



BLOOMING AI

Developing an AI-powered Open-Source Learning Platform based on Bloom's taxonomy

Work Package 2 – Activity 2.1

* * *

MAPPING STATE-OF-THE-ART PRACTICES: REPORT

Alexandra Lazareva (Associate Professor, University of Agder) alexandra.lazareva@uia.no

Frédéric Clavert (Assistant Professor, University of Luxembourg) frederic.clavert@uni.lu

Marianthi Grizioti, (Assistant Professor, University of Athens) mgriziot@eds.uoa.gr

Ioana Duta-Visescu (Postdoctoral Researcher, University of Luxembourg) ioana.visescu@uni.lu

Vebjørn Sæterdal Landmark (Student Assistant, University of Agder) vebjorn.s.landmark@uia.no

Fridtjof Emil Reme (Student Assistant, University of Agder) fridtjof.e.reme@uia.no

OBJECTIVE

The aim of this report is to outline the results of state-of-the-art review targeting to identify best practices and methodologies for integrating GenAI with Bloom's taxonomy. The results of the report are based on a systematic review of academic literature. In the following, the review method is outlined, and then selected results are presented. It is important to emphasise that this report does not cover all the papers included in the review. This report draws on the most relevant methodologies and practices identified for the BloomingAI project.

A Zotero folder with all the identified literature, an [Excel sheet](#) outlining the tags for each item, and an [overall summary](#) of the results are available for a closer and more detailed inspection.

METHOD

The review was performed by three reviewers in the BloomingAI project team. Four databases were chosen: cairn.info (<https://shs.cairn.info>), JSTOR (<https://www.jstor.org>), ERIC (<https://eric.ed.gov>), IEEE Xplore (<https://ieeexplore.ieee.org/Xplore/home.jsp>). The search for peer-reviewed literature published since 2017 was performed using the keywords (“artificial intelligence” OR AI) AND “higher education” AND (Bloom OR “Bloom’s taxonomy”). After removing the duplicates, 138 papers were identified. This included journal papers, conference papers, and books. The team decided to include relevant literature regardless of the subject domain. Both empirical and theoretical papers were included. 10 items were excluded, resulting in the total of 128 items (see Figure 1.)

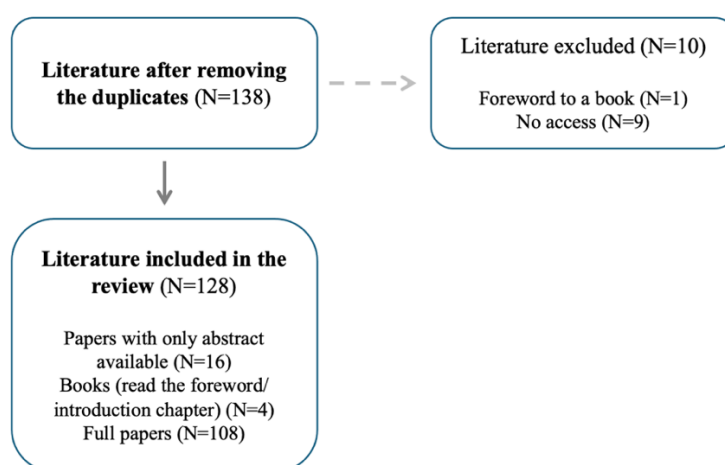


Figure 1. Literature review results

The team proceeded to tag the literature in terms of (1) general themes and key findings to identify best practices, as well as (2) research method and limitations to identify existing methodologies. At least two of the reviewers read each paper; in addition, two student assistants were trained and completed an additional round of tagging the total of 91 papers for increased inter-rater reliability.

This report highlights the most relevant findings and refers to specific examples from the literature. The variation in the conclusions was somewhat expected as the review included studies carried out in various subject domains.

FINDINGS

Findings Overview

Both empirical studies (N=68) and conceptual/theoretical/opinion pieces (N=60) were identified and included in the final review.

Some of the conceptual papers touched upon the issues of ethics; this aspect is left out from this report as it is addressed in the report on the ethical risks of AI in education (task 2.2).

Bloom's taxonomy was more present in some papers than in others; this report is focused on papers where Bloom's taxonomy had a significant role.

GenAI and Practical Implications for Assessment Pedagogy

A part of papers was dedicated to the issues of exam questions, both when it comes to exam question generation (e.g., Akdeniz et al., 2025; So et al., 2025) and automation of exam question classification. The study by Akdeniz et al. (2025) demonstrates that GenAI is effective in generating questions/tasks aiming at lower-order cognitive questions. The study by So et al. (2025), however, demonstrates GenAI's ability to generate multiple choice questions of different difficulty levels and question types aligning with all the levels of Bloom's taxonomy (So et al., 2025).

Several papers demonstrate that GenAI has a vast potential to effectively automate exam question classification (e.g., Banujan et al., 2023; Gani et al., 2023; Gavhane & Pagare, 2025). The key implication of these findings is that GenAI has a large potential in terms of reducing the workload for educators.

Turning to the identified conceptual/theoretical pieces, the issue of assessment is also widely addressed. Some papers raise the issue of and advocate for redefining assessment (e.g., Chen & Nunes, 2025; Lubbe et al., 2025) in the age of GenAI. Felipe et al. (2025) specifically discuss GenAI's implications for redesigning home exams (specifically tailored to levels 3-6 in Bloom's taxonomy). Nault et al. (2025) suggest a framework for assessment combining Bloom's and ICAP (interactive, constructive, active, and passive) model, this way linking two taxonomies where one is focused on the goals and the other one on the means of achieving those goals. These frameworks lack empirical validation.

The review paper by Lubbe et al. (2025) is focused specifically on assessment pedagogy, advocating for the creation of assessments on the higher-order thinking level. It is suggested to integrate "AI evaluation" as a separate skill. Lubbe et al. (2025) also claim that assessments should focus not only on the output but also on the process. The results of their review provide several practical recommendations; however, the limitation of this paper is that it is – same as the pieces mentioned above – purely theoretical and does not

build on empirical research. At the same time, empirical evidence can be found in the study by Toker & Akgun (2025). This study involved 123 student participants completing three tasks of increasing complexity, aligned with Bloom's taxonomy, across four groups: (1) control, (2) e-textbook, (3) Google, and (4) ChatGPT. The ChatGPT group initially demonstrated the highest AI plagiarism rates during the lower-order cognitive tasks, but their performance improved on higher-order cognitive tasks that required analysis, evaluation, and creation. The key takeaway from the study is that it is central to design assessments fostering higher-order thinking as a strategy for minimising AI plagiarism.

Methodologies for Content Creation with GenAI

A number of identified papers discuss how GenAI can assist educators in the overall learning content creation. For example, the case study by Robinson (2024) emphasises the importance of an objective-driven approach when designing content, to avoid fragmented AI-generated content and move towards coherent learning journeys. To achieve that, Robinson (2024) combines backward design, Bloom's taxonomy and constructivist theory, and emphasise the importance of ongoing validation and human oversight when designing learning content with the help of GenAI. Similarly, Faraon et al. (2023) suggest how learning activities can be tailored to Bloom's taxonomy.

While the papers mentioned above are conceptual and lack empirical validation, there are also empirical studies on the issue. The study by Alidadi et al. (2024) which is focused on the domain of programming, presents promising evidence that GenAI can generate tasks at the “apply” and “analyse” levels of Bloom's taxonomy. At the same time, as the authors themselves note, this can be related to the nature of the programming subject itself as it requires problem-solving skills (Alidadi et al., 2024). Another study by Idris (2025) develops and investigates algorithms to transform traditional educational content into enhanced video lectures with the help of Bloom's taxonomy. Bloom's taxonomy is involved here specifically to ensure the generation of higher-order thinking content (Idris, 2025).

Best Practices for GenAI-Assisted Teaching and Learning

Most part of this report is dedicated to outlining some of the identified best practices and methodologies for integrating GenAI with Bloom's taxonomy in the context of teaching and learning in higher education. Theoretical pieces emphasise the importance of the AI agent being an *assistant* (e.g., Chan & Colloton, 2024), repeatedly pointing out the central role of the educators and students as active participants of the process. Some works suggest the need for developing such skills in students as “higher-order prompting”, and even revising the Bloom's taxonomy, adding “discovery” at the base of the pyramid (Jackson, 2025).

Some empirical pieces also focus on the issue of prompting. De los Ángeles Domínguez-González et al. (2023) analysed pre-service teachers' abilities to write prompts, demonstrating that the largest number of generated prompts corresponded to the highest levels of Bloom's taxonomy ("apply" and "create"), which highlights the potential of GenAI for the development of pedagogical skills. At the same time, in their empirical piece, Gereade (2024) argues that students are more engaged in fundamental understanding at the initial stages of learning, which makes them more comfortable asking lower-order thinking questions before naturally moving on to formulating higher-order queries. In their study, they conclude that the lower-order queries are still dominating in number, which suggests that students generally would benefit from training on formulating various types of higher-order thinking questions (Gereade, 2024).

Another theoretical piece, positioned in the domain of programming education, discusses the possibility of a "reverse" Bloom taxonomy – suggesting that since GenAI can handle the "creating" task effectively, more emphasis should be put on the students' ability to *understand* and critically *analyse* solutions proposed by GenAI (Pesovski et al., 2024).

Some papers focus on testing and comparing different GenAI tools' performance on the different levels of Bloom's taxonomy. For example, the study by Cheong (2025) demonstrates that ChatGPT performs sufficiently up to level 3 in Bloom's taxonomy, and the accuracy goes down as the cognitive level goes up. Murali et al. (2024) tested several chatbots in three subjects (introduction to programming through C, fundamentals of data structures, and computer security). Both students and all chatbots were given questions aligned with the Blooms taxonomy levels 1 to 4 in the three different subjects. The results of the study show that the chatbot responses were sufficient for lower-order thinking level questions but could not answer convincingly when asked for an algorithm. Thus, Murali et al. (2024) conclude that chatbots cannot be "depended on" for higher-order learning but can be used to assist students in passing basic courses.

The review process identified multiple empirical studies in various subject domains investigating GenAI's impact on various aspects of student learning. It is important to point out that most of these studies are based on self-reported data (e.g., student surveys), and only a few studies also involved learning performance assessment as a measure.

Essien et al. (2024) investigate GenAI's effects on the development of critical thinking skills in students. Their study, based on students self-reported data, demonstrates positive effects on the lower-level skills such as understanding and application, but no statistically significant impact on the highest levels, e.g., creating (Essien et al., 2024). At the same time, the study by Sanchez-Rodriguez et al. (2025) demonstrates that ChatGPT can effectively supports students' understanding at both lower and higher levels of Bloom's taxonomy, including skills such as analyzing, creating, and optimizing. However, the study

was carried out as a single-group pretest-posttest design with a relatively small sample (N=39) and also based on self-reported data (surveys).

The experimental study by Chang et al. (2022), positioned in nursing education, demonstrates that incorporating knowledge-based chatbot systems can have a positive impact on student learning, both in terms of their learning performance, and critical thinking and learning satisfaction (the latter two based on self-reported data).

Other works focused on integrating large language models with Bloom's taxonomy to target personalised learning, for example by means of a personal AI tutor. For example, in their study, Madhubhashana et al. (2024) classified learners in three groups – beginner, intermediate, and advanced. Through a simulation, this way of assisting learners demonstrated to improve student engagement in online large group learning settings. Such findings have important implications, especially when combined and considered together with research applying GenAI for predicting learner performance (e.g., Ane & Nepa, 2024).

Keskikallio et al. (2024) introduce a theoretically and empirically validated guided process (IntelliBloom) whose aim is to assist teachers in integrating educational AI services into their pedagogical practice. This process is grounded in Bloom's digital taxonomy and consists of five stages, aiming at increasing teachers' pedagogical AI literacy. The process was tested with 15 pre-service teachers, and data were collected through a survey and interviews (Keskikallio et al., 2024). The results demonstrate how IntelliBloom can guide teachers through the process of aligning AI tool with learning objectives using Bloom's digital taxonomy. At the same time, the results reveal some of the possible challenges. For example, some informants discuss that their work is usually "more chaotic and less structured". This is an important implication demonstrating that a guided process may not fully capture the nuances of unique professional practices and workflows.

Kinto-Ramirez et al. (2025) develop a methodology combining narrative-based instruction, cognitive scaffolding and AI. The methodology is grounded in problem-based learning, Bloom's taxonomy, and self-regulated learning. ANECAI (Narrative Learning and Collaborative AI-Assisted Assessment) methodology follows five phases: real-life contextualization, theoretical development through storytelling, scaffolded questioning, an AI-assisted synthesis phase, and a collaborative peer assessment. The study demonstrated that students exposed to ANECAI achieved a 15% improvement in conceptual retention and an 88.2% increase in self-regulated learning strategies. These findings emphasise the effectiveness of this methodology in fostering student autonomy and critical thinking. It is important to note that in this experimental study, self-reported data were combined with pre- and post-intervention knowledge assessments. However,

the amount of students was rather small (control and experimental groups of N=12 and N=17).

IDENTIFIED LIMITATIONS

The review identified several cross-cutting limitations. Firstly, the identified relevant theoretical pieces and frameworks are often lacking empirical validation. Secondly, many of the empirical papers are single-course and often based on a limited amount of participants and self-reported data. Finally, there is a lack of longitudinal studies in the area.

CITED WORKS

Akdeniz, H., Clark, T., & Roberts, J. L. (2025). Can AI generate questions aligned with Bloom's Taxonomy? A framework for gifted education to support teachers. *Journal of Advanced Academics*, 36(4), 671-694.

Alidadi, M., Taghiyareh, F., & Shahhoseini, N. (2024, October). Evaluating LLM-Generated Persian Questions for Teaching Conditional Programming Using Bloom's Taxonomy. In *2024 11th International Symposium on Telecommunications (IST)* (pp. 726-730). IEEE.

de los Ángeles Domínguez-González, M., Hervás-Gómez, C., Díaz-Noguera, M. D., & Reina-Parrado, M. (2023). Attention to diversity from artificial intelligence. *The European Educational Researcher*, 6(3), 101-115.

Ane, T., & Nepa, T. (2024). Artificial intelligence prediction model for educational knowledge representation through learning performance. *Research on Education and Media*, 16(2).

Banujan, K., Kumara, S., Prasanth, S., & Ravikumar, N. (2023). Revolutionising Educational Assessment: Automated Question Classification Using Bloom's Taxonomy and Deep Learning Techniques--A Case Study on Undergraduate Examination Questions. *International Journal of Education and Development using Information and Communication Technology*, 19(3), 259-278.

Chan, C. K. Y., & Colloton, T. (2024). *Generative AI in higher education: The ChatGPT effect*. Routledge.

Chang, C. Y., Kuo, S. Y., & Hwang, G. H. (2022). Chatbot-facilitated nursing education. *Educational Technology & Society*, 25(1), 15-27.

Chen, Y., & Nunes, B. P. (2025, July). Quantifying Assessment Difficulty in GenAI-Enhanced Education: Integrating Reinterpreted RBT and SOLO Taxonomy. In *2025 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 382-386). IEEE.

Cheong, M. (2025). ChatGPT's Performance Evaluation in Spreadsheets Modelling to Inform Assessments Redesign. *Journal of Computer Assisted Learning*, 41(3), e70035.

Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, 49(5), 865-882.

Faraon, M., Granlund, V., & Rönkkö, K. (2023, September). Artificial intelligence practices in higher education using Bloom's digital taxonomy. In *2023 5th International Workshop on Artificial Intelligence and Education (WAIE)* (pp. 53-59). IEEE.

Felipe, A. L., Khwakhali, U. S., & Nguyen, T. N. (2025, July). A Framework for Assessment Design in the Era of Generative AI: Case Study of Take-Home Assignment in Software-related Courses. In *2025 10th International STEM Education Conference (iSTEM-Ed)* (pp. 1-6). IEEE.

Gani, M. O., Ayyasamy, R. K., Sangodiah, A., & Fui, Y. T. (2023). Bloom's Taxonomy-based exam question classification: The outcome of CNN and optimal pre-trained word embedding technique. *Education and Information Technologies*, 28(12), 15893-15914.

Gavhane, J. M., & Pagare, R. (2025, April). Revolutionizing Academic Evaluation: Bloom's Taxonomy Meets Deep Learning and NLP. In *2025 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-10). IEEE.

Gerede, Ç. E. (2024, October). Are we asking the right questions to ChatGPT for learning software design patterns?. In *2024 9th International Conference on Computer Science and Engineering (UBMK)* (pp. 1092-1097). IEEE.

Idris, M. (2025, August). Educational Content Enhancement Algorithm: Integrating Cognitive Load Theory and Bloom's Taxonomy for AI-Driven Educational Video Generation. In *2025 IEEE International Conference on High Performance Computing and Communications (HPCC)* (pp. 1631-1638). IEEE.

Jackson, J. (2025, April). "Who's Doing the Thinking Here?": A Pedagogy-First Approach to Integrating Large Language Models in Higher Education. In *2025 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-3). IEEE.

Keskikallio, J., Ulvsgård, A., Villavicencio, V., Faraon, M., & Rönkkö, K. (2024, September). Guided by IntelliBloom: Promoting Artificial Intelligence Literacy Among Teachers in Higher Education. In *2024 6th International Workshop on Artificial Intelligence and Education (WAIE)* (pp. 318-330). IEEE.

Kinto-Ramírez, H., Yunes-Rojas, J. A., Nieto-Jalil, J. M., Morales-Nava, R., Sayeg-Sánchez, G., & Lozano-Aponte, J. (2025, July). Integrating Narrative Learning and AI for Collaborative Problem Solving in Engineering Education. In *2025 6th International Conference of the Portuguese Society for Engineering Education (CISPEE)* (pp. 1-7). IEEE.

Lubbe, A., Marais, E., & Kruger, D. (2025). Cultivating independent thinkers: The triad of artificial intelligence, Bloom's taxonomy and critical thinking in assessment pedagogy. *Education and Information Technologies*, 30(12), 17589-17622.

Madhubhashana, S., Ranhinda, V., Jayakody, D., Shavinda, L., Nawarathne, M., & Haddela, P. (2024, December). Adapting pedagogical frameworks to online learning: Simulating the tutor role with LLMs and the QUILT model. In *2024 6th International Conference on Advancements in Computing (ICAC)* (pp. 581-586). IEEE.

Murali, R., Dhanalakshmy, D. M., Avudaiappan, V., & Sivakumar, G. (2024, May). Towards assessing the credibility of chatbot responses for technical assessments in higher education. In *2024 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE.

Nault, K., Nieto-Taborda, M., & Ruhi, U. (2025, May). Linking ICAP and Bloom's Taxonomy: A Comprehensive Framework to Assess Students' Cognitive Engagement in the Context of GenAI. In *2025 5th International Conference on Artificial Intelligence and Education (ICAIE)* (pp. 460-464). IEEE.

Pesovski, I., Vorkel, D., & Trajkovic, V. (2024, November). AI-Powered Education: Rethinking the Way Programming is Taught Using AI Tools and Reversed Bloom's Taxonomy. In *2024 21st International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1-6). IEEE.

Robinson, B. (2024). Harnessing AI for structured learning: the case for objective-driven design in e-learning. In *International Journal on E-Learning* (pp. 457-470). Association for the Advancement of Computing in Education (AACE).

Sánchez-Rodríguez, E., Merino-Soto, C., Chen, M. H., Zavala, G., & Chans, G. M. (2025, April). AI-supported learning: Integrating ChatGPT to enhance cognitive skills in STEM education. In *2025 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-5). IEEE.

So, C. C., Lee, C. H., Yang, D., Loh, A. W. K., & Sze, C. K. (2025). Automatic Generation of High-Quality MCQs With LLMs for Artificial Intelligence Education. *IEEE Access*.

Toker, S., & Akgun, M. (2025). Reducing AI plagiarism through assessment of higher-order cognitive skills. *Innovations in Education and Teaching International*, 62(5), 1665-1681.