Proceedings of the 20th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 675–680 DOI: 10.15439/2025F8073 ISSN 2300-5963 ACSIS, Vol. 43

Applications and Challenges of Artificial Intelligence in Educational Course Design and Delivery

Daniel Dan
School of Applied Data Science, Modul University Vienna, Vienna, Austria
ORCID: 0000-0002-7251-7899
Email: daniel.dan@modul.ac.at

Anna Wróblewska, Bartosz Grabek, Michał Taczała
Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland
ORCIDs: A: 0000-0002-3407-7570, B: 0009-0008-4917-8017, M: 0009-0004-4407-2361
Emails: anna.wroblewsk1, bartosz.grabek.stud, michal.taczala.stud@pw.edu.pl

Minoru Nakayama
School of Interdisciplinary Mathematical Sciences, Meiji University, Tokyo, Japan
ORCID: 0000-0001-5563-6901
Email: nakayama@meiji.ac.jp

***Abstract*—This survey reviews AI goals and tools for: (1) preparing educational materials, (2) interacting with teachers and students, and (3) assessing the results and providing feedback with (semi-)automatic methods. As a summary, we provide the crucial challenges to be tackled and discuss the associated ethical concerns.**

***Index Terms*—Artificial Intelligence, Automated Assessment, Higher Education, Large Language Models, Learning Analytics**

I. Introduction

RECENT advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP), particularly the rise of Large Language Models (LLMs), have transformed human-computer interaction across domains. Significant changes are evident in education, where artificial intelligence supports students in their everyday studies and helps teachers deliver material more engagingly and effectively [1].

A recent survey indicates that teachers now use AI more often than students [2]. Our complementary survey reviews NLP methods for course design, delivery, and personalised assessment, including technical details and their implementation. We discuss how these solutions are designed, integrated into higher-education workflows, and evaluate their effectiveness for both students and teachers. The listed solutions form a strong basis for fully integrated, technology-aided education, and we present an example AI-driven engineering course.

Our research questions are:

- What current approaches apply NLP to course design and delivery?
- How do these methods fit educational settings?
- How can they be integrated into different course aspects?

The survey follows the class flow. Section II reviews NLP for course-content creation; Section III examines technology that enhances live teaching. Section V covers automated assessment, proctoring, and anti-cheat tools, and Section VI, explores personalised learning systems. Finally, we outline implementation challenges and ethics.

II. Course Content Creation & Modification

Designing a university course involves transforming ideas into course materials that are engaging, academically sound, and easy-to-follow [3], [4]. AI now supports instructors in this process. Designing teaching material must consider principles that support learning [5], [6], and how students learn and are assessed [7]. In practice, a course is split into several lectures – seminar, frontal, or interactive. While face-to-face teaching remains dominant, innovative educational technologies reduce student passivity, enabling them to interact and actively learn the class material [8]. Student assessment remains delicate and error-prone. Modern methods emphasize continuous homework and assignments, a mid-term checkpoint, and a final written, oral, or project-based exam [9]. At course launch, instructors draft a syllabus, a "contract" that defines objectives, lecture sequence, assessment, and permitted tools [10]. Some universities now ban student use of generative AI for homework [11]. Rather than prohibiting evolving tools, we advocate channeling them, e.g., letting chatbots handle low-stakes assessment to save instructor time. AI already cuts preparation time: a 2024 survey of 144 instructional designers found that 67% reported moderate-to-significant efficiency gains, and 64% let ChatGPT draft objectives or quizzes first [12].

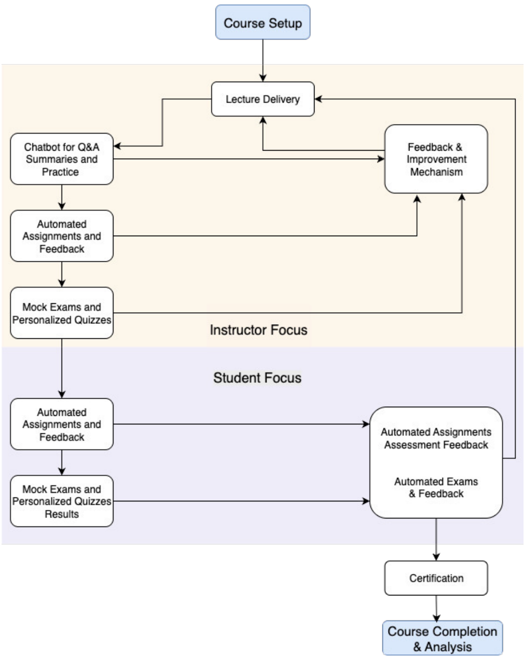

Fig. 1. Flowchart of subsequent steps in educational academic course delivery and students' assessment, up to the final certification. Note: light yellow are boxes more in the focus of the instructor, and grey ones are more for students

Figure 1 summarises the course flow. In an AI-enhanced version, instructors become supervisors and facilitators, while AI acts as a coach, evaluator, and feedback provider [13]. We will touch especially on the important ethical aspects and challenges while working on the creation of educational tools. The crucial questions here are:

1) How can AI/LLMs generate lecture materials and exercises?
2) How can technology structure the course and each session?
3) How can automated tools improve grading and feedback?
4) How do we evaluate the quality of AI-prepared material?
5) What are the pros and cons of teacher–AI collaboration?

A. How can AI/LLMs generate lecture materials and exercises?

AI tools can now draft lesson plans aligned with standards [14], [15], [16]. By specifying duration, lesson count, learner profile, and desired outcomes, instructors obtain varied analogies, examples, and explanations. Prompt engineering is continually refined [17], and multiple versions can be generated to suit individual students or fresh perspectives [18].

B. How can we use AI for detailed lesson planning and time allocation?

ChatGPT-style assistants quickly produce weekly and daily plans – topics, activities, assignments, assessments [19]. They can also recommend external videos and readings and link them to objectives, ensuring self-directed study complements the course [20]. Beyond text, AI can seed graphic-novel tasks, game-based projects, and debate topics to deepen engagement [21], [22], [23].

C. Grading and feedback generation

Timely, actionable feedback is hard at scale [24]. AI can create individualized quizzes and instant comments, drawing on prior performance, while human oversight preserves fairness and limits bias [25], [26].

D. Materials quality evaluation

AI-generated content still requires a documented two-pass human review: a subject-matter expert checks accuracy, then an instructional designer ensures pedagogy and accessibility [27], [28], [29]. Automated fact-checkers, readability analyzers, and rubric-based scores (e.g., AIGDER $\geq$ 85/100) flag issues for revision [30], [31], [32], [33]. Student "muddiest-point" feedback pinpoints residual confusion for rapid remediation [34], [35], [36], [37], [38].

E. Interaction between instructors and AI

Instructors work in cycles—prompt, review, refine, and embed pedagogy—when collaborating with AI. Attitudes range from scepticism about bias and deskilling [39] to enthusiasm for efficiency and creativity [40]. Empirical studies show that training and guided use shift beliefs and improve outcomes [41], [42].

III. TEACHING

Generative AI tools can create diversified and personalised teaching materials and even automate routine "paper work" such as emails, deadline checks and consistency control [43]. Yet quality is essential; recent studies stress rigorous evaluation, from re-using credible sources [46] to iterative assessment protocols [44], [45].

Beyond content creation, AI supports real-time observation of teaching and learning. A design study [47] foregrounds learning-analytics dashboards -single screens aggregating key indicators- for ongoing improvement. Surveys classify multiple dashboard types and confirm they mainly serve teacher supervision [48]

A. Interactive content generation

Automatic content or question generation hinges on carefully crafted LLM prompts that spell out context, audience, and difficulty. Retrieval-Augmented Generation can instead feed the model verified knowledge before output [38].

B. Engaging lecture delivery

During lectures, AI can act as a creative co-teacher—suggesting analogies or examples [49], powering avatar instructors that field questions [50], and turning text into images, videos or podcast teasers for multimedia slides [51], [52]. Students can critique AI answers to hone critical thinking and visualise experiments via dynamic images.

C. Real-time learning process supervision

Real-time supervision captures how, not just what, students learn. AI dashboards visualise engagement and performance from digital environments [53]; open-learner models update each learner's knowledge profile [54]. Dashboards must stay interpretable to be actionable [55], [56], [57]. Predictive analytics within these dashboards can flag at-risk or low-engagement students for prompt support [58], [59], [60], [61].

D. AI chatbot Q&A summaries and practice

LLM-powered chatbots have spread rapidly across higher education. Surveys find enthusiasm tempered by doubts: ease of use is clear, but cognitive gains are still debated [62], [63]. When offered as optional study partners, chatbots can scaffold distance or dyslexic learners through personalised encouragement, summaries and practice prompts [64], [65].

IV. Aid in Academic Activities

A. Coding assistance

One of the core competencies acquired by engineering and computer-science students is coding. Programming is indispensable for many assessed projects. Requirements vary—Mathematical Computation students code in R, Computer-Science majors in Python or Java, Robotics students in low-level hardware languages—but all must solve complex problems. AI-powered coding assistants (ChatGPT, Gemini, Claude, Grok) now provide instant code generation, refactoring, explanations and debugging via LLMs plus web search.

Recent work compared ChatGPT-assisted learners with self-directed peers: final scores were similar, but ChatGPT users debugged more, refined AI code and, citing ease-of-use, planned to keep using the tool [66].

These findings show how AI tools scaffold programming at every level, offering step-by-step guidance and fast debugging. Effectiveness, however, hinges on good prompts and critical evaluation; instructor-led prompt training markedly boosts outcomes [67]. LLMs still struggle with large code-bases, may hallucinate and often default to generic patterns. Over-reliance can erode academic integrity, so institutions need guardrails that balance tool use with independent problem-solving.

B. Academic writing aids

A review [68] groups modern AI writing tools into literature management, drafting, plagiarism checks, data analysis and other features. They accelerate brainstorming by mining vast corpora and adding predictive analytics throughout the research workflow.

V. Assessment

A. Assessment generation

Automated assessment is now a popular NLP topic because it saves teachers time and money. Its goal is to achieve scores similar to human graders with far less effort and cost. Many methods exist (e.g., semantic similarity), but large language models (LLMs) currently dominate. Their chief flaw is hallucination—inventing incorrect information, which undermines reliability. One mitigation is retrieval-augmented generation (RAG): a standard LLM linked to a vetted knowledge base. Retrieving evidence relevant to each question reduces hallucination and equalises scores [69], though irrelevant hits can still occur. LLMs further support AI-assisted grading [70]; yet their human-like prose raises doubts about assessment validity. Human oversight, therefore, remains essential [71].

LLM-generated assessments can still be inconsistent. The "LLM-as-a-judge" pattern—a second model auditing the first—adds a control layer that flags misgraded answers and outliers with high accuracy [72].

B. Automated final exam / exam proctoring

AI-driven NLP already handles full exams—multiple-choice and essay tasks. Studies confirm feasibility [73] but show score distributions can narrow, and domain-knowledge gaps (e.g., biomedical items) cut accuracy [74]. Essay scoring benefits from LLM text processing [75], yet quality hinges on prompt tuning: focusing on different rubric points shifts grades [76]. Performance also depends on language coverage [77], and research now extends LLM scoring to oral exams [78]. Item-response theory (IRT) frameworks could anchor AI scoring, replacing today's opaque heuristics [**?**].

C. Anti-cheat solutions

Anti-cheat tools remain vital [79]. Traditional detectors such as Turnitin catch plagiarism [80], while stylometric authorship checks compare writing style [81]. These methods falter on short answers. GPT-based classifiers now spot AI-generated text but still produce false positives. The most reliable—though intrusive—approaches monitor cursor or keystrokes during exams [82], [83].

VI. Personalized Learning

Personalized learning represents a fundamental shift in the educational process, aiming to replace the singular, standardized educational experience with a customised, learner-centred one delivered via intelligent tutoring systems, content recommendations or adaptive assessments.

A. Adaptive assessments

This protocol first creates a large RAG-based question bank, each item labelled for difficulty and topic. The instructor sets the time window, duration, item count and proctoring; the system then randomises—or adaptively selects—questions for every student. Learners log in while camera, safe-browser and image-recognition tools safeguard integrity. Built-in timers show time left and auto-submit; real-time alerts flag suspicious activity. After all submissions, RAG compares answers with a rubric, awards partial credit and exports grades to the LMS. Plagiarism scans and on-the-fly follow-up questions protect originality, while analytics refine subsequent tests.

Considerably less research views these tests from the learner's side. Co-designing AI-based assessments with students can boost critical thinking and creativity [**?**], and meta-awareness that items are machine-generated maps to Bloom's highest levels—analysis, synthesis and evaluation [85].

B. Personalized feedback

Providing reasonable feedback is a crucial part of the whole AI-supported exam-conducting ecosystem. It helps students fill gaps in knowledge, focus on the right areas, and understand the material faster. Research shows that the best feedback is timely, specific, and actionable [86]. Equally important is wording it pedagogically so it truly motivates learning. The most common solution today is a data-driven approach: transformer models (e.g., BERT) can evaluate many aspects of an answer through their understanding of context [87]. Yet merging expert knowledge with AI feedback is challenging; LLM comments are often generic and miss key details [86]. LLMs may also hallucinate—studies report that about 15% of feedback can be totally incorrect [86]. Prompt design, therefore, matters: telling a model to "behave as a teacher," giving step-by-step instructions, and asking it to analyse first and comment second all raise quality.

Despite these gains, the model is still far from perfect, so retrieval-augmented generation (RAG) can further improve performance. RAG combines an LLM with local data sources such as presentations or notes, ensuring feedback stays strictly aligned with course material [88]. A complementary method splits the task among several specialised agents, each handling a smaller part; this helps catch mistakes and adjust tone [88].

Feedback during learning is tentative diagnostic information, so a relationship of mutual trust between learner and instructor is essential. When automation intervenes, its effect on that relationship must be monitored; simple chatbots can bore users and discourage sustained dialogue.

LLM-based feedback systems have now been piloted in real courses, and learner perceptions are being evaluated [89]. Careful messaging can promote reflection on one's learning situation [90], and design choices—topic coverage, timing, tone—must be optimised for each activity [91].

VII. Challenges & Future Directions

In recent years, many studies have explored the possibilities, challenges and threats of AI tools in education. Yet most merely list what AI can do; few run carefully designed experiments to test usability and reliability. The key task now is to conduct comparative studies—contrasting AI-based recommendations, testing, grading and feedback with long-standing practices. For instance, we should compare LLM resources with earlier crowd-sourcing platforms used by students before chatbots.

We must also continuously measure how well up-to-date, fine-tuned and well-prompted LLMs acquire and refresh knowledge.

Moreover, ethical challenges loom large; automated tools must remain especially sensitive to educational contexts. As AI is introduced, we must be aware of data and algorithmic biases. Models trained on large corpora can mirror societal, cultural or institutional prejudices, prioritising students by past performance or learning style and producing unfair outcomes tied to race, language, socioeconomic status or disability. Thus transparency and explainability are needed for every recommendation or assessment, yet remain elusive because LLMs are inherently opaque.

Student privacy is also at stake, particularly under learning surveillance. Over-reliance on automation, coupled with inadequate informed consent, can erode professional judgement on both sides of the classroom. Unequal digital access—whether through language or technology—may widen existing educational gaps.

Mitigation demands careful design: human-in-the-loop oversight, informed student choice, and channels for learner voice. Done well, AI frees teachers to spend more quality, personalised time with each student.

Acknowledgments. This work was supported by the EU Horizon Europe OMINO (Overcoming Multilevel Information Overload) grant No. 101086321 and by the Polish Ministry of Education and Science (International Projects Co-Financed program). The views expressed are solely those of the authors and do not necessarily reflect those of the European Union or the European Research Executive Agency, which are not responsible for them.

References

[1] Y. Walter, "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., vol. 21, no. 1, p. 15, Feb. 2024, doi:10.1186/s41239-024-00448-3. https://doi.org/10.1186/s41239-024-00448-3

[2] Quizlet, Inc., "State of AI in Education," 2023. [Online]. Available: https://www.prnewswire.com/news-releases/quizlets-state-of-ai-in-education-survey-reveals-teachers-are-surprise-ai-champions-301884427.html

[3] J. Mason and S. Goodenough, "Course creation," in *Distance Teaching for Higher and Adult Education*, Routledge, 2018, pp. 100-120. https://doi.org/10.4324/9780429430930-8

[4] E. Dickey and A. Bejarano, "GAIDE: A Framework for Using Generative AI to Assist in Course Content Development," in 2024 IEEE Frontiers in Education Conf. (FIE), Oct. 2024, pp. 1-9. https://doi.org/10.1109/FIE61694.2024.10893132

[5] J. Singer, R. W. Marx, J. Krajcik, and J. C. Chambers, "Constructing extended inquiry projects: Curriculum materials for science education reform," Educ. Psychol., vol. 35, no. 3, pp. 165-178, 2000.https://doi.org/10.1207/S15326985EP3503_3

[6] R. A. Reiser, "A history of instructional design and technology: Part II: A history of instructional design," Educ. Technol. Res. Dev., vol. 49, no. 2, pp. 57-67, 2001. https://doi.org/10.1007/BF02504928

[7] J. W. Pellegrino, J. D. Bransford, and M. S. Donovan, *How People Learn: Bridging Research and Practice*, Washington, DC: National Academies Press, 1999.

[8] N. Friesen, *The Textbook and the Lecture: Education in the Age of New Media*, Baltimore, MD: JHU Press, 2017.

[9] A. J. Nitko, "Conceptual frameworks to accommodate the validation of rapidly changing requirements for assessments," Curriculum Assess., vol. 1, no. 2, pp. 143-163, 2001.

[10] J. L. Hess and M. S. Whittington, "Developing an effective course syllabus," Nacta J., pp. 23-27, 2003.

[11] K. de Fine Licht, "Generative artificial intelligence in higher education: Why the 'banning approach' to student use is sometimes morally justified," Philos. Technol., vol. 37, no. 3, p. 113, 2024. https://doi.org/10.1007/s13347-024-00799-9

[12] L. McNeill, "Automation or innovation? A generative AI and instructional design snapshot," in The IAFOR Int. Conf. on Education-Hawaii 2024 Official Conf. Proc., 2024, pp. 187-194. https://doi.org/10.22492/issn.2189-1036.2024.17

[13] R. R. Hansen et al., "The effects of an AI feedback coach on students' peer feedback quality, composition, and feedback experience," Læring og Medier, vol. 31, pp. 1-31, 2025. https://doi.org/10.7146/lom.v17i31.148831

[14] C. Lammert, S. DeJulio, S. Grote-Garcia, and L. M. Fraga, "Better than Nothing? An Analysis of AI-Generated Lesson Plans Using the Universal Design for Learning & Transition Frameworks," The Clearing House: A Journal of Educational Strategies, Issues and Ideas, vol. 97, no. 5, pp. 168-175, 2024. https://doi.org/10.1080/00098655.2024.2427332

[15] A. Bucchiarone et al., "Designing and Generating Lesson Plans combining Open Educational Content and Generative AI," in Proc. 27th Int. Conf. Model Driven Eng. Lang. Syst., ACM/IEEE, 2024, pp. 78-86. https://doi.org/10.1145/3652620.3687773

[16] A. Walter, "Utilizing Language-Generating Artificial Intelligence in Educational Planning: A Case Study," J. Interdiscip. Teacher Leadersh., vol. 8, no. 1, pp. 29-59, 2024.

[17] A. Przegalinska and T. Triantoro, *Converging Minds: The Creative Potential of Collaborative AI*. Boca Raton, FL: CRC Press, 2024. https://doi.org/10.1201/9781032656618

[18] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," IEEE Access, vol. 8, pp. 75264-75278, 2020. https://doi.org/10.1109/ACCESS.2020.2988510

[19] M. Javaid, A. Haleem, R. P. Singh, S. Khan, and I. H. Khan, "Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system," BenchCouncil Trans. Benchmarks, Stand. Eval., vol. 3, no. 2, Art. 100115, 2023. https://doi.org/10.1016/j.tbench.2023.100115

[20] A. Plasencia and N. Navas, "MOOCs, the flipped classroom, and Khan Academy practices: The implications of augmented learning," in *Innovation and Teaching Technologies: New Directions in Research, Practice and Policy*, Springer, 2014, pp. 1-10. https://doi.org/10.1007/978-3-319-04825-3_1

[21] F. Holguin, "Exploring AI in Graphic Novels," 2024.

[22] K. Guo, Y. Zhong, D. Li, and S. K. W. Chu, "Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation," Comput. Educ., vol. 203, Art. 104862, 2023. https://doi.org/10.1016/j.compedu.2023.104862

[23] U. Lee et al., "Can ChatGPT be a debate partner? Developing ChatGPT-based application "DEBO" for debate education, findings and limitations," Educ. Technol. Soc., vol. 27, no. 2, pp. 321-346, 2024.

[24] D. D. Stevens, *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning*, 2nd ed., New York, NY: Routledge, 2023.

[25] J. Gnanaprakasam and R. Lourdusamy, "The role of AI in automating grading: Enhancing feedback and efficiency," in *Artificial Intelligence and Education-Shaping the Future of Learning*, IntechOpen, 2024, pp. 153-167. https://doi.org/10.5772/intechopen.1005025

[26] I. T. Awidi, "Comparing expert tutor evaluation of reflective essays with marking by generative artificial intelligence (AI) tool," Comput. Educ. Artif. Intell., vol. 6, Art. 100226, 2024. https://doi.org/10.1016/j.caeai.2024.100226

[27] M. Bazluki, K. Gyabak, and B. Uderman, "Instructor feedback on a formal online course quality assurance review process," Online J. Distance Learn. Admin., vol. 21, no. 2, pp. 1-9, 2018.

[28] W. W. Lee, "Subject Matter Experts and Instructional Designers: Making Distinctions," Perform. Instr., vol. 33, no. 8, pp. 23-25, 1994. https://doi.org/10.1002/pfi.4160330807

[29] R. Mancilla and B. Frey, "Instructional design staffing for online programs," Online J. Distance Learn. Admin., vol. 26, no. 2, Art. n2, 2023.

[30] M. Smith, J. Turner, E. Sanford-Moore, and H. H. Koons, "The Lexile framework for reading: An introduction to what it is and how to use it," in Pacific Rim Objective Measurement Symp. 2015 Conf. Proc., 2016, pp. 411-424. https://doi.org/10.1007/978-981-10-1687-5_27

[31] E. Hoes, S. Altay, and J. Bermeo, "Leveraging ChatGPT for efficient fact-checking," PsyArXiv, Apr. 2023. https://doi.org/10.31234/osf.io/qnjkf

[32] Q. Huang, C. Lv, L. Lu, and S. Tu, "Evaluating the Quality of AI-Generated Digital Educational Resources for University Teaching and Learning," Systems, vol. 13, no. 3, Art. 174, 2025. https://doi.org/10.3390/systems13030174

[33] K. Shattuck, "Focusing research on quality matters," Am. J. Distance Educ., vol. 29, no. 3, pp. 155-158, 2015. https://doi.org/10.1080/08923647.2015.1061809

[34] H. A. Seneviratne and M. Thenabadu, "Evaluation of Student Perceptions on "Muddiest Point" Classroom Assessment Technique Implemented As A Formative Assessment Method," Int. J. Humanit. Art Soc. Stud., vol. 6, no. 1, 2021.

[35] N. Scaria, S. Dharani, and D. Subramani, "Automated Educational Question Generation at Different Bloom's Skill Levels Using Large Language Models: Strategies and Evaluation," in *Int. Conf. on Artificial Intelligence in Education*, Cham: Springer, 2024, pp. 165-179. https://doi.org/10.1007/978-3-031-64299-9_12

[36] S. Moore, E. Costello, H. A. Nguyen, and J. Stamper, "An Automatic Question Usability Evaluation Toolkit," in *Artificial Intelligence in Education*, A. M. Olney et al., Eds., Cham: Springer Nature Switzerland, 2024, pp. 31-46. https://doi.org/10.1007/978-3-031-64299-9_3

[37] E. Costello, J. C. Holland, and C. Kirwan, "Evaluation of MCQs from MOOCs for common item writing flaws," BMC Res. Notes, vol. 11, Art. 849, Dec. 2018. [Online]. Available: https://doi.org/10.1186/s13104-018-3959-4. https://doi.org/10.1186/s13104-018-3959-4

[38] R. Meissner et al., "LLM-generated competence-based e-assessment items for higher education mathematics: methodology and evaluation," Front. Educ., vol. 9, Art. 1427502, Oct. 2024. https://doi.org/10.3389/feduc.2024.1427502

[39] S.-V. Oprea, I. Nica, A. Bâra, and I.-A. Georgescu, "Are skepticism and moderation dominating attitudes toward AI-based technologies?" Am. J. Econ. Sociol., vol. 83, no. 3, pp. 567-607, 2024. https://doi.org/10.1111/ajes.12565

[40] A. Kulal, H. U. Rahiman, H. Suvarna, A. Abhishek, and S. Dinesh, "Enhancing public service delivery efficiency: Exploring the impact of AI," J. Open Innov. Technol. Mark. Complex., vol. 10, no. 3, Art. 100329, 2024. https://doi.org/10.1016/j.joitmc.2024.100329

[41] S. K. Amemasor et al., "A systematic review on the impact of teacher professional development on digital instructional integration and teaching practices," Front. Educ., vol. 10, Art. 1541031, 2025. https://doi.org/10.3389/feduc.2025.1541031

[42] M. Diliberti et al., *Using Artificial Intelligence Tools in K-12 Classrooms*, RAND, 2024.

[43] S. Triberti, R. Di Fuccio, C. Scuotto, E. Marsico, and P. Limone, ""Better than my professor?" How to develop AI tools for higher education," Front. Artif. Intell., vol. 7, Art. 1329605, Apr. 2024. https://doi.org/10.3389/frai.2024.1329605

[44] G. W. Soad, N. F. Duarte Filho, and E. F. Barbosa, "Quality evaluation of mobile learning applications," in *2016 IEEE Frontiers in Education Conference (FIE)*, 2016, pp. 1-8. https://doi.org/10.1109/FIE.2016.7757540

[45] X. Zeng, X. Peng, and C. Lu, "Survey on the quality assessment factors of educational APP," in *2017 Int. Symp. on Educational Technology (ISET)*, 2017, pp. 196-200. https://doi.org/10.1109/ISET.2017.52

[46] R. Topal and F. Shargh, "Teaching students how to find and identify reliable online sources: A series of exercises," J. Polit. Sci. Educ., vol. 19, no. 3, pp. 475-484, 2023. https://doi.org/10.1080/15512169.2022.2163899

[47] E. Mazzullo, O. Bulut, T. Wongvorachan, and B. Tan, "Learning Analytics in the Era of Large Language Models," Analytics, vol. 2, no. 4, pp. 877-898, Nov. 2023. https://doi.org/10.3390/analytics2040046

[48] B. A. Schwendimann et al., "Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research," IEEE Trans. Learn. Technol., vol. 10, no. 1, pp. 30-41, 2017. https://doi.org/10.1109/TLT.2016.2599522

[49] Z. Wang, J. Valdez, D. B. Mallick, and R. G. Baraniuk, "Towards Human-Like Educational Question Generation with Large Language Models," in *Artificial Intelligence in Education*, M. M. Rodrigo et al., Eds., Cham: Springer Int'l Publishing, 2022, pp. 153-166. https://doi.org/10.1007/978-3-031-11644-5_13

[50] M. C. Fink, S. A. Robinson, and B. Ertl, "AI-based avatars are changing the way we learn and teach: benefits and challenges," Front. Educ., vol. 9, Art. 1416307, 2024. https://doi.org/10.3389/feduc.2024.1416307

[51] A. Kokala, "Revolutionizing Content Creation: Leveraging AI-Driven Podcast Generation with NotebookLM and Personalized Insights," Int. Res. J. Mod. Eng. Technol. Sci., vol. 6, 2024.

[52] S. Wang, Z. Ning, A. Truong, M. Dontcheva, D. Li, and L. B. Chilton, "PodReels: Human-AI Co-Creation of Video Podcast Teasers," in Proc. 2024 ACM Designing Interactive Systems Conf., 2024, pp. 958-974. https://doi.org/10.1145/3643834.3661591

[53] P. Kannan and D. Zapata-Rivera, "Facilitating the use of data from multiple sources for formative learning in the context of digital assessments: informing the design and development of learning analytic dashboards," in *Frontiers in Education*, vol. 7, 2022, Art. 913594. https://doi.org/10.3389/feduc.2022.913594

[54] R. Bodily, J. Kay, V. Aleven, I. Jivet, D. Davis, F. Xhakaj, and K. Verbert, "Open learner models and learning analytics dashboards: a systematic

review," in *Proc. 8th Int. Conf. on Learning Analytics and Knowledge*, 2018, pp. 41-50. https://doi.org/10.1145/3170358.3170409

[55] I. Molenaar and C. A. N. Knoop-van Campen, "How teachers make dashboard information actionable," IEEE Trans. Learn. Technol., vol. 12, no. 3, pp. 347-355, 2018. https://doi.org/10.1109/TLT.2018.2851585

[56] M. Sahin and D. Ifenthaler, "Examining of Learners' Dashboard Interaction in Computer Classification Testing Environment," in *2022 Int. Conf. on Advanced Learning Technologies (ICALT)*, 2022, pp. 152-154. https://doi.org/10.1109/ICALT55010.2022.00052

[57] N. Valle, P. Antonenko, K. Dawson, and A. C. Huggins-Manley, "Staying on target: A systematic literature review on learner-facing learning analytics dashboards," Br. J. Educ. Technol., vol. 52, no. 4, pp. 1724-1748, 2021. https://doi.org/10.1111/bjet.13089

[58] B. Gras, A. Brun, and A. Boyer, "For and by student dashboards design to address dropout," in Companion Proc. 10th Int. Conf. Learn. Analytics & Knowledge, Workshop on Addressing Drop-Out Rates in Higher Education (ADORE '20), 2020.

[59] E. M. Queiroga et al., "A learning analytics approach to identify students at risk of dropout: A case study with a technical distance education course," Appl. Sci., vol. 10, no. 11, Art. 3998, 2020. https://doi.org/10.3390/app10113998

[60] J. Smink and F. P. Schargel, *Helping Students Graduate: A Strategic Approach to Dropout Prevention*, Eye on Education, 2004.

[61] G. Akçapınar and M. N. Hasnine, "Discovering the effects of learning analytics dashboard on students' behavioral patterns using differential sequence mining," Procedia Comput. Sci., vol. 207, pp. 3818-3825, 2022. https://doi.org/10.1016/j.procs.2022.09.443

[62] C. McGrath, A. Farazouli, and T. Cerratto-Pargaman, "Generative AI chatbots in higher education: a review of a merging research area," Higher Education, 2024, pp. 1-17. https://doi.org/10.1007/s10734-024-01288-w

[63] D. Ravšelj et al., "Higher education students' perceptions of ChatGPT: A global study of early reactions," PLOS ONE, vol. 20, no. 2, Art. e0315011, Feb. 2025.

[64] S. D'Urso, F. Sciarrone, and M. Temperini, "Boulez: a chatbot-based federated learning system for distance learning," in Proc. Int. Conf. Information Visualisation (IV), IEEE, 2023, pp. 210-215. https://doi.org/10.1109/IV60283.2023.00045

[65] V. De Marco, F. Sciarrone, and M. Temperini, "TutorChat: a chatbot for the support to dyslexic learner's activity through Generative AI," in Proc. Int. Conf. Advanced Learning Technologies (ICALT), IEEE, 2024, pp. 155-157. https://doi.org/10.1109/ICALT61570.2024.00051

[66] D. Sun, A. Boudouaia, C. Zhu, and Y. Li, "Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study," Int. J. Educ. Technol. High. Educ., vol. 21, Art. 14, Feb. 2024. https://doi.org/10.1186/s41239-024-00446-5

[67] H. Güner and E. Er, "AI in the classroom: Exploring students' interaction with ChatGPT in programming learning," Educ. Inf. Technol., vol. 1, 2025, doi:10.1007/s10639-025-13337-7. https://doi.org/10.1007/s10639-025-13337-7

[68] M. Khalifa and M. Albadawy, "Using artificial intelligence in academic writing and research: An essential productivity tool," Comput. Methods Programs Biomed. Update, vol. 5, Art. 100145, Jan. 2024, doi:10.1016/j.cmpbup.2024.100145. https://doi.org/10.1016/j.cmpbup.2024.100145

[69] E. Cano-Pleite et al., "Exploring the Integration of an AI-Assisted Midterm Exam in a Thermal Engines Course," in Proc. 17th Annu. Int. Conf. Educ. Res. Innovation (ICERI2024), Seville, Spain, IATED, 2024, pp. 6877-6881. https://doi.org/10.21125/iceri.2024.1659

[70] G. Kortemeyer and J. Nöhl, "Assessing confidence in AI-assisted grading of physics exams through psychometrics: An exploratory study," Phys. Rev. Phys. Educ. Res., vol. 21, no. 1, p. 010136, 2025. https://doi.org/10.1103/PhysRevPhysEducRes.21.010136

[71] S. Dhara, S. Chatterjee, R. Chaudhuri, A. Goswami, and S. K. Ghosh, "Artificial intelligence in assessment of students' performance," in *Artificial Intelligence in Higher Education*, CRC Press, 2022, pp. 153-167. https://doi.org/10.1201/9781003184157-8

[72] L. Zheng et al., "Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena," arXiv:2306.05685, 2023.

[73] J. Floden, "A comparison between human and AI grading of exams in higher education using ChatGPT," Br. Educ. Res. J., vol. 51, pp. 201-224, 2025. https://doi.org/10.1002/berj.4069

[74] W. Hersh and K. F. Hollis, "Results and implications for generative AI in a large introductory biomedical and health informatics course," npj Digital Medicine, vol. 7, Art. 247, 2024. https://doi.org/10.1038/s41746-024-01251-0

[75] A. Pack, A. Barret, and J. Escalante, "Large language models and automated essay scoring of English language learner writing: insights into validity and reliability," Comput. Educ. Artif. Intell., vol. 6, pp. 1-9, 2024. https://doi.org/10.1016/j.caeai.2024.100234

[76] M. Stahi, L. Biermann, A. Nehring, and H. Wachsmuth, "Exploring LLM prompting strategies for joint essay scoring and feedback generation," in Proc. Workshop on Innovative Use of NLP for Building Educational Applications (BEA), 2024, pp. 1-16.

[77] W. Li and H. Liu, "Applying large language models for automated essay scoring for non-native Japanese," Humanities Soc. Sci. Commun., vol. 11, Art. 723, 2024. https://doi.org/10.1057/s41599-024-03209-9

[78] A. Nitze, "Future-proofing Education: a prototype for simulating oral examinations using large language models," in Proc. STEMEIC '23 Conf., 2023, pp. 1-6.

[79] S. I. Salim, R. Y. Yang, A. Cooper, S. Ray, S. K. Debray, and S. Rahaman, "Impeding LLM-assisted Cheating in Introductory Programming Assignments via Adversarial Perturbation," in Proc. EMNLP 2024, 2024, pp. 445-463. https://doi.org/10.18653/v1/2024.emnlp-main.27

[80] M. El-Muwalla and A. Badran, "Turnitin: Building Academic Integrity Against Plagiarism to Underpin Innovation," in *Higher Education in the Arab World*, Springer, 2020, pp. 261-268. https://doi.org/10.1007/978-3-030-37834-9_11

[81] R. Crockett and K. Best, "Stylometric Comparison of Professionally Ghost-Written and Student-Written Assignments," in *Integrity in Education for Future Happiness*, European Network for Academic Integrity, 2020, pp. 35-49. https://doi.org/10.11118/978-80-7509-772-9-0035

[82] M. Psyridou, F. Prezja, M. Torppa, M.-K. Lerkkanen, A.-M. Poikkeus, and K. Vasalampi, "Machine Learning Predicts Upper Secondary Education Dropout as Early as the End of Primary School," Sci. Rep., vol. 14, Art. 1-12, 2024. https://doi.org/10.1038/s41598-024-63629-0

[83] N. U. R. Junejo, M. W. Nawaz, Q. Huang, X. Dong, C. Wang, and G. Zheng, "Accurate Multi-Category Student Performance Forecasting at Early Stages of Online Education Using Neural Networks," arXiv:2412.05938, 2024. https://doi.org/10.1038/s41598-025-00256-3

[84] A. Smolansky, A. Cram, C. Raduescu, S. Zeivots, E. Huber, and R. F. Kizilcec, "Educator and student perspectives on the impact of generative AI on assessments in higher education," in *Proc. 10th ACM Conf. Learn. @ Scale (L@S '23)*, Copenhagen, Denmark, 2023, pp. 378-382. https://doi.org/10.1145/3573051.3596191

[85] T. H. Bers, "Assessing Critical Thinking in Community Colleges," New Dir. Community Colleges, vol. 2005, pp. 15-25, 2005. https://doi.org/10.1002/cc.192

[86] E. Mazzullo and O. Bulut, "Automated Feedback Generation for Open-Ended Questions: Insights from Fine-Tuned LLMs," in Proc. FM-EduAssess at NeurIPS 2024 Workshop, PMLR, Edmonton, AB, Canada, 2024, pp. 1-18.

[87] Q. Jia, M. Young, Y. Xiao, J. Cui, C. Liu, P. Rashid, and E. Gehringer, "Automated Feedback Generation for Student Project Reports: A Data-Driven Approach," in Proc. 54th ACM Tech. Symp. Comput. Sci. Educ. (SIGCSE 2023), Toronto, ON, 2023, pp. 916-922.

[88] S. C. E. Fung, M. F. Wong, and C. W. Tan, "Automatic Feedback Generation on K-12 Students' Data Science Education by Prompting Cloud-Based LLMs," in Proc. 18th Int. Conf. Educ. Data Mining (EDM 2024), 2024. https://doi.org/10.1145/3657604.3664673

[89] T. Ruwe and E. M.-Paus, "Embracing LLM feedback: the role of feedback providers and provider information for feedback effectiveness," Front. Educ., vol. 9, Art. 1461362, 2024. https://doi.org/10.3389/feduc.2024.1461362

[90] A. Kinder, F. J. Briese, M. Jacobs, N. Dern, N. Glodny, S. Jacobs, and S. Leßmann, "Effects of adaptive feedback generated by a large language model: A case study in teacher education," Comput. Educ. Artif. Intell., vol. 8, Art. 100349, 2025. https://doi.org/10.1016/j.caeai.2024.100349

[91] S. Riazi and P. Rooshenas, "LLM-Driven feedback for enhancing conceptual design learning in database systems courses," in *Proc. SICSE TS*, 2025, pp. 1-7. https://doi.org/10.1145/3641554.3701940