within academic courses that follow educational standards. It also sets the stage for future work, including issues of ethics, technology, and law in robotics training, as well as the design of custom systems for more advanced learning.

Keywords: Immersive Learning, Virtual Reality, Extended Reality, Artificial Intelligence, Robotic Education.

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.324S
Part of ISBN: 978-86-89755-37-4

1 Introduction

Education in the field of robotics faces a number of challenges in the era of digital transformation. On the one hand, there is a lack of qualified labor in the production, logistics and automation sectors, while on the other hand, the demand for employees with knowledge in robotics, automation and mechatronics is constantly increasing. Despite the growing global skills gap in robotics and manufacturing, it is clear that traditional education methods are failing to develop the necessary practical and digital skills.

New trends in industry show that robotics is one of the most in-demand skills today. Research conducted by Deloitte and the Manufacturing Institute in 2021 predicts that by 2030, 2.1 million manufacturing jobs will remain unfilled in the United States [1]. The World Economic Forum predicts a shortage of over 7 million skilled workers in manufacturing by 2030. At the same time, only 30% of the workers on the production line have the necessary skills in accordance with the new digitization trends, and the demand for robotics is the most pronounced [1].

It can be concluded that the growing gap between what the industry requires and the skills in the field of technology possessed by experts requires a change in educational learning models. In this context, technologies of virtual reality (VR), augmented reality (XR) in integration with artificial intelligence (AI) provide new opportunities for learning through interaction and simulations in realistic but controlled conditions. Immersive learning enables students to face complex industrial scenarios (e.g. control of a production line or robotic packaging) in realistic conditions. In this way, cognitive assimilation of knowledge is improved and operational risks are reduced.

The combination of VR, XR, and AI can significantly transform education, providing scalable, efficient, and ethically sound learning models to advance education in the domains of robotics and industrial automation, potentially closing the existing gap between academic learning and industry demands.

This paper is organized as follows: In Section 1, we provide an introduction that outlines the current state of robotics and the challenges of integrating XR and AI technologies, with a particular emphasis on the shortage of industrial skills; Section 2 reviews relevant research and presents the proposed framework; Section 3 elaborates on the integration of XR solutions; Section 4 discusses the methodology for validation and implementation in academic courses; Section 5 highlights the expected results in terms of enhancing student skills, fostering interactive learning, and reducing operational risks; finally, the conclusion summarizes the contributions and outlines future development perspectives.

2 Related work

It is undeniable that new technological trends have imposed new requirements in the organization and management of teaching and learning. Contemporary technological achievements and the application of AI tools emphasize the need to identify and analyze issues related to their implementation in the education sector.

The influence of AI in education has increased significantly and the academic sphere is becoming more efficient and personalized, but also global, contextually intensive and asynchronous [2]. A number of authors point out that the competitiveness of higher education institutions is conditioned by increasing the efficiency of learning methods, i.e. the application of AI tools in education [[3], [4], [5]]. Kee & Zhang [6] and Neroni et al. [7] state that the implementation of IVR can increase student motivation and engagement, while simultaneously reporting difficulties in student concentration due to technology and the social environment [8]. Sung et al. [9] confirm the improvement of students' attitude and greater enjoyment of learning compared to classic video materials, while Pande et al. [10] point out that IVR promotes long-term knowledge retention and greater interest compared to video content, although perceived learning benefits diminish over time. On the other hand, Al-Azawei et al. [11] concluded that no significant differences were found in performance or ease of use compared to interactive e-testing in learning management systems (e.g. Moodle).

There is a possibility of the appearance of positive and negative emotions during the application of IVR, where the intensity of emotions depends on the design of the IVR environment and the students' self-confidence [12]. Positive effects of IVR learning can occur in the field of mechanical and electrical engineering [13], in learning foreign languages compared to mobile applications [14] and for improving oral presentation skills during public presentations [15].

Perez & Keleş [16] propose a model for the future development of education based on VR technology, which by encouraging deeper connections between cognitive processes and physical activities improves student learning outcomes. A new framework for Virtual Laboratory Environments (VLEs), focuses on learning through bodily experience, offering students the opportunity to physically interact with virtual samples and machines from the fields of mechanical and materials engineering. Compared to traditional (non-corporeal VR methods), research results have shown significant improvements in students' understanding and retention of knowledge, along with better test scores. The authors conclude that immersive VR environments can significantly enhance the learning experience for engineering students through hands-on and interactive experiences.

Stracke et al. [17] with a systematic analysis of scientific works on the application of immersive virtual reality in higher education, conclude on the one hand that a large number of studies highlight the positive impact of IVR on the degree of student engagement, encouraging interactive and research activities, supporting specific learning processes, increasing the enjoyment of learning and complete immersion in IVR environments. However, the authors also state that the implementation of IVR does not automatically lead to an improvement in the learning process, as numerous studies show no significant differences in educational outcomes between IVR and traditional learning methods, and recommend a careful analysis of the specific conditions of IVR application (physical side effects, the need to accept technology and individual differences among students), as well as its meaningful integration into the curriculum.

Creating highly engaging and personalized learning environments by integrating immersive technology and AI has the potential to transform traditional education systems. Bekteshi [18] states in his work that the application of immersive technologies (VR, AR, Mixed Reality - MR) and AI can significantly improve educational outcomes by: facilitating conceptual understanding through experiential and interactive learning, increasing the rate of knowledge retention through engaging, gamified and immersive educational scenarios, by providing safe and controlled simulations. Immersive technologies increase student interest by making teaching interactive and visually striking, while AI personalizes environments by adapting content and pace to individual student needs while providing real-time feedback [18]. The integration of immersive technologies (experiential context necessary for deeper learning) and AI (enhances interactivity and adaptability of VR/AR environments) represents a powerful combination for improving education [18].

Compared to previous research, this paper proposes a comprehensive and scalable framework that integrates VR, XR, and AI technologies with the aim of advancing education in the field of robotics. Unlike existing approaches, the proposed framework provides practice-oriented, transferable, and ethically grounded training aligned with contemporary educational standards.

3 Proposed framework

The proposed framework for immersive learning in higher education consists of the integration of AI and VR and is based on the previously developed XR4Human-SERVE 5.0 approach [19]. This project demonstrated the feasibility of neuroergonomics – the interaction of humans and other parts of the system and the work environment itself – with an XR-enabled assembly workstation, which was validated in the context of an SME with multimodal interfaces and cognitive assessment integrated into a single XR headset. XR4Human-SERVE 5.0, Horizon Europe established a comprehensive and integrated immersive learning environment that leveraged the synergistic potential of VR, XR and AI to enhance robotics education. Our current framework aims to overcome the shortcomings of traditional robotics curricula, which are effective in providing theoretical knowledge and limited laboratory practice. Practical teaching in higher education institutions, in most cases, cannot provide students with a certain amount of practical engagement, often due to the unavailability and affordability of certain expensive equipment. For these reasons, and the very requirements of the industries and their realities, a new framework is needed to implement practical training for students. Our framework relies on constructivist learning theories, which refer to the field of active learning and the discovery of cause-and-effect relationships that are part of the study, as well as contextualized interaction in the acquisition of knowledge and the development of skills [20].

The basis of our framework consists of three interdependent modules that address critical issues in current robotics education:

- Immersive simulation environments;
- Adaptive learning systems powered by AI;
- Assessment and feedback mechanisms.

Immersive simulation environments, as the first module, consist of high-quality XR simulations that replicate complex industrial scenarios such as robotic assembly lines, warehouse logistics operations, and automated inspection in quality control. Unlike traditional university labs, specially created virtual environments offer scalability and repeatability. Students can experiment with and operate simulated robotic machines (most often replicas of robotic industrial arms) without the risk of equipment damage or safety hazards. Such an approach is consistent with previous research showing that immersive robotics-focused learning games and VR platforms, such as Robotics Academy and IL-PRO, successfully provide safe, repeatable, and industrially relevant training environments [21]. The design of the simulation is based on the fundamental principles of cognition, which state that learning is most effective when cognition is based on sensorimotor and hands-on experiences. Our research is also consistent with studies showing that active physical engagement improves conceptual understanding, as students who physically interacted with objects while actively participating in tasks outperformed those who merely observed [22]. Thus, by engaging students in physically and cognitively rich contexts, the first module of our framework demonstrates that learning through practice is more effective than traditional learning. Furthermore, the proposed framework for immersive learning bridges the gap between abstract robotics concepts on the one hand and applied industrial practice on the other.

Our second module uses adaptive AI systems to enable personalization of the educational model and adaptability to a large group of students. Machine learning algorithms analyze patterns obtained in interactions with students. The module's algorithms track success and achievements in solving performance tasks such as: response time, error rate and navigation strategies and, based on the results, dynamically adjust the difficulty of tasks, introduce task difficulty scaling mechanisms or offer targeted micro-interventions. Thanks to the system's adaptability and constant feedback, students receive a differentiated educational experience aligned with their prior knowledge, learning pace and cognitive preferences. In addition to personalization - adjusting the difficulty level to each student individually, we expect AI to function as an intelligent mentor, a guide, offering contextual advice, guiding reflective thinking and simulating collaborative problem solving through natural language dialogue, especially in robotics education, as shown by Kal et al. [23]. ARtonomous is an example of a successfully implemented project that integrates adaptive AI in robotics education. ARtonomous allows high school students to train virtual robots using reinforcement learning within an adaptive simulation environment - students dynamically engage with changing task difficulty based on real-time performance [24].

Assessment and feedback mechanisms, as the third module, provide comprehensive, continuous and formative assessment of results. Assessment mechanisms integrate multimodal assessment strategies that capture behavioral data, interaction logs, and task results for each student. Dashboards visualize student progress in real time, and the results are available to both students and instructors. The capabilities of an automated feedback mechanism can foster self-regulated learning, i.e., the student has the ability to manage and continuously regulate his or her own learning process. Yang et al. [25] found that frequent testing combined with feedback significantly enhances long-term memory in students. Immersive technologies emphasize the importance of multimodal feedback such as visual feedback, tactile feedback, and adjustable sensory channels, as well as the importance of personalizing educational materials to support inclusive education and meet the diverse needs of students [26]. The third framework aims to digitize existing robotics curricula, and to transform pedagogical practice using immersive, adaptive, and ethically grounded educational experiences.

4 Proposed development and integration of XR platform

The development of the XR platform represents a practical implementation of the proposed framework Immersive Learning in Higher Education: Integrating AI and VR. The development itself requires a balance between technical reliability, pedagogical goals, and user-centered approach and design. This platform will be the backbone of immersive robotics education because it supports both individual and team learning approaches, as suggested in the literature [27].

To develop the XR platform, we require:

- Hardware devices;
- Software architecture;
- Integration of artificial intelligence.

Hardware devices are: VR headsets that provide a sense of immersion in the VR space thanks to high-resolution displays; haptic gloves and feedback devices that allow tactile interaction with objects in the virtual environment; and motion tracking systems, which are often integrated with other devices, VR headsets, vests, and gloves. There are also more advanced systems that can further enrich our stay in the VR world, such as integration with exoskeletons or wearable sensors that can capture additional feedback. Studies highlight the key role of VR/AR integrated with haptics in enabling multisensory, embodied learning for STEM contexts [28], while recent work in engineering education highlights the frameworks of haptic interaction for interactive learning [29].

A software platform built on engines such as Unity and Unreal Engine that have model design capabilities, advanced physics simulation capabilities, and real-time rendering capabilities. The modular design allows developers to recombine existing parts of the system, without disrupting the overall system, and create new scenarios for robotics training. The cloud, which stores data, provides the ability for multiple users to access the platform simultaneously and enables teamwork on common tasks. The integration of artificial intelligence and learning models that allow for dynamic adjustment of task complexity based on student performance in real time enhances the platform's functionality. We have reinforcement learning (which involves training machine learning models using computer programs) supporting the generation of personalized scenarios. New scenarios allow students to encounter new variations of problems that improve the transfer of knowledge to real-world robotics practice. Natural language processing (NLP), which combines computational linguistics, predictive artificial intelligence, and deep learning models to process human language, and the modules function as a virtual assistant with whom we can talk, guide us through the training process, and provide us with answers to the questions we ask. By successfully integrating artificial intelligence and the software platform, we ensure the compatibility of task difficulty and a personalized approach to students.

The design of the platform should follow the principles of human-computer interaction (HCI) focused on the ease of use of the proposed frameworks for immersive learning. The platform should have intuitive interfaces that are easy to understand, visual cues that have universal meaning, and progressive scaling of tasks. The platform should have the ability to adjust visual elements, such as text size, accessible audio descriptions, i.e. tutorials for solving tasks, and be compatible with assistive devices. The platform itself must meet ethical principles and foster inclusiveness among diverse student populations [[27], [30]].

The platform design uses recognized standards such as SCORM, xAPI, and LTI for compatibility with Learning Management System (LMS) systems, and facilitates integration into existing curricula and allows for tracking student achievement across platforms and institutions. Interoperability is supported by recent initiatives in XR-based robotics education that emphasize the importance of interoperability and collaborative learning in industrial contexts [31].

The development of the XR platform also follows a design-based research (DBR) methodology. The iterative approach is a circular model for developing a platform that involves iterative cycles of prototyping, testing, and refinement, through several iterations: in the first phase, a prototype is tested, then feedback is used to improve that prototype – this process is used until we are satisfied with the functionality of the platform itself. Early prototypes are tested with a small group of students and educators, then the design is modified and pedagogical fit and technical robustness are taken into account. Later iterations, nearing the completion of the platform, involve larger-scale testing with more diverse cohorts to improve scalability and reliability. DBR is established as a methodology for developing technologically advanced learning environments, especially for balancing theory with iterative design in authentic contexts [32]. The XR immersive learning platform aims to be a high-quality, simple and pedagogically effective network that connects theoretical robotics education with current industrial needs [28].

5 Validation and deployment of platform at academic courses

Validation and implementation of the platform represent a phase of transition from the conceptual framework of platform development, presented in the previous chapters, to practical application in the educational process in higher education. This phase requires rigorous empirical, systematic, testing, systematic integration into curricula and programs, and long-term sustainability planning [33].

Validation begins with pilot projects conducted in selected robotics courses at the Metropolitan University of Belgrade, such as operations research, machine learning, artificial intelligence, and virtual/augmented reality (VR/AR) in video games. These pilot projects will represent mixed-method research designs, combining quantitative and qualitative approaches to obtain a holistic picture of the effectiveness of the platform. Experimental settings will include control and experimental groups: the control group engages in conventional, traditional, teaching, while the experimental group uses the XR platform.

The experimental approach will allow for comparative analysis of outcomes such as skill acquisition, conceptual understanding, and motivation [33]. Quantitative data collected and processed by the platform includes the time it took the student to complete the tasks, the number of errors they made, and the retention time for each individual task. Thanks to the data obtained, we will have objective evidence of the development of skills and the percentage of students' task-solving efficiency. In addition to quantitative data, qualitative data will address students' perceptions of engagement during practical work on tasks, the degree of motivation, and the ease, i.e. difficulty, of using the platform. Quantitative and qualitative data will provide triangulated evidence of the educational impact of the platform, as shown in the literature [34].

Feedback collected during the platform development cycle from the initial pilot project enables further, iterative, refinement of the technical and pedagogical aspects of the platform. For example, if learners have difficulty navigating complex interfaces and report a problem, then usability adjustments will be made, i.e. interface redesign. Similarly, performance data can reveal where AI-driven personalization requires modification of task difficulty. Iterative feedback was intended to ensure that the platform evolved in response to real-world use and the needs of different industries [28]. Implementation requires extensive engagement of academic staff who should be trained in the technical functioning of the XR platform and its pedagogical integration. Organized workshops and professional development sessions should be designed to adapt specific teaching subjects to immersive learning. For example, educators will be guided on combining XR activities with theoretical lectures, assignments and lab exercises, as outlined in the supporting literature [27].

The implementation should take into account ethical principles and legal issues, i.e. the security of the data entered by the student should be ensured. Data governance policies are established to govern the collection, storage and analysis of student data, ensuring compliance with data protection regulations such as the GDPR [35]. Students should follow informed consent protocols that are implemented to guarantee transparency and their autonomy. Accessibility standards are built into the implementation to ensure the inclusiveness of all students - that no student is disadvantaged due to disability or resource constraints. The goal of the implementation is to create a sustainable model for immersive learning in the workplace, which includes establishing maintenance protocols, providing ongoing technical support, and updating technical maintenance. Feedback loops allow for continuous data collection from instructors and students to drive iterative improvements to the platform. The platform is intended as a tool for robotics education, but also as a scalable model that can be adapted to other technical disciplines, such as mechanical engineering, mechatronics, or even healthcare training [28]. The XR platform is intended to become an integral and sustainable part of academic programs through rigorous validation and careful implementation of all components. The long-term contribution of the learning platform in higher education is reflected in the possibilities of a new form of immersive, adaptive and ethically responsible education that prepares students for the complexities of the Fifth Industrial Revolution.

6 Expected results

Significant industrial, operational and pedagogical outcomes are expected to be generated by implementing the suggested framework for combining AI and XR, which directly addresses the observed skills gap in industrial automation and robotics. Students are primarily expected to significantly improve their robotic systems operation, programming and troubleshooting skills. By providing a variety of contextual experiences that are usually unavailable in physical labs due to resource limitations, the realistic simulations in this mastery learning environment enable repeated and intentional implementation of complex activities, which has been proven to significantly improve the development of performance and diagnostic skills [36]. AI-driven personalization further enhances this by offering adaptable challenge levels and tailored feedback, thereby targeting each learner's unique speed of skill development and encouraging higher cognitive engagement and self-regulated learning [[37], [38]].

This framework leverages scalable XR technologies in order to provide affordable access to industrial simulations of sophisticated robotic factories that would otherwise be too expensive or logistically challenging to establish in a classroom [39]. By making high-end practical training a crucial part of robotics education, this experience allows students from various institutional backgrounds to achieve a shared standard of practical competency, thereby helping to close the worldwide industrial skills gap. Long-term sustainability and continuous upkeeping with constantly developing technologies and industry standards are ensured by the platform's design for modular content expansion and reuse.

Moreover, it is expected that the framework's ability to simulate a realistic industrial setting will enhance systemic thinking ability and strong situational awareness of graduates entering the workforce. Industrial environments' physical characteristics and dynamic operating challenges, such as production bottlenecks, system malfunctions and time limitations, are intended to be replicated by the highly precise simulations. Since learners must combine past knowledge, strategic reasoning and sensory data in order to solve problems, which are crucial indicators of proficient skills in complex situations, this exposure is expected to enhance decision-making and cognitive load management under pressure [[40], [41]]. Students are additionally trained in dispersed cognition and team coordination as crucial industry skills by the multi-user collaborative functionality [42]. A highly engaging learning environment that increases student enjoyment and motivation, leading to better learning outcomes, is created by integrating multimodal interaction through visual, auditory and tactile feedback [[9], [17]].

Lastly, the most significant outcome is the substantial reduction of operational risk. The framework acts as a vital risk-free testing ground where catastrophic failures and safety protocols can be experienced without real-world repercussions by enabling students to make mistakes and experience failure situations in a virtual setting. This firsthand experience of failure states and safety procedures is essential for creating a deeply rooted safety culture and procedural memory, thereby lowering the likelihood of accidents and expensive mistakes in actual industrial operations [43]. The AI system's objective performance data collection makes it possible to precisely identify deficient skills before they manifest in dangerous situations, ensuring that graduates join the workforce skilled, as well as risk-aware and safety-conscious [44]. By giving learners ongoing feedback and ensuring that their practical skills are firmly established before being applied in the real-world industrial setting, the AI-powered assessment module contributes to developing a future workforce that is more competent and safety-conscious.

7 Conclusion

This paper proposed a framework that combines VR, XR and AI in order to advance robotics education. With this approach and use of immersive simulations, students can practice in realistic and safe environments. In this way they can build various technical skills, and prepare for the demands of modern industry. The framework also helps in reducing the global skills gap because students have access to training that is practical, low-cost, and closer to real-world scenarios. The proposed learning process uses adaptive AI and constant interactions with students so that learning becomes more flexible, personalized, and engaging. Overall, this approach creates a strong foundation for connecting academic education with industrial needs.

In the future, the proposed framework should be tested in real academic courses to measure the impact on student skills and motivation. Also, there is a need for feedback from learners that gained knowledge with use of immersive simulations and used it in real industry environments. It will also be important to explore in more detail ethical and legal aspects, such as data privacy and accessibility for all learners. Further development of the proposed approach can include new robotics scenarios and integration with other fields like healthcare or mechanical engineering. By continuing this research, the framework can grow into a widely used tool for practical and safe training in many areas of education.

Acknowledgments. The underpinning project content inspiring this study is submitted for funding to MASTER 2nd open call for cascade funding. MASTER is a project co-funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Robotics Tomorrow: Addressing the Skills Gap in Manufacturing through Robotics Training. RoboticsTomorrow.
2. Owoc, M.L., Sawicka, A., Weichbroth, P.: Artificial intelligence technologies in education: Benefits, challenges and strategies of implementation, http://arxiv.org/abs/2102.09365, (2021).
3. Chiroma, H., Mohd Shuib, N.L., Abubakar, A.I., Zeki, A.M., Gital, A.Y., Herawan, T., Abawajy, J.H.: Advances in teaching and learning on Facebook in higher institutions. IEEE Access. 5, 480–500 (2017).
4. Yang, A.-M., Li, S.-S., Ren, C.H., Liu, H.-X., Han, Y., Liu, L.: Situational Awareness System in the Smart Campus. IEEE Access. 6, 63976–63986 (2018).
5. Kim, W.-H., Kim, J.-H.: Individualized AI tutor based on developmental learning networks. IEEE Access. 8, 27927–27937 (2020).
6. Kee, T., Zhang, H.: Digital experiential learning for sustainable horticulture and landscape management education. Sustainability. 14, 9116 (2022).
7. Neroni, M.A., Oti, A., Crilly, N.: Virtual Reality design-build-test games with physics simulation: opportunities for researching design cognition. Int. J. Des. Creat. Innov. 9, 139–173 (2021).
8. Young, G., Stehle, S., Walsh, B., Tiri, E.: Exploring virtual reality in the higher education classroom: Using VR to build knowledge and understanding. J. Univers. Comput. Sci. 26, 904–928 (2020).
9. Sung, B., Mergelsberg, E., Teah, M., D'Silva, B., Phau, I.: The effectiveness of a marketing virtual reality learning simulation: A quantitative survey with psychophysiological measures. Br. J. Educ. Technol. 52, 196–213 (2021).
10. Pande, P., Thit, A., Sørensen, A.E., Mojsoska, B., Moeller, M.E., Jepsen, P.M.: Long-term effectiveness of immersive VR simulations in undergraduate science learning: lessons from a media-comparison study. Res. Learn. Technol. 29, (2021). https://doi.org/10.25304/rlt.v29.2482.
11. Al-Azawei, A., Baiee, W.R., Mohammed, M.A.: Learners' experience towards E-assessment tools: A comparative study on virtual reality and Moodle quiz. Int. J. Emerg. Technol. Learn. 14, 34 (2019).
12. Ślósarz, L., Jurczyk-Romanowska, E., Rosińczuk, J., Kazimierska-Zając, M.: Virtual reality as a teaching resource which reinforces emotions in the teaching process. SAGE Open. 12, (2022). https://doi.org/10.1177/21582440221118083.
13. Kamińska, D., Sapiński, T., Aitken, N., Rocca, A.D., Barańska, M., Wietsma, R.: Virtual reality as a new trend in mechanical and electrical engineering education. Open Phys. 15, 936–941 (2017).
14. Nicolaidou, I., Stavrou, E., Leonidou, G.: Building primary-school children's resilience through a web-based interactive learning environment: Quasi-experimental pre-post study. JMIR Pediatr. Parent. 4, e27958 (2021).
15. Boetje, J., van Ginkel, S.: The added benefit of an extra practice session in virtual reality on the development of presentation skills: A randomized control trial. J. Comput. Assist. Learn. 37, 253–264 (2021).
16. Perez, R.P., Keleş, Ö.: Immersive virtual reality environments for embodied learning of engineering students, http://arxiv.org/abs/2503.16519, (2025).
17. Stracke, C.M., Bothe, P., Adler, S., Heller, E.S., Deuchler, J., Pomino, J., Wölfel, M.: Immersive virtual reality in higher education: a systematic review of the scientific literature. Virtual Real. 29, (2025). https://doi.org/10.1007/s10055-025-01136-x.
18. Bekteshi, L.: Education in the era of AI and immersive technologies - A systematic review. Journal of Research in Engineering and Computer Sciences. 3, 01–14 (2025).

19. XR4Human-SERVE 5.0, https://www.master-xr.eu/oc-project/xr4human-serve-5-0/, last accessed 2025/09/12.
20. Yan, W., Li, B., Lowell, V.L.: Integrating Artificial Intelligence and Extended Reality in language education: A systematic literature review (2017–2024). Educ. Sci. (Basel). 15, 1066 (2025).
21. Vassigh, S., Peterson, E., Bogosian, B., Tubella, J.: The Robotics Academy: An immersive learning game for training industrial robots, http://dx.doi.org/10.7275/YAX0-5Z24, (2022). https://doi.org/10.7275/YAX0-5Z24.
22. Kontra, C., Goldin-Meadow, S., Beilock, S.L.: Embodied learning across the life span. Top. Cogn. Sci. 4, 731–739 (2012).
23. von CAD-KAS PDF-Editor, V.M. der D. von C.-K.P.-E.V.M. der D.: Förderkreis der Angewandten Informatik, https://www.uni-muenster.de/imperia/md/content/angewandteinformatik/aktivitaeten/publikationen/enhancing_ai_tutoring_in_robotics_education_-_2024.pdf?utm_source=chatgpt.com., last accessed 2025/09/12.
24. Dietz, G., King Chen, J., Beason, J., Tarrow, M., Hilliard, A., Shapiro, R.B.: ARtonomous: Introducing middle school students to reinforcement learning through virtual robotics. In: Interaction Design and Children. pp. 430–441. ACM, New York, NY, USA (2022).
25. Yang, C., Luo, L., Vadillo, M.A., Yu, R., Shanks, D.R.: Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. Psychol. Bull. 147, 399–435 (2021).
26. Creed, C., Al-Kalbani, M., Theil, A., Sarcar, S., Williams, I.: Inclusive AR/VR: accessibility barriers for immersive technologies. Univers. Access Inf. Soc. 23, 59–73 (2024).
27. Khlaif, Z.N., Mousa, A., Sanmugam, M.: Immersive extended reality (XR) technology in engineering education: Opportunities and challenges. Technol. Knowl. Learn. 29, 803–826 (2024).
28. Sanfilippo, F., Blazauskas, T., Salvietti, G., Ramos, I., Vert, S., Radianti, J., Majchrzak, T., Oliveira, D.: A perspective review on integrating VR/AR with haptics into STEM education for multi-sensory learning. Robotics. 11, 41 (2022).
29. Mutis, I., Oberemok, M.: Haptic technology interaction framework in engineering learning: A taxonomical conceptualization. Comput. Appl. Eng. Educ. 33, (2025). https://doi.org/10.1002/cae.70009.
30. Johnson-Glenberg, M.C.: Immersive VR and education: Embodied design principles that include gesture and hand controls. Front. Robot. AI. 5, (2018). https://doi.org/10.3389/frobt.2018.00081.
31. Orsolits, H., Valente, A., Lackner, M.: Mixed Reality-based robotics education—supervisor perspective on thesis works. Appl. Sci. (Basel). 15, 6134 (2025).
32. Wang, F., Hannafin, M.J.: Design-based research and technology-enhanced learning environments. Educ. Technol. Res. Dev. 53, 5–23 (2005).
33. Miller, A.J., Kalafatis, S.: Mixed reality equipment training. In: Proceedings of the 2023 7th International Conference on Virtual and Augmented Reality Simulations. pp. 105–113. ACM, New York, NY, USA (2023).
34. Kopácsi, L., Karagiannis, P., Makris, S., Kildal, J., Rivera-Pinto, A., de Munain, J.R., Rosel, J., Madarieta, M., Tseregkounis, N., Salagianni, K., Aivaliotis, P., Barz, M., Sonntag, D.: The MASTER XR platform for robotics training in manufacturing. In: 30th ACM Symposium on Virtual Reality Software and Technology. pp. 1–2. ACM, New York, NY, USA (2024).
35. Huang, L.: Ethics of artificial intelligence in education: Student privacy and data protection. Sci. Insights Educ. Front. 16, 2577–2587 (2023).
36. Anders Ericsson, K.: Deliberate practice and acquisition of expert performance: A general overview. Acad. Emerg. Med. 15, 988–994 (2008).
37. Zimmerman, B.J., Schunk, D.H.: Handbook of Self-Regulation of Learning and Performance. Routledge. (2011).

38. VanLEHN, K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educ. Psychol. 46, 197–221 (2011).
39. Bailenson, J.: Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do. W. W. Norton & Company (2018).
40. Driskell, J.E., Salas, E., Driskell, T.: Foundations of teamwork and collaboration. Am. Psychol. 73, 334–348 (2018).
41. Kim, J., Hancock, P.: Modeling operator performance under stress and fatigue: What can a cognitive architectural model tell us? In: Advances in Human Factors and Ergonomics Series. pp. 1–10. CRC Press (2010).
42. Fiore, S.M., Rosen, M.A., Smith-Jentsch, K.A., Salas, E., Letsky, M., Warner, N.: Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts. Hum. Factors. 52, 203–224 (2010).
43. Reason, J.: The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries. Ashgate Publishing (2008).
44. Salas, E., Tannenbaum, S.I., Kraiger, K., Smith-Jentsch, K.A.: The science of training and development in organizations. Psychol Sci Public Interest. 13, 74–101 (2012).

„Play and Learn“,

Video Games and Digital Play in Early Childhood Education

Miljan Stevanović[1] [0009-0007-8912-5625], Petar Pejić[2][0000-0003-4155-8038], and Miloš Nikolić[2] [0009-0009-9023-9050]

[1] Faculty of Digital Arts, Belgrade Metropolitan University, Tadeuša Košćuška 63, 11000 Belgrade, Serbia
[2] Faculty of Information Technology, Belgrade Metropolitan University, Tadeuša Košćuška 63, 11000 Belgrade, Serbia
[2] Faculty of Information Technology, Belgrade Metropolitan University, Tadeuša Košćuška 63, 11000 Belgrade, Serbia
miljan.stevanovic@metropolitan.ac.rs
petar.pejic@metropolitan.ac.rs
milos.nikolic@metropolitan.ac.rs

Abstract. “Play and Learn”, Digital Video Game in Early Childhood Education, explores the possibilities of video game localization, digital game-based learning (DGBL) and software design, i.e., video games for early childhood education, for children aged 2 to 6. When designing a video game, a special focus is on the cultural and linguistic adaptation of the game. The game design is intended to use multiple models – to be based on story (narrative), audio and visual information in order to engage children whose native language is Serbian. The use of the native, local, language is important for a more accessible and understandable transfer of tasks and the interaction between the child and the video game itself. The adoption of DGBL technologies has transformed early childhood education by introducing interactive tasks adapted to the learning and age of children on mobile devices and mobile devices. The first part of this paper shows that educational video games improve children's motivation and creativity, and also encourage mutual cooperation in solving given problems. The second part of the paper presents the pedagogical framework adopted when designing the game. The third part answers the question of why cultural and linguistic localization is important. And the fourth part presents the preschool educational game "Play and Learn" (in Serbian “Igraj i uči”) that integrates the presented pedagogical principles and the importance of linguistic localization. integration of pedagogical principles, user-friendly interfaces and developmental appropriateness. Moreover, cultural and linguistic localization plays a key role in maximizing educational impact, because children learn best in their native language and therefore localized educational games can strengthen the child's personality, increase his understanding of the given problem and provide more interesting learning experiences.

Keywords: Educational video game, Early childhood education, Serbian Language localization.

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.338S
Part of ISBN: 978-86-89755-37-4

1 Developmental Benefits and Learning Outcomes of Digital Game-based Learning

Digital game-based learning (DBL) has positive effects on young children's development in several areas. A systematic review of educational video games by Alotaibi [1] concluded that early childhood game-based learning has moderate to large positive effects on cognitive, social, and emotional outcomes, and on children's motivation and engagement.

Educational videogames can:

• Strengthen children's cognitive skills – reasoning and problem-solving abilities;

• Games often exercise memory, attention, and logic. For example, digital play can foster early language skills, mathematical reasoning, and spatial cognition in preschoolers.

• Improve critical thinking, creativity, and problem-solving when children engage in well-designed learning games.

One literature review noted that DGBL has "has the potential to create new forms of childhood learning" [2].

In some games, we have digital guides that help the child in the process of playing and encourage team play. By solving joint tasks in games, children develop social intelligence – connecting with peers, respecting the rules of turn-taking and resolving conflict during the game [3]. It is also desirable that certain elements of the game, in addition to developing social intelligence, be designed to develop emotional intelligence in children. For example, interactive stories (similar to classic children's fairy tales) can foster skills in managing emotions, expressing themselves and understanding each other. Papoutsi et al., evaluated the serious game "The Park of Emotions" and experimentally validates as an effective intervention tool, highlighting the importance of implementing digital games aimed at improving emotional intelligence and emotional skills throughout childhood and preadolescence [4]. Engaging, playful context of games often keeps children motivated and persistently involved, which may further support emotional regulation. Researchers have observed that children can be more engaged and stay on-task longer during digital play than in some traditional activities [5].

Digital games can contribute to both motor skills and physical activity if the development of the games was properly designed. Many preschool games require tapping, dragging, or drawing on touchscreens, which can help refine fine motor control and hand–eye coordination—children with frequent tablet use in early toddlerhood showed significantly better performance on fine motor assessments than non-users [6]. Some evidence suggests that interactive tablet play can improve these skills, though findings are mixed and dependent on usage patterns. Notably, active video games encourage physical movement. One study found that exergaming, that refers to active video games that are also a form of exercise, encourage physical movement in preschool-aged children [7]. Video games or applications for mobile phones that are based on movement (dancing, exploring nature, or applications that help with exercise) require children to actively, physically, participate in the tasks.

Educational games are designed to, in addition to physical engagement, also strengthen early skills such as writing and numeracy. For example, game-based literacy applications have successfully supported the development of reading skills in preschool children, in accordance with the cultural and other requirements of the specific target environment [8]. Math-oriented games can teach counting, shapes, and basic problem-solving in an engaging format [9]. Research indicates that well-designed games can improve young children's knowledge in areas like vocabulary, letter recognition, and basic math, often by embedding learning objectives within play tasks [10]. Some studies have specifically targeted creativity—for instance, a digital game called Thinking Paradise was developed to train preschoolers' creative thinking, yielding improvements across fluency, originality, elaboration, anti-block, and title scores in the Torrance Test [11]. Game-based learning must be aligned with the pace of children's development as well as with the curriculum and program for preschool children. An interactive approach to teaching, unlike the traditional one, can contribute to greater success in mastering the material and greater engagement of children. Video games can make learning fun and interactive, because the process itself is seen as a form of entertainment.

The design of video games in early childhood must be of high quality and have a clear context. The emphasis should be on purposefulness, that they be appropriate to the child's developmental level and that they be used according to the instructions of the educator. Educators have a role in selecting appropriate educational content, encouraging teamwork, but also in being aware of the needs of each individual child and, based on those needs, scaling the difficulty of the game. With these supports in place, the consensus of recent literature is that digital play can be a valuable tool to promote children's learning and development in the early years [[12], [1]].

2 Psychological Frameworks: Piaget and Vygotsky

Classic child development theories provide a useful lens for understanding why play – including digital play – is so critical in the preschool years. Both Jean Piaget and Lev Vygotsky viewed play as a fundamental, "building block" activity through which children construct knowledge. As Piaget emphasized, intelligence is an adaptation that develops through organization and assimilation of experiences, with play acting as a key mechanism for symbolic capacity and cognitive growth [13]. Similar to Piaget, Vygotsky argued that play allows children to act "beyond their average age" by engaging in activities that expand their abilities with social support (with the help of educators).[14] The process of play (with physical toys or video games) allows young children to experiment, as they have complete freedom to interpret the world in their own way. This process of cognitive maturation is essential for later skills such as reading, writing, or arithmetic.

Piaget's theory deals with the progression of children through different stages of cognitive development – for example, preschoolers (aged around 2–6 years) are in the preoperational stage, which is characterized by intuitive thinking and symbolic play [13]. At this stage, children learn best through concrete, sensory experiences and active exploration. Piaget explained that intelligence builds upon sensorimotor activity and gradually develops into operations that allow for logical thinking [13]. Digital video games, in order to be in line with Piaget's principles, should have visually appealing content and the possibility of direct manipulation of the game content that would suit the child's developmental level. For example, the game can use basic and strong colors, clear and stylized drawings, and the game interface itself should be intuitive, i.e. clear enough for the child to use. All of these elements should attract the attention of young children. Piaget also defined assimilation and accommodation as two key processes by which children learn: they assimilate, adopt, new experiences into existing frameworks and adapt to their own frameworks when new experiences challenge previous structures [13]. Educational video games can support both assimilation and accommodation by allowing children to repeat actions, experiment, and receive feedback. As Piaget states - children are "little scientists" who learn through trial and error because the interactive and exploratory nature of educational video games enables them to do so.

Vygotsky adds that social interaction is also important for cognitive progress. He introduced the concept of the Zone of Proximal Development (ZPD) – the range of tasks a child cannot do alone but can accomplish with guidance from a more skilled partner [14]. Play creates an ideal context for such guided participation. In his view, during play “a child always behaves beyond his average age,” stretching to new skills with the help of others. This idea of scaffolding is central: when a caregiver or teacher plays alongside the child, asks questions, or gives hints, digital play becomes a scaffolded activity that pushes the child's learning forward. Van der Veer [15] explains that Vygotsky saw language and thought as merging in early childhood, where children first repeat instructions aloud before internalizing them as inner speech. This process transforms play into a powerful context for problem solving and self-regulation.

There is evidence that well-designed educational software can facilitate this adult-child interaction. For instance, literacy games such as Alphablocks show, a British preschool television series featuring animated letter-characters in Alphaland, that when adults guide children during gameplay, it creates a “meeting of minds” where both builds meaning together. This directly reflects Vygotsky's idea that learning is socially co-constructed [15]. Role-playing games – whether in a physical or digital environment – allow children to practice following rules, taking on roles, and abstract thinking, which Vygotsky linked to the development of higher mental [14] functions.

Piaget suggests that educational games should be concrete, but also intuitive and exploratory to suit the developmental level of preschool children, while Vygotsky emphasizes social interaction and the role of language in learning. When designing educational video games, both views should be taken into account: they should be interactive and encourage independence in the child, but also require teamwork and adult participation. This blend maximizes learning within the child's developmental stage and translates foundational child development principles into practical digital learning design.

3 Cultural and Language Localization of Video Game

When designing video games for use in early childhood education, the cultural and linguistic context of children's play should be taken into account. Research has shown that young children learn best in their native language and within the family setting. UNESCO notes that “evidence tells us that learning first in one's mother tongue leads to better outcomes in the future – for individuals, cultures, and nations" [16]. In practical terms, this means an educational game for Serbian preschoolers will be most effective if its language is Serbian and it reflects Serbian cultural context. Using the child's mother tongue not only improves comprehension and learning of content, but also affirms the child's identity and comfort, making the learning experience more natural [[17], [18]].

For young children, who have not yet developed basic language skills, games should provide narration, accompanying audio instructions, and text in their native language. For example, if a game was first developed for an English-speaking area, it is easiest to simply translate the words into Serbian. However, effective localization goes further - it can include using common Serbian nursery rhymes, local names, or voices with familiar accents to make the game seem like the source. The goal is to have the child engage with the content without language being a barrier. This is supported by guidelines in early education that initial literacy and learning activities be conducted in the child's first language whenever possible [8]. For example, a game that teaches letters would be designed around the Serbian alphabet – in Cyrillic and Latin script for children from the Serbian-speaking area. It should also include specific characters Đ, Č, Š, etc., and how they would be pronounced by a Serbian teacher or parent.

Stories, characters, and overall visuals are often the most important elements of children's play. These elements should be aligned with the child's cultural background as they can significantly increase their engagement. The story, themes, and characters that appear in the games should be inspired by literary works by recognized local authors or from historical sources and folk tales. For example, a game for children from the Serbian-speaking area may include: local animals (such as hedgehogs or bears that are common in Serbian children's stories), traditional motifs or folkloric characters, or scenarios such as the setting of a "Slavic fairy tale" that Serbian children may recognize from books and cartoons [[19], [20]]. The use of local cultural sources can make the game more interesting for the child, as they recognize parts of the world around them.

At the same time, designers should be mindful of cultural neutrality vs. specificity. A study by Nikolopoulou [8] on early childhood software localization pointed out that some features of software are easier to localize than others. Software with a "dominance of pictures, animation and sound, [and] culture-independent content" and with neutral interfaces and storylines can be more readily adapted to different countries. In other words, using universally appealing imagery (like basic shapes, happy cartoon animals) and avoiding highly culture-specific references can ease localization to multiple languages. However, to truly connect with a target culture, certain customizations are beneficial – such as including a few locally loved songs or cultural references that make the experience unique for that audience. It's a balance: core educational content might remain the same globally (e.g. counting apples), but the style and context can be tweaked to fit local expectations and curricular needs. For example, if a game teaches about foods, a Western version might use apples and pizza, whereas a localized Serbian version might include plums and pita bread to reflect foods children recognize [[17], [18]].

Localization in educational games for early childhood is not just translation – it's cultural adaptation. Ensuring the game speaks the child's language (literally and figuratively) can improve both learning outcomes and enjoyment. As a child's right, early education in the mother tongue is recommended, and this extends to digital learning tools. Game developers and educators should thus collaborate to create or adapt games that align with local languages and cultures [[21], [10]].

4 „Play and Learn“

The educational video game "Play and Learn" (in Serbian "Igraj i uči") was developed at the Metropolitan University, center in Niš, within the Laboratory for Video Games during the year 2025 and is continuously being developed. The game is designed as a modular platform, which will be developed in the coming years and will represent a long-term and creative project of the laboratory. The development team consisted of associate professor dr. Petar Pejić, assistant Miloš Nikolić and a group of fourth-year students from the Faculty of Information Technologies (study program: Video Game Development) and the Faculty of Digital Arts (study program: Graphic Design). The collaboration between students from two different faculties is reflected in the integration of the technical, software, side of the project with original visual and audio design. Joint work, which is part of the students' professional practice, contributed to the successful realization of the video game "Play and Learn".

The video game was developed in the Unity game engine and is optimized for Android mobile devices, making it easily accessible to preschoolers. The entire visual identity (see Fig. 1) was originally created by students of the Faculty of Digital Arts. The narration and music in the game have also been adapted to provide an authentic and culturally localized experience in the Serbian language. The game has a two-dimensional visual style that is suitable for children, with colorful and rounded shapes and interactive elements that are suitable for children aged 2 to 6 years (see Fig. 2). Game design emphasizes simplicity, recognition and positive emotional engagement. "Play and Learn" is organized as a modular island system. Each island represents a different thematic unit (e.g. pirates, winter) and contains 6 to 12 minigames. This structure provides a variety of experiences while maintaining a clear pedagogical framework and lends itself to future content expansion.

Fig. 1. Visual identity of video game "Play and Learn".

Fig. 2. Two-dimensional visual style that is suitable for children of video game "Play and Learn".

Children interact primarily through the mechanisms of dragging, arranging and selecting objects. These mechanics support the development of fine motor skills and logical thinking. Currently, there are three basic types of mini-games: assembling shapes (children drag and join geometric figures into correct patterns, see Fig. 3); shape matching and color matching (children group objects based on color associations); and order of sizes (children arrange objects according to relative size which encourages pattern recognition). All these tasks are intuitive, not non-verbal, and are aligned with the developmental levels of preschool children.

Fig. 3. Drag and join geometric figures into correct patterns suitable.

Unlike most preschool games that focus on mastering language and numeracy, "Play and Learn" emphasizes: development of motor skills (fine hand-eye coordination through mechanisms of dragging and arranging elements); pattern recognition (shape, size, color); and cognitive flexibility and problem solving (choosing correct schedules).

The game is based on the theories presented by Piaget and Vygotsky. According to Piaget's theory, children aged 2 to 6 learn best through concrete tasks, sensory and symbolic play. Game also supports (as already described in the paper) the process of assimilation and accommodation where the child learns by repeating actions, experimenting with solutions and having the availability of immediate feedback. From a Vygotsky point of view, the design also supports the Zone of Proximal Development – children can perform tasks independently, but also adults can guide them through their mastery.

The game is currently free to download and available for families and educators. Pilot testing with preschool children is planned, which will provide empirical feedback on the effectiveness of the game and on future improvements. The game is conceived as a laboratory platform that future generations of students will build on, with new islands and mini-games, as well as new visual identities for those parts of the game. The upcoming expansion will include additional themes, new mechanics and potential localization into other languages if the game proves successful in Serbia.

5 Conclusion

"Play and Learn", Video Games and Digital Play in Early Childhood Education is an analysis of video game-based learning and its localization. The paper highlights the potential of video games as a tool for education, and their fun and interactive approach to imparting knowledge and skills. Video games, viewed as multimodal products, require careful translation and localization to ensure accessibility and engagement of children. The prevalence of mobile phones and tablets makes digital learning based on video games accessible to a wider group of users. Thanks to well-designed mechanics, these games can improve children's motivation, stimulate their creativity, and develop problem-solving and collaboration skills. Of course, content control by parents or educators is always needed, because due to their design, problems such as overuse must be taken into account.

Cultural and linguistic localization, in our case into the Serbian language, appear as a key factor in improving the educational value of preschool games. This requires the integration of local languages in which the video game will be used, the adoption of local narrative and cultural elements in the game design.

The video game "Play and Learn" is designed for educational purposes in early childhood. It combines well-thought-out pedagogical design, child-friendly interfaces and cultural and linguistic localization of the Serbian-speaking area. Collaboration between experts (educators, psychologists, linguists and game designers) allows for the creation of engaging, developmentally appropriate and culturally relevant digital learning environments.

Acknowledgments. We would like to thank a group of fourth-year students from the Faculty of Information Technologies (study program: Video Game Development) and the Faculty of Digital Arts (study program: Graphic Design) who participated in the realization of the video game "Play and Learn".

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Manar S. Alotaibi, Game-based learning in early childhood education: a systematic review and meta-analysis. Front. Psychol. 15:1307881. doi: 10.3389/fpsyg.2024.1307881 (2024)
2. Behnamnia, N., Kamsin, A., Ismail, M. A. B., and Hayati, S. A. A review of using digital game-based learning for preschoolers. J. Comput. Educ. 10, 603–636. doi: 10.1007/s40692-022-00240-0 (2022)
3. Garaigordobil M, Berrueco L, Celume MP. Developing Children's Creativity and Social-Emotional Competencies through Play: Summary of Twenty Years of Findings of the Evidence-Based Interventions "Game Program". J Intell. 2022 Oct 2;10(4):77. doi: 10.3390/jintelligence10040077. PMID: 36278599; PMCID: PMC9590021. (2022)
4. Papoutsi, C.; Drigas, A.; Skianis, C.; Skordoulis, C.; Pappas, M. “The Park of Emotions”: A Serious Game for the Development of Emotional Intelligence in Children. Appl. Sci. 2024, 14, 6067. https://doi.org/10.3390/app14146067 (2024)
5. Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. Putting education in “educational” apps: Lessons from the science of learning. Psychological Science in the Public Interest, 16(1), 3–34. https://doi.org/10.1177/1529100615569721 (2015)
6. Souto PHS, Santos JN, Leite HR, Hadders-Algra M, Guedes SC, Nobre JNP, Santos LR, Morais RLS. Tablet Use in Young Children is Associated with Advanced Fine Motor Skills. J Mot Behav. 2020;52(2):196-203. doi: 10.1080/00222895.2019.1602505. Epub 2019 Apr 22. PMID: 31007146 (2020)
7. Zan Gao et al., Effects of exergaming on motor skill competence, perceived competence, and physical activity in preschool children, Journal of Sport andHealth Science, https://doi.org/10.1016/j.jshs.2018.12.001 (2018)
8. Nikolopoulou, K. Early Childhood Educational Software: Specific Features and Issues of Localization. Early Childhood Education Journal. 35. 173-179. doi: 10.1007/s10643-007-0168-5. (2007)
9. DePascale, M. & Ramani, G. The Role of Math Games for Children's Early Math Learning: A Systematic Review. doi: 10.31234/osf.io/xkvhb. (2024)
10. Stephen, C., & Plowman, L. Information and Communication Technologies in Pre-school Settings: a review of the literature. International Journal of Early Years Education, 11(3), 223–234. https://doi.org/10.1080/0966976032000147343. (2003)
11. Xiong, Z. & Liu, Q. & Huang, X. The influence of digital educational games on preschool Children's creative thinking. Computers & Education. 189. 104578. doi. 10.1016/j.compedu.2022.104578. (2022)
12. Li, H., He, H., Luo, W. et al. Early Childhood Digital Pedagogy: A Scoping Review of Its Practices, Profiles, and Predictors. Early Childhood Educ J. https://doi.org/10.1007/s10643-024-01804-8 (2024)
13. Piaget, J. The origins of intelligence in children. New York: International Universities Press. United States of America. (1952)
14. Vygotsky, L. S. (1935/1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press. ISBN 0--674-57629-2, Printed in the United States of America. (1978)
15. van der Veer, R. Vygotsky's Theory. doi: 10.1002/9781119171492.wecad101. (2020)
16. Bühmann, D., Trudell, B., Mother tongue matters: Local language as a key to effective learning. Paris: UNESCO, https://unesdoc.unesco.org/ark:/48223/pf0000161121. (2008)

17. Hamari, J. & Koivisto, J. & Sarsa, H. Does Gamification Work? — A Literature Review of Empirical Studies on Gamification. Proceedings of the Annual Hawaii International Conference on System Sciences. doi: 10.1109/HICSS.2014.377. (2014)
18. Deterding, S. & Dixon, D. & Khaled, R. & Nacke, L. From Game Design Elements to Gamefulness: Defining Gamification. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011. 11. 9-15. doi: 10.1145/2181037.2181040. (2011)
19. Jakovljević, T., Games for Children as a Model of Folklore Communication, City Museum of Novi Sad, Serbia, https://hrcak.srce.hr/en/file/68729. (2009)
20. Zhu, Y., Zhou, R. & Zhang, Y. Designing preschool children's educational games for enlightenment through decision analysis methods. Multimed Tools Appl 83, 78331–78360. https://doi.org/10.1007/s11042-024-19803-7. (2024)
21. Haugland, S. W., & Wright, J. L. Young Children and Technology: A World of Discovery. Boston, MA: Allyn & Bacon. ISBN: 978-0205175246 (ISBN-10: 0205175244). (1997)

The Illusion of Understanding: Digital Empathy and Emotional Mimicry in AI-Driven Education

Tamara Vučenović[1[0000-0003-4152-4372]] and Marija Maksimović[2[0009-0006-6013-079X]],

[1,2] Belgrade Metropolitan University, Serbia
tamara.vucenovic@metropolitan.ac.rs
marija.maksimovic@metropolitan.ac.rs

Abstract. The phenomenon of increasing algorithmic comfort can be described as a process in which apparent security turns into a serpent embrace-like pressure, with increasing automation producing a tension between comfort and suffocation, as well as a dynamic of resistance, adaptation, and release. In the age when education is shaped according to code, the rule prevailed: *"no more endless hours creating something."* Efficiency has become the new holy grail, but it is increasingly being forgotten that this convenience comes at a price – the loss of creative effort and human authenticity.

Within the Education 4.0 paradigm, contemporary education strives to develop the skills of the future – creativity, critical thinking, and digital literacy – while at the same time increasingly relying on large language models (LLMs) like ChatGPT. However, we cannot help but wonder if this approach, while opening up possibilities for new forms of learning, actually leads to a superficial automation of the process that ignores the key component of learning – the emotional dimension. The emotional mimicry that large language models achieve through imitating empathy and using apparently empathic phrases actually remains on the surface, as confirmed by research showing that students who use LLM tools record less cognitive effort, as well as a lower quality of reasoning – which indicates emotional shallowness and thought passivation.

The findings of a survey conducted on a sample of 60 students show that young users simultaneously recognize and support the practical benefits of AI tutors, yet express concerns about cognitive addiction and illusory empathy. This study examines the consequences of LLM tutor-based educational practices on thought identity, emotional depth, and confidence development. The findings, including the results of the survey, indicate that interaction with algorithmic tools reduces creative and semantic brain activity, encouraging superficial automatism and digital dementia, i.e., a decline in memory, concentration, and critical reasoning. In response, the paper proposes digital emotional literacy as a key pedagogical strategy and safeguard against the manipulation of empathy, the addictive reliance on tutors, and the substitution of emotional reality for algorithmic comfort.

Keywords: Artificial Intelligence in Education, Emotional Intelligence, Simulation of Empathy.

Research Paper
DOI: https://doi.org/10.46793/eLearning2025.348V
Part of ISBN: 978-86-89755-37-4

1 Introduction

In the context of the growth of emotional mimicry and reliance on algorithmic tutors, the emergence of digital dementia is particularly worrying - a phenomenon of reduced memory, focus and critical reasoning due to the increasingly frequent delegation of thought processes to technology. We came to this phenomenon through the daily use of artificial intelligence and undermining its penetration into all segments of not only business, but also education. In this way, we created a symbiotic, increasingly unbreakable bond with smart tutors.

Due to the rapid penetration of generative AI systems into education, digital natives stand out among the most affected groups [1] - generations born after 1980, who grew up in the environment of the Internet and digital technologies. Although they are fluid in the use of technology, according to Prensky [1], their excessive reliance on digital tools can lead to lower emotional literacy, as it favors quick, superficial interactions instead of deep interpersonal connections, which limits the development of empathy and emotional depth. Those with less emotional and digital literacy in the group of digital migrants are the most vulnerable and always at greater risk [2].

The risk is reflected in the unpreparedness to face the challenges that artificial intelligence brings, and due to the lack of developed skills, they often find themselves in a foggy ability to distinguish simulated empathy from the real one. The explanation is simple: lacking these skills, they are less resistant to errors and illusions of understanding and less able to distinguish the simulated empathy of such algorithms from real empathy [2].

Various authors have warned that phenomena such as simulated empathy in AI carry the risk of emotional manipulation of the user. Although an advanced chatbot may appear empathetic, it has no real feelings behind the words which Kurian [2] describes as empathy gap, i.e. lack of emotional empathy, and even the opposite of empathy, due to the misleading effect on the user. Studies have documented cases of humans perceiving AI companions as friends or partners [3], which raises the ethical question of reciprocity - the relationship is always one-sided and inauthentic [4].

This can lead to a reduction in human interaction or an impairment in the understanding of real emotions [3]. Accordingly, in the next phase of AI development, the key question is how (and whether) machines can go beyond mere imitation and achieve some form of true emotional intelligence instead of programmed emotional performance [2]. If this is not possible, the imperative of transparency is imposed - the user must know how to communicate with the algorithm in order to avoid unethical manipulation of emotions [4].

The combination of that emotional insensitivity and digital credulity (trusting the algorithm) makes young people especially vulnerable. In short, *"empathy among algorithms"* is currently a one-way street: algorithms imitate empathy, and people (especially children) project it onto them. However, that illusion can cost a young person's development dearly if it is not addressed through education and protective measures.

Further in the paper, we consider the occurrence of emotional mimicry in AI tutors and its implications.

2 Literature Review

People inherently have tendency towards emotional mimicry by spontaneously imitating the expressions of feelings of others. For example, if the interlocutor seems sad or worried, we often adopt a similar facial expression or tone of voice ourselves. This unconscious imitation functions as a social "glue": by harmonizing emotions, communication participants confirm that they share a similar state, which strengthens empathy and trust [5]. Interestingly, research shows that people show similar reactions when their interlocutor is a machine. The theory "computers as social actors" [6] showed that users unconsciously apply social norms when interacting with computers. Despite knowing that a computer has no feelings, we respond to a kind voice or message with politeness and compassion, almost as if we were communicating with a person. This phenomenon - anthropomorphizing, i.e. attributing human characteristics to non-human entities - explains why a digital assistant that responds with "I understand, I'm here to help" can provide some comfort to the user. Children and teenagers especially easily perceive digital characters (chat-bots, voice assistants) as quasi-friends - they address them in a personalized way, trust them and expect understanding from them, often not making a clear distinction between simulation and real empathy [7]. The human brain responds to such "social signals"—a warm tone of voice or caring words—similarly whether they come from a human or an algorithm, as long as the signals are believable [6].

The phenomenon known as the "Google effect" or the so-called digital amnesia, indicates that the availability of information only a click away from us changes the way we remember it. When we know they are readily available, we are less likely to remember them permanently [8]. In an educational context, AI tutors further relieve students by taking on routine cognitive load, which can allow focus on more complex tasks [9]. However, research points out that with excessive reliance on such systems, cognitive dependence occurs, followed by a decrease in independent analytical and critical thinking [10].

German neurologist Manfred Spitzer [11] warned that such patterns lead to the phenomenon of digital dementia, a condition in which the excessive use of digital devices leads to impaired memory, poor focus and learning disorders, similar to the symptoms of dementia in old age. Recent research confirms that early and chronic exposure to screens can negatively affect the structure and function of the brain, increasing the risk of cognitive and emotional disorders in young people [12]. In other words, the brain accustomed to constant and fast digital stimuli becomes lazy in memory and focus, which further illuminates the risk of cognitive passivation and loss of autonomy in the era of AI tools.

Can AI "outdo" human empathy? Children as the most vulnerable digital natives who grow up in the center of the empathic gap - must learn where the alive ends and the algorithmic begins. Research shows that they often perceive chatbots as "quasi-friends" and trust them with their feelings, even though algorithms have neither a moral compass nor real empathy [2]. Dangerous examples from practice, such as Alexa or Snapchat AI, show that such trust can be risky [13]. The youngest members of society are especially susceptible to these challenges. Their developing brains can be influenced by the increasing amount of time spent on digital devices. Increasing digital dementia - exposure to a large amount of information and distractions - has a negative impact on our ability to focus, think deeply and remember [14]. Paradoxically, sometimes AI responses are rated more empathetic than human responses [15, 16], which illuminates how easily we mix simulation and reality.

3 Methodology and results

The research was conducted through an anonymous online survey (Google Forms) [17]. in September 2025. The sample consisted of 60 students of various study programs aged 20–25. All respondents had previous experience with AI tools, most often ChatGPT. The questionnaire combined closed and open questions where the quantitative data was analyzed descriptively, while the open answers were processed by thematic analysis.

Key findings

- **Frequency of AI use** - 47% of students use AI almost daily, 30% occasionally, 17% rarely, and only 7% never use it.
- **Experiencing AI empathy** - More than half (55%) rate AI messages as mechanical and without impact, while 35% admit that they still like them and find them [potentially motivating.
- **Comparison with lecturers** - 57% of students stated that at least sometimes the AI seemed more pleasant or understandable to them than the lecturer; only 27% say they have never felt it.
- **Trust in AI statements of understanding** - Half (47%) partly believe phrases like "I understand that it is difficult for you", but a third do not believe them at all, while only a small number perceive them as sincere.
- **Digital emotional literacy** - 67% of students think that education about emotional literacy should be introduced in the digital environment, while 22% remain undecided, and 8% think that it is not necessary. In the open answers, it was suggested that this topic should be included at least in a few lessons and that the lecturers should first develop these competencies themselves.

4 Discussion

The results show that students recognize the paradox of AI tutoring: on the one hand, they perceive AI as more approachable and patient than some lecturers, and on the other hand, they clearly distinguish programmed kindness from real human empathy. This double awareness points to the existence of an intuitive digital emotional literacy - the ability to enjoy the benefits of AI, without fully trusting its "empathy".These findings can also be viewed through the concept of a digital 'filter bubble' [14], where algorithms shape our experience of reality so that it increasingly resembles what we already think and feel. This narrows the space for dealing with different perspectives, which further complicates the development of empathy and critical thinking.

Secondly, students are aware that AI reduces cognitive effort, but also the risk of passivation: it is easier to ask than to think. Such self-reflection is valuable - it shows that young people are aware of the "trade-off" effect of modern technology: they gain efficiency, but lose some of the exercise. These results strongly match the concept of cognitive unloading [18], and the theses of Spitzer [11], about the decline of motivation for deeper thinking.

Interestingly, even those skeptical of AI "empathy" see value in learning about it; their caution is precisely the argument that such literacy should be widely developed. Such attitudes are practically a call to action for the educational system. While digital literacy has traditionally been associated with technical skills, students clearly signal that education must also include the social-emotional dimension of digital intelligence.

In addition to quantitative findings, students' open-ended responses provide valuable insights into their attitudes and feelings. They supplement the figures with personal experiences, dilemmas and suggestions, revealing how young people perceive the presence of AI in learning – not only as a tool, but also as a phenomenon that changes their habits, expectations and way of thinking.

Qualitative insights from respondents

"AI-generated emotion makes the interaction more interesting, but I mainly use it for practical purposes."

"AI is effective on its own in many areas. I believe it supports both creative and empathic work."

"I believe that we rely too heavily on AI and that it may negatively affect certain professions."

"It can be addressed in one or several classes, but it should not take too much time away from core courses."

"We should learn how to use AI—for example, as a tool for information gathering—but its analyses should not be taken for granted, and critical thinking should never be abandoned."

"Lecturers need to be trained in emotional simulation and interface psychology in order to prepare students for digital reality."

"Illusion is a sufficiently accurate, yet incomplete term. Its meaning depends on numerous factors, which is why even this concept may feel like an illusion today."

Graphical representation of results

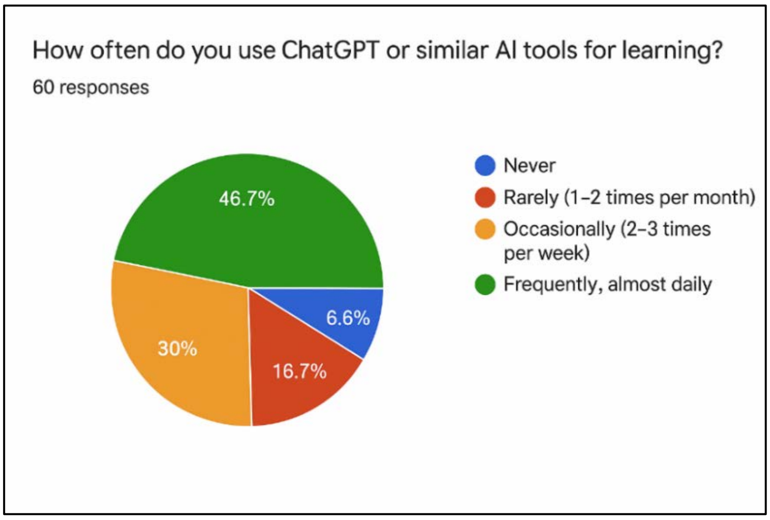


Fig. 1. Frequency of AI tutor use among students (N = 60). The majority report daily or occasional use, while only 7% have never used AI.

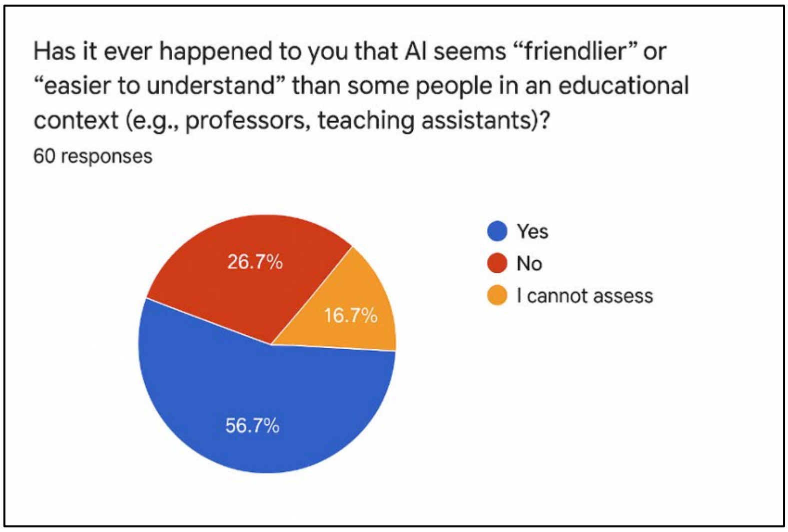


Fig. 2. Perceptions of AI empathy compared to lecturers. More than half of the students state that AI has sometimes appeared more pleasant or understanding.

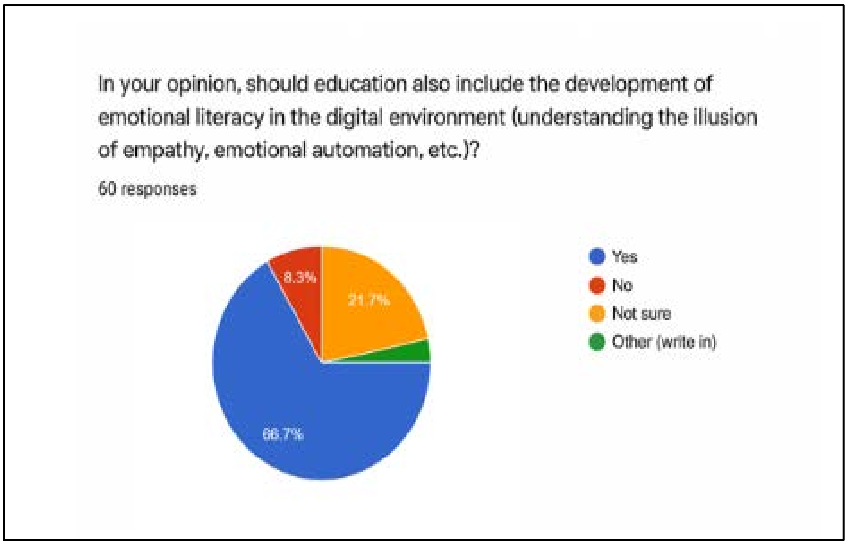


Fig. 3. Attitudes toward introducing digital emotional literacy. Approximately two-thirds of students support this initiative.

5 Conclusion

Our mini-research in the form of a survey showed that students have already recognized both the strengths and weaknesses of artificial intelligence in learning. AI helps them progress through the syllabus more quickly, but it often encourages shortcuts and reduced effort. Paradoxically, many perceive the "empathy" of chatbots as more pleasant than human empathy, even though they are aware that it is only an algorithmic illusion.

This double awareness, simultaneous trust and doubt, opens up an important space for education to assume a new responsibility: developing digital emotional literacy and introducing training on how to use AI without losing one's own thinking and emotional capacities. Continued research and practice in this area are essential. Empathy, creativity, and authenticity must remain at the core of education, as these are priceless human skills that no algorithm will ever be able to fully replicate.

If neglected, there is a risk that younger generations will become hostages of "digital dementia"—accustomed to the numbing comfort of quick answers, but increasingly weakened in memory, attention, and critical thinking. The algorithmic "filter bubble" further deprives us of emotional expression and contributes to a weakening of empathy. If students grow up in such a closed environment, we risk raising a generation that will find it increasingly difficult to engage in independent thinking.

Therefore, it is crucial that the educational system does not turn into a uniform army of AI pioneers, but instead into a space of free choice and critical inquiry. Not all ease is beneficial: sometimes it is precisely the difficulty of the path that shapes our authenticity. If we forget this, we risk becoming mere pieces on the board of technological progress pawns moved by someone else's rules, rather than active players with our own value and meaning.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Prensky, M. (2001). Digital natives, digital immigrants. On the Horizon, 9(5), 1–6.
   https://doi.org/10.1108/10748120110424816
2. Kurian, N. (2023). AI's empathy gap: The risks of conversational Artificial Intelligence for young children's well-being and key ethical considerations for early childhood education and care. Contemporary Issues in Early Childhood, 26(1), 1–8. https://doi.org/10.1177/14639491231206004
3. Ta, V., Holländer, A., & Krämer, NC (2021). Too much information? Exploring the relationship between information disclosure and the development of parasocial relationships with virtual agents. Computers in Human Behavior, 120, 106756. https://doi.org/10.1016/j.chb.2021.106756
4. Scheutz, M. (2012). The inherent dangers of unidirectional emotional bonds between humans and social robots. In P. Lin, K. Abney, & GA Bekey (Eds.), Robot ethics: The ethical and social implications of robotics (pp. xx–xx). MIT Press.
5. Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893
6. Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153
7. Rutledge, P. B. (2025, June 24). Kids and chatbots: When AI feels like a friend. *Psychology Today.* https://www.psychologytoday.com/intl/blog/positively-media/202506/kids-and-chatbots-when-ai-feels-like-a-friend
8. Storm, BC, Stone, SM, & Benjamin, AS (2017). Using the internet to access information inflates future use of the internet to access other information. Memory, 25(6), 717–723. https://doi.org/10.1080/09658211.2016.1210171
9. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Boston, MA: Center for Curriculum Redesign.
10. Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 26(2), 582–599.
11. Spitzer, M. (2012). *Digitale Demenz: Wie wir uns und unsere Kinder um den Verstand bringen.* München, Germany: Droemer Knaur

12. Kanbaj, A., Demir, H., & Yilmaz, F. (2025). Digital dementia: Early exposure to digital media and cognitive risks in adolescents. Journal of Neurological Sciences, 42(1), 55–67.
13. Paul, K. (2023, May 3). Snapchat's My AI chatbot offers unsafe advice to teens, investigation finds. The Guardian.https://www.theguardian.com/
14. Vučenović, T., & Stojanović, M. (2024). Digital Communications: Management, Marketing Strategies and Practical Examples. Novi Sad: Academic book.
15. Ayers, JW, Poliak, A., Dredze, M., Leas, EC, Zhu, Z., Kelley, JB, Faix, DJ, Goodman, AM, Longhurst, CA, & Hogarth, M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Internal Medicine, 183(6), 589–596. https://doi.org/10.1001/jamainternmed.2023.1838
16. Chen, D., Leung, C., Chen, B., Miao, C., & Barnes, B. (2024). Physician and artificial intelligence chatbot responses to cancer patient questions posted on a social media forum: A comparison of empathy and quality. *JAMA Oncology.* Advance online publication. https://doi.org/10.1001/jamaoncol.2024.0836
17. https://docs.google.com/forms/d/e/1FAIpQLSc6SDCNyuF5xGg3o3eXcqNI7EtiAn-jVoopWvBMGvBXk7-S6w/viewform
18. Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science, 333*(6043), 776–778. https://doi.org/10.1126/science.1207745

Current Trends in Social Entrepreneurship Education on Massive Open Online Courses platforms

Irena Konecki[1][0000-0002-3190-6441]

[1] Faculty of Organization and Informatics University of Zagreb, Croatia
irena.konecki@foi.unizg.hr

Abstract. This paper examines current trends in social entrepreneurship education (SEE) as delivered through Massive Open Online Course (MOOC) platforms in North America. The analysis identifies dominant themes in course content as well as the teaching methods employed. The most frequently addressed topics include impact measurement, innovation processes, resource acquisition, scaling, and sustainability. All of the courses analyzed incorporate video lectures and assigned readings, with all but one integrating detailed case studies. Guest lectures and structured discussion prompts emerge as the next most frequently employed teaching methods. A comparative assessment with face-to-face SEE courses highlights the need for the development of instructor-led online formats that incorporate more complex assignments subject to expert evaluation. By presenting the range of pedagogical tools used across MOOCs to address similar themes, this study provides valuable insights for educators seeking to update or redesign both face-to-face and online SEE curricula.

Keywords: Social Entrepreneurship, Social Entrepreneurship Education, Massive Open Online Courses, Teaching Methods, MOOCs

1 Introduction

Social entrepreneurship integrates entrepreneurial processes with the pursuit of a social mission. It encompasses enterprises whose primary objective is the sustainable reduction of exclusion, marginalization, or the suffering of specific social groups [1]. The European Union distinguishes between two types of social enterprises [2]: those that provide social services or goods to vulnerable populations, and those that foster social and professional integration through the employment of disadvantaged individuals. In essence, such enterprises incorporate into market activities individuals who might otherwise remain excluded from production and/or consumption due to reduced productivity, prejudice, low purchasing power, or related constraints [3].

In response to the growing prominence of social enterprises, universities have increasingly developed courses designed to cultivate the competencies necessary for launching and successfully managing social ventures. The subsequent chapter provides a concise literature review on social entrepreneurship education (SEE), with particular emphasis on teaching methods and the competencies such education seeks to foster. Drawing on this review, the authors formulated the research questions guiding the study. The third section outlines the methodological framework, followed by data

Research Paper
DOI: https://doi.org/10.46793/ eLearning2025.357K
Part of ISBN: 978-86-89755-37-4

analysis in the fourth section, while the fifth section presents the discussion and conclusions.

2 Theoretical background

The first formal course on social entrepreneurship was introduced by Gregory Dees at Harvard University in the mid-1990s. Soon thereafter, other leading U.S. universities incorporated social entrepreneurship into their curricula, further legitimizing the field within higher education [4], [5]. The concept quickly diffused to Europe, where universities in the United Kingdom, France, Belgium, and Italy were among the early adopters [4], [5]. Empirical research has shown that students exposed to social entrepreneurship education (SEE) perceive social entrepreneurship as both more desirable and more feasible than their counterparts who lack such exposure [6].

Perić and Delić [7] argued that the integration of social entrepreneurship education represents an integral dimension of universities' broader social responsibility. Croatian higher education institutions responded gradually by embedding the topic into existing entrepreneurship courses [8]. A notable milestone was the introduction of the elective course Social Entrepreneurship and Social Innovation at VERN Polytechnic in 2012 [9]. At present, only one mandatory SEE course exists in Croatia: it is offered within the undergraduate program Economics of Entrepreneurship at the Faculty of Organization and Informatics, University of Zagreb [10], [11]. Comparative research between Croatia and the United States indicates that Croatian SEE offerings remain less focused on innovation and scaling social enterprises, while also providing fewer opportunities for experiential projects with service-learning components [12]. Furthermore, significant scope for improvement persists in developing student competencies in sales and financial management.

Technological advancements have simultaneously reshaped the higher education sector by enabling the widespread delivery of courses online, independent of students' geographic location. Massive Open Online Courses (MOOCs) experienced particularly rapid expansion between 2017 and 2021, more than doubling in size during this period, although growth has since decelerated. The online learning platform market is projected to generate USD 60.25 billion in revenue by 2025, with an expected annual growth rate of 5.81% from 2025 to 2029 [13]. MOOCs platforms themselves may be conceptualized as social enterprises, as they create social value by mitigating inequalities in access to educational resources through free or low-cost online offerings available globally [14].

Paunescu and Vidović [15] examined the availability of SEE within MOOCs platforms. Coursera emerged as the most significant provider, with content developed by 11 universities. The University of Pennsylvania was the most active contributor, followed by the University of Illinois at Urbana-Champaign, ESSEC Business School, Copenhagen Business School, and the University of Colorado Boulder. Beyond supporting the creation of new social enterprises, MOOCs-based SEE contributes to the cultivation of sustainability-oriented and socially entrepreneurial mindsets among learners across the public, private, and civil society sectors [16].

The competencies required to successfully establish and manage a social venture, as articulated by practitioners, include problem-solving, effective team building, financial management, leadership and capacity development, stakeholder communication, interpersonal skills, marketing and sales, strategic management, outcome measurement, and the ability to foster collaborative relationships [17]. Brock and Steiner [4] identified recurrent thematic areas in SEE curricula, including social problems and needs, innovation, scaling social enterprises, resource mobilization, opportunity recognition, sustainable business outcomes, and outcome measurement. To investigate whether MOOCs-based SEE continues to emphasize these themes more than a decade later, the following research question is posed:

RQ1: What topics dominate social entrepreneurship education on MOOCs platforms?

Prior research has highlighted that students are more effective in developing social ventures when intrinsically motivated, particularly through emotional attachment to a social entrepreneurial idea [18], [19]. Consequently, SEE pedagogies should aim to strengthen compassionate love, considered the most distinctive trait of individuals motivated to engage in social entrepreneurship [20]. Brock and Steiner [4] also documented a wide range of teaching methods in SEE, including lectures, venture analysis, discussions, case studies, practical projects, service learning, consulting, business plan development, as well as less common approaches such as guest lectures, volunteering, internships, and direct involvement in social entrepreneurial ventures. To determine the extent to which these methods are employed in MOOCs-based SEE, the following research question is formulated:

RQ2: Which teaching methods are most commonly employed in social entrepreneurship education on MOOCs platforms?

3 Methodology

This study employs a desk research approach to collect qualitative data on the availability of social entrepreneurship education through MOOCs platforms. The initial step in identifying courses whose curricula predominantly address social entrepreneurship involved determining which MOOCs platforms to include in the analysis. While numerous online articles compile lists of available platforms, this study drew upon *The Massive List of MOOC Platforms Around the World in 2025* [21], as it represents the most recent and comprehensive source, categorizing platforms by global region. Subsequently, all North American MOOCs platforms listed in Table 1 were systematically searched using the keywords *"social entrepreneurship"*, *"social entrepreneur"*, and *"social enterprise"*. The resulting data were manually organized in an Excel spreadsheet to facilitate the identification of patterns in course content and teaching methods. The analysis, therefore, relies on the accuracy and completeness of the information publicly available on the MOOCs platforms.

Table 1. MOOCs platforms.

Region	MOOCs platform/Country
North America	Coursera/the United States
	edX/the United States
	Udacity/the United States
	Canvas Network/the United States
	Kadenze/the Unites States
	Complexity Explorer/the United States
	MéxicoX / Mexico - nula
Europe	FutureLearn/the United Kingdom
	France Université Numérique (FUN) / France
	EduOpen / Italy
	Federica Web Learning / Italy
	European Multiple MOOC Aggregator (EMMA) / Europe
	OpenHPI / Germany
	MOOC.fi / Finland
	Prometheus / Ukraine - 2
	Open Education (openedu.ru) / Russia
Asia (excluding China)	SWAYAM / India
	NPTEL / India
	JMOOC / Japan
	gacco / Japan
	OpenLearning / Japan
	K-MOOC / Korea
	ThaiMOOC / Thailand
	Edraak (Arabic) / Jordan
	Campus-Il / Israel
	UTas / Australia
China	XuetangX / China
	Chinese University MOOC / China
	Zhihuishu / China
	Xue Yin Online / China
	Open Education (openedu.tw) / Taiwan
	eWant — education you want / Taiwan
	A massive list of all Chinese language MOOC platforms

4 Results

Of the seven North American MOOCs platforms reviewed, only two—Coursera and edX—offer courses in social entrepreneurship. Each platform currently provides six SEE courses.

On Coursera, two of the identified courses are designed as stand-alone offerings. Three additional courses may also be taken individually, although they form part of a broader specialization entitled *Social Entrepreneurship*, which delivers a more comprehensive curriculum in the field. The final course, *Social Entrepreneurship*, is embedded within the wider specialization *Business Strategies for a Better World*, which extends beyond social entrepreneurship to include other business-related topics.

The six courses available on edX are all structured as stand-alone offerings. Notably, two of these are archived, meaning that while students can still access the teaching materials, they are unable to obtain certification. An overview of the identified courses is presented in Table 2.

Table 2. SEE courses on MOOCs platforms.

Platform	Course title	University
Coursera	Creating Change through Social Entrepreneurship	Yale University, USA
Coursera	Social Entrepreneurship	University of Pennsylvania, USA
Coursera	Identifying Social Entrepreneurship Opportunities	Copenhagen Business School, Denmark
Coursera	Social Business Model and Planning for Social Innovation	Copenhagen Business School, Denmark
Coursera	Unleashing the Impact of your Social Enterprise	Copenhagen Business School, Denmark
Coursera	Social Impact Strategy: Tools for Entrepreneurs and Innovators	University of Pennsylvania, USA
edX	Social Entrepreneurship and Systems Change	Harvard University, USA
edX	Disciplined Approach to Social Entrepreneurship	Curtin University, Australia
edX	Social Entrepreneurship – von der Idee zur Umsetzung	Social Entrepreneurship Akademie, Germany
edX	Enabling Entrepreneurs to Shape a Better World	Social Entrepreneurship Akademie, Germany
edX	Social Franchising	Católica Lisbon School of Business and Economics, Portugal
edX	Business and Impact Planning for Social Enterprises	Massachusetts Institute of Technology, USA

Figure 1 illustrates the distribution of SEE courses by university. Among the institutions represented on the analyzed MOOCs platforms, four are based in the United States, three in the European Union, and one in Australia.

With respect to language, all but one of the courses are delivered in English; the exception is a course offered in German. The extent of accessibility for non-English speakers varies considerably, ranging from a course with no alternative language options to three courses that provide translation into as many as 24 languages.

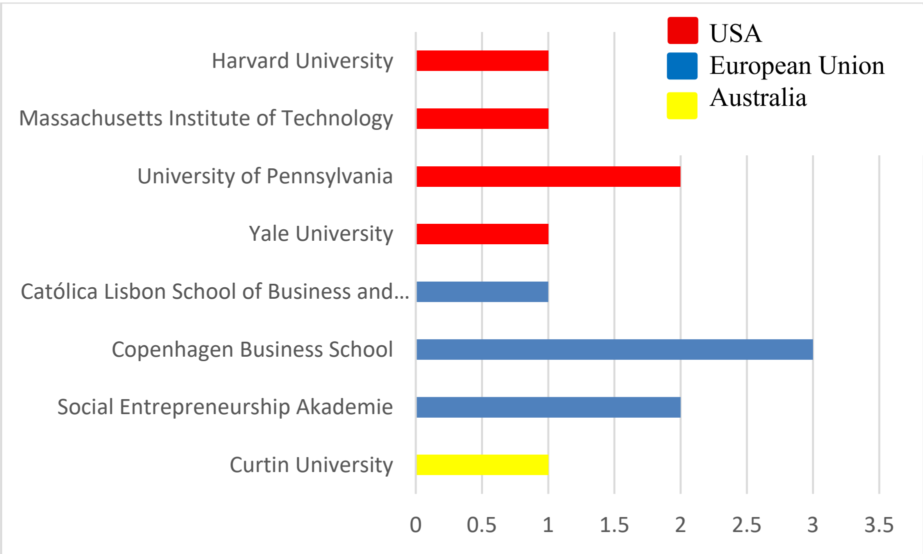


Fig. 1. Number of SEE courses on MOOCs platforms by university.

On the edX platform, enrollment data were available for only one course, which recorded 5,297 students. By contrast, the Coursera courses achieved substantially higher enrollment figures, ranging from 8,823 to 55,493 students. Enrollment numbers are, however, strongly influenced by the time of course creation and therefore cannot be regarded as a direct indicator of course quality.

Nevertheless, two courses stand out in terms of participation. The course *Social Impact Strategy: Tools for Entrepreneurs and Innovators*, offered by the University of Pennsylvania (USA), has enrolled 55,493 students to date, representing the highest figure observed. The second most popular course, *Identifying Social Entrepreneurship Opportunities*, developed by Copenhagen Business School (Denmark), has attracted 32,978 students to date.

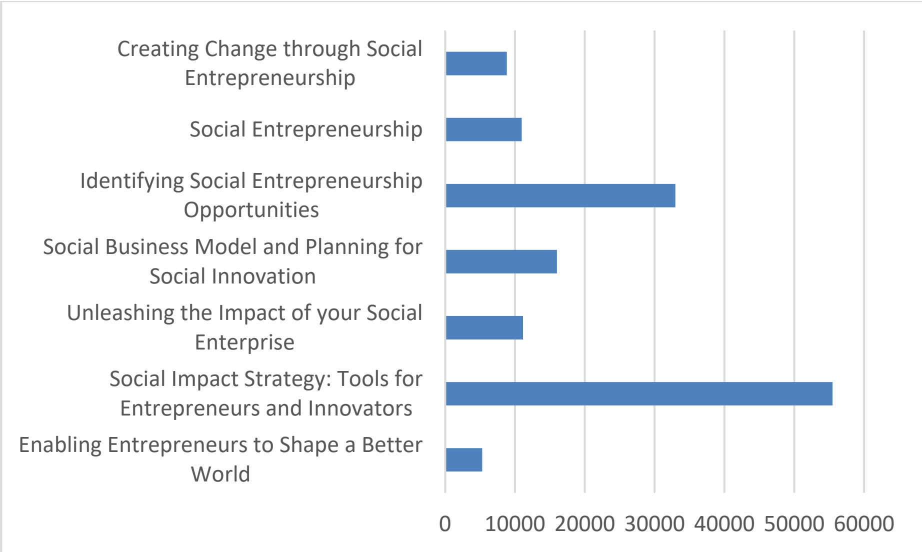


Fig. 2. Total number of enrolled students (data available for seven courses).

With the exception of one course that specifies prior experience as a prerequisite, all of the analyzed offerings are designed at the beginner level. In terms of workload, the courses vary considerably, ranging from three to forty hours of study time, thereby providing options suitable for learners with different preferences. All currently available certificate-granting courses operate on a flexible schedule. Only one course was instructor-paced, though it is now archived and no longer actively facilitated.

The dominant thematic areas within SEE courses on MOOCs platforms are presented in Figure 1. The most frequently addressed topics include outcome measurement, innovation processes, resource acquisition, scaling, and sustainability. A variety of approaches to outcome measurement are introduced, such as Social Return on Investment (SROI), balanced scorecards, use-case diagrams, theory of change, performance criteria frameworks, and the IOOI (input–output–outcome–impact) model. Innovation is most commonly addressed through user-driven design, emphasizing an understanding of beneficiaries. To cultivate empathy, several courses employ tools such as empathy maps, mind maps, and stakeholder maps, while others explore prototyping, problem trees, and pivoting strategies.

Resource acquisition is primarily examined in three dimensions: (1) funding opportunities for social enterprises, including impact investing and crowdfunding; (2) stakeholder collaboration; and (3) human resource strategies, particularly the recruitment of individuals aligned with the organization's mission and the formation of effective teams. One course, for example, draws upon Belbin's team role framework.

Scaling is conceptualized as both organizational growth and partnership development, while sustainability is framed through the construction of durable business models. Under this theme, courses introduce tools such as the business model spectrum, revenue and distribution models specific to social enterprises, competitive analysis, positioning assessments, the logic model, the value proposition canvas, and the business model canvas. One course further expands this toolkit with process mapping, stakeholder mapping, root cause analysis, and system boundary definition.

Approximately half of the courses also address opportunity evaluation, communication strategies, business plan development, and the legal structures of social enterprises. Opportunity evaluation typically involves data collection and problem analysis, incorporating demographic, regional, and infrastructural factors. Some courses distinguish between *impact opportunities* and *market opportunities*. Communication-related content includes pitching, storytelling, communication planning, and the formulation of impact opportunity statements.

A number of additional topics appear more sporadically across the curricula. For instance, several courses provide a broad overview of current social needs, highlight the role of failure in entrepreneurial learning, or address market analysis in greater depth. Others emphasize management techniques, student self-assessment of strengths and weaknesses, or knowledge of the broader social entrepreneurship ecosystem. One course specifically focuses on intellectual property.

Overall, the courses offered on Coursera and edX are largely aligned in their coverage of key topics. Minor differences emerge, however, with edX courses tending to place stronger emphasis on sustainability, while Coursera courses devote comparatively greater attention to the development of business plans.

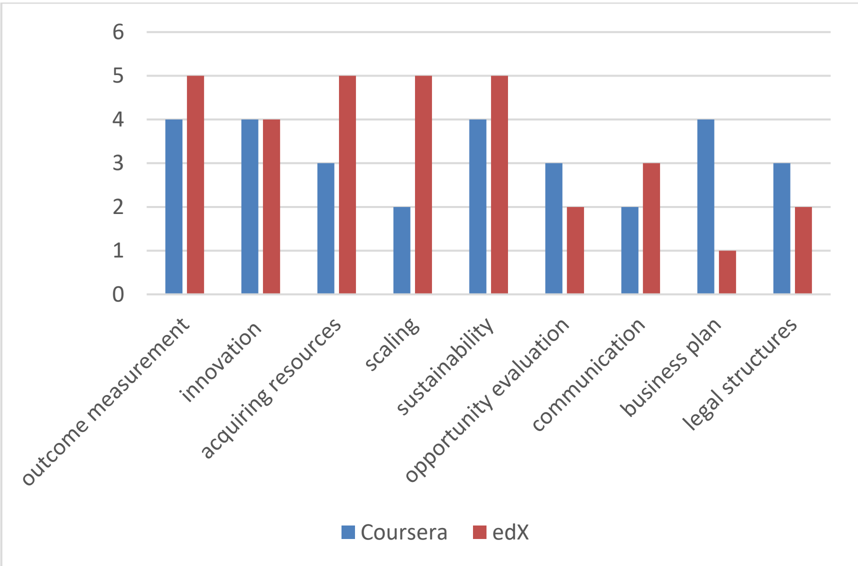


Fig. 1. Topics covered in SEE courses on MOOCs platforms.

Figure 2 presents a comparison of the topics addressed in SEE courses developed in the United States and the European Union. Courses originating in the U.S. place comparatively greater emphasis on outcome measurement and sustainability.

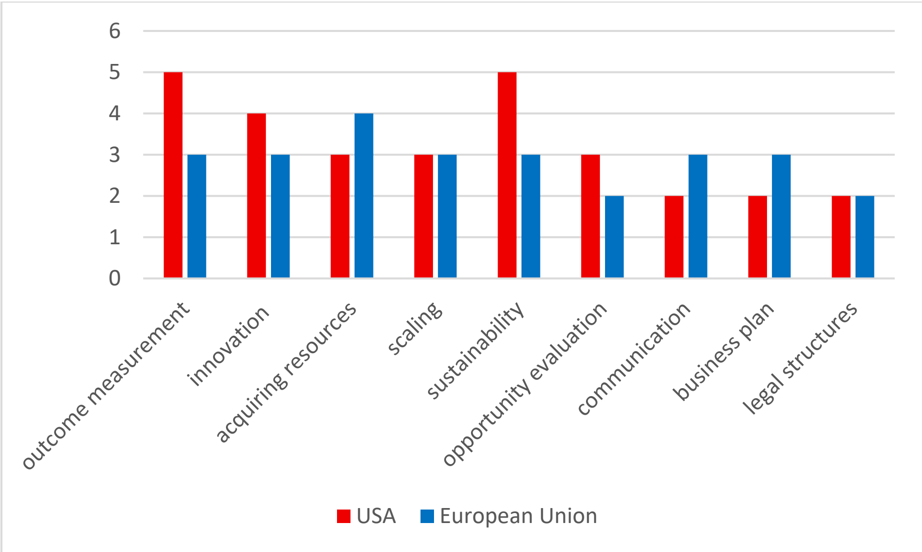


Fig. 2. Topics covered in SEE courses developed in the USA and the European Union (the course developed in Australia is not included here).

Figure 3 summarizes the teaching methods employed across the analyzed SEE courses. All courses incorporate video lectures and assigned readings, while all but one make extensive use of detailed case studies. Guest lectures represent the next most frequently adopted instructional method, followed by discussion prompts, which were identified in seven courses. However, the effectiveness of these prompts appears limited, as they typically elicit individual student statements without fostering substantive peer-to-peer interaction.

To assess students' knowledge and competencies, seven courses employ quizzes, and an equal number utilize practical assignments, three of which are evaluated through peer review. Overall, the pedagogical approaches do not differ substantially between the two platforms. A minor distinction can nevertheless be observed: edX courses make more frequent use of quizzes compared to those hosted on Coursera.

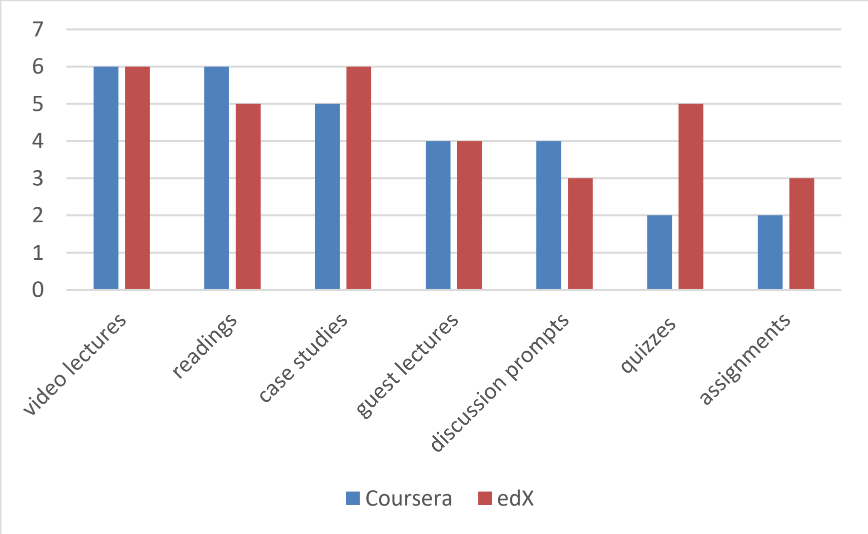


Fig. 3. Teaching methods used in SEE courses on MOOCs platforms.

Figure 4 presents a comparison of teaching methods employed in SEE courses developed in the United States and the European Union. Courses originating in the European Union make comparatively greater use of guest lectures than those developed in the United States.

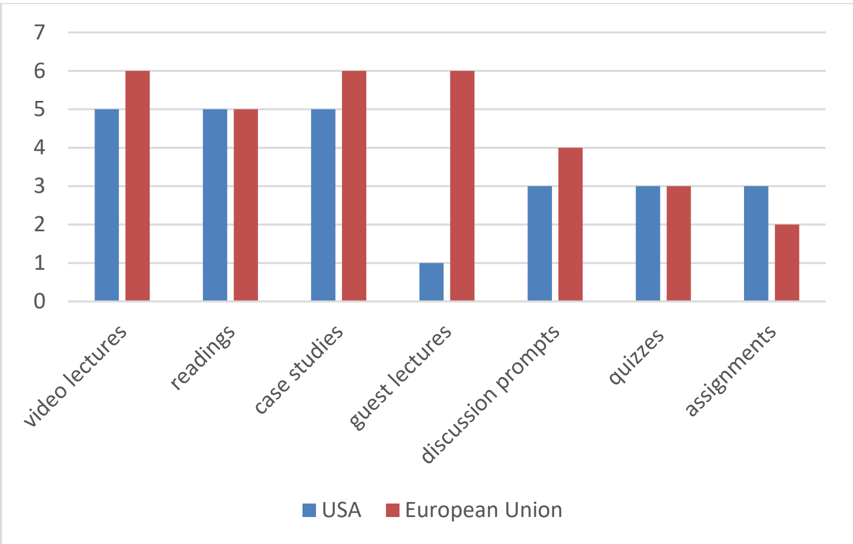


Fig. 4. Teaching methods used in SEE courses developed in the USA and the European Union (the course developed in Australia is not included here).

5 Conclusion

Individuals interested in social entrepreneurship may benefit considerably from introductory-level courses available on Coursera and edX. A primary limitation, however, concerns the accessibility of these courses in multiple languages, which varies substantially across providers. In terms of content, MOOCs-based SEE generally addresses the same thematic areas as traditional face-to-face university courses. Nevertheless, their shorter duration restricts opportunities for more in-depth engagement with contemporary social problems and needs.

Self-paced online education offers the advantage of continuous accessibility, yet it remains limited compared to instructor-led formats due to its reliance on less interactive teaching methods. In particular, the absence of human instructors precludes the use of pedagogical approaches such as service learning, volunteering, and internships within social enterprises. As a result, the analyzed MOOCs provide condensed, highly focused content delivered within a relatively short timeframe. While low cost and flexible scheduling represent clear advantages, interactive learning opportunities are underutilized because they require greater instructor involvement and consequently increase course expenses. Some courses attempt to address this gap through peer evaluation, though its effectiveness remains debatable.

As previously noted, discussion prompts rarely result in meaningful interaction, as students tend to share individual opinions without responding to their peers. Two potential improvements are proposed. First, online discussion forums could be organized by economic sectors and target groups, allowing students with similar interests—such as supporting homeless populations—to connect with peers who have completed the course, review their reflections, learn from them and potentially initiate collaborations. Second, courses could incorporate a service-learning component whereby students volunteer within their chosen sector of interest and subsequently share their reflections with peers. Such measures would not significantly increase instructor workload, while enhancing experiential and collaborative learning. By contrast, embedding group projects may be less feasible, given the self-paced nature of MOOCs and the heterogeneous levels of student motivation.

Although nearly all analyzed courses are positioned at the beginner level, there is a clear absence of instructor-led formats capable of providing tailored guidance to teams developing social ventures. While such courses would inevitably be more resource-intensive, they would better facilitate the acquisition and application of entrepreneurial skills rather than focusing solely on theoretical knowledge.

Among the less frequently addressed topics, the role of failure warrants particular attention. It is argued that every entrepreneurship curriculum should explicitly address entrepreneurial failure, as it constitutes an inherent and often recurring element of entrepreneurial practice.

Finally, the analyzed courses on North American MOOCs platforms employ different practical tools to address similar topics in social entrepreneurship. Future research should extend this analysis by comparing SEE on platforms belonging to diverse geographical regions to identify similarities, differences, and potential best practices.

References

1. Martin, R. L., Osberg, S.: Social entrepreneurship: the case for definition. Stanford Social Innovation Review **5**(2), 27–39 (2007)
2. European Commission: Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee and the Committee of the regions, Social Business Initiative, Creating a favourable climate for social enterprises, key stakeholders in the social economy and innovation. European Commission (2011)
3. Kedmenec, I.: The Impact of Individual, Social and Cultural Factors on the Social Entrepreneurial Intentions of Business Students, Ph.D. thesis, University of Maribor, (2015).
4. Brock, D., Steiner, S.: Social entrepreneurship education: Is it achieving the desired aims? SSRN Electronic Journal, (2009)
5. Vidović, D.: Socijalno poduzetništvo u Hrvatskoj, Ph.D. thesis, University of Zagreb, (2012).
6. Kedmenec, I., Rebernik, M., Tominc, P.: Social Entrepreneurship Education and Its Association with Perceived Desirability and Feasibility of Social Entrepreneurship among Business Students. Croatian Journal of Education: Hrvatski časopis za odgoj i obrazovanje **18**(4), 1035–1065 (2016)
7. Perić, J., Delić, A.: Developing social responsibility in Croatian Universities: A benchmarking approach and an overview of current situation. International Review on Public and Nonprofit Marketing **13**, 69–80 (2016)
8. Škrtić, M., Mikić, M.: Poduzetništvo [Entrepreneurship]. Sinergija nakladništvo, Zagreb (2011)
9. Vidović, D. (2019). Social enterprises and their ecosystems in Europe. Country report - Croatia. European Commission (2019)
10. Bilić, I., Lovrinčević, M.: Moving towards social entrepreneurship in south east Europe: Can we rate academic contribution? In: Soliman, K. S. (ed.) International Business Information Management Association, IBIMA, pp. 2236–2247. IBIMA, Madrid (2018)
11. Kedmenec, I.: Social entrepreneurial intention: Does the classroom matter? In: Skala, K. (ed.) 43rd international convention on information, communication and electronic technology, MIPRO, pp. 1435–1440. MIPRO, Rijeka (2020)
12. Konecki, I., Toplek, A., & Detelj, K.: Social entrepreneurship education at Croatian universities. Revija za socijalnu politiku, **30**(3), 253-268 (2023)
13. Statista, Online Learning Platforms – Worldwide, https://www.statista.com/outlook/emo/online-education/online-learning-platforms/worldwide, last accessed 2025/08/30
14. Arıker, Ç.: Massive open online course (MOOC) platforms as rising social entrepreneurs: Creating social value through reskilling and upskilling the unemployed for after COVID-19 conditions. In Research Anthology on Business Continuity and Navigating Times of Crisis. IGI Global Scientific Publishing, Khosrow-Pour, M. (ed.) pp. 607–629 (2022)
15. Paunescu, C., Vidović, D.: Social Entrepreneurship Education in Academia: A MOOC Scheme Analysis. In: Dima, A. M. (ed.) Proceedings of the international conference on economics and social sciences, pp. 491–501. Sciendo, Bucharest (2020)

16. Calvo, S., Lyon, F., Morales, A., Wade, J.: Educating at scale for sustainable development and social enterprise growth: The impact of online learning and a massive open online course (MOOC). Sustainability, **12**(8), 3247 (2020)
17. Miller, T., Wesley, C., Williams, D.: Educating the Minds of Caring Hearts: Comparing the Views of Practitioners and Educators on the Importance of Social Entrepreneurship Competencies. Academy of Management Learning & Education **11**(3), 349–370 (2012)
18. Amundam, D. N.: Enhancing potential social innovative thinking, responsible, social entrepreneurship education: a curriculum content and teaching method model. Journal of Entrepreneurship Education, **22**(5), 1–21 (2019)
19. García-González, A., Ramírez-Montoya, M. S.: Social Entrepreneurship Competency in Higher Education: An Analysis Using Mixed Methods. Journal of Social Entrepreneurship, published online, 91–109 (2020)
20. Kedmenec, I., Rebernik, M., Perić, J.: The impact of individual characteristics on intentions to pursue social entrepreneurship. Ekonomski pregled: mjesečnik Hrvatskog društva ekonomista Zagreb **66**(2), 119–137 (2015)
21. The Report, Massive List of MOOC Platforms Around the World in 2025, https://www.classcentral.com/report/mooc-platforms/?utm_source=chatgpt.com, last accessed 2025/08/30

INTERNATIONAL Conference on e-Learning (16 ; 2025 ; Beograd)
E-Learning 2025 : proceedings / The Sixteenth International Conference on e-Learning, Belgrade, 25-26 September 2025. ; editors Dragan Domazet, Nemanja Zdravković. - Belgrade : Metropolitan University, 2025 (Niš : Control P). - 369 str. : ilustr. ; 30 cm

Tiraž 20. - Bibliografija uz svaki rad.

ISBN 978-86-89755-37-4

1. Domazet, Dragan, 1947- [уредник] 2. Zdravković, Nemanja, 1988- [уредник]
а) Електронско учење -- Зборници б) Вештачка интелигенција -- Примена -- Зборници в) Образовна технологија -- Зборници г) Информациона технологија -- Образовање -- Зборници

COBISS.SR-ID 182987785