Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #48417

BOARD # 99: Work in Progress: AI in online laboratory teaching - A Systematic Literature Review

Mr. Johannes Kubasch, University of Wuppertal

Johannes Kubasch is a mechanical engineer and research associate at the Chair of Technical and Engineering Education at the University of Wuppertal. As a engineer in automotive engineering, he initially worked in the automotive supply industry in the development of airbag systems before moving to the University of Wuppertal to work in the field of engineering education. In the past, he worked on the AdeLeBk.nrw project to digitize the university training of prospective teachers at vocational schools and to adapt the learning content to the requirements of later professional practice.

Dr. Dominik May, University of Wuppertal

Dr. May serves as a Professor for Technical Education and Engineering Education Research at the School of Mechanical Engineering and Safety Engineering at University of Wuppertal. His work revolves around generating both fundamental and practical knowledge that defines, informs, and enhances the education of engineers.

His primary research thrust centers around the development, implementation, practical utilization, and pedagogical value of online laboratories. These laboratories span a range of formats, including remote, virtual, and cross-reality platforms. Dr. May's scholarly pursuits extend into the sphere of online experimentation, particularly within the context of engineering and technical education. Prior to his role at the University of Wuppertal, Dr. May held the position of Assistant Professor within the Engineering Education Transformations Institute at the University of Georgia (Athens, GA, USA).

Central to Dr. May's scholarly endeavors is his commitment to formulating comprehensive educational strategies for Technical and Engineering Education. His work contributes to the establishment of an evidence-based foundation that guides the continual transformation of Technical and Engineering Education. Additionally, Dr. May is actively involved in shaping instructional concepts tailored to immerse students in international study contexts. This approach fosters intercultural collaboration, empowering students to cultivate essential competencies that transcend cultural boundaries.

Beyond his academic role, Dr. May assumes the position of President at the "International Association of Online Engineering (IAOE)," a nonprofit organization with a global mandate to advocate for the broader advancement, distribution, and practical application of Online Engineering (OE) technologies. His leadership underscores his commitment to leveraging technological innovation for societal progress. Furthermore, he serves as the Editor-in-Chief for the "International Journal of Emerging Technologies in Learning (iJET)," a role that facilitates interdisciplinary discussions among engineers, educators, and engineering education researchers. These discussions revolve around the interplay of technology, instruction, and research, fostering a holistic understanding of their synergies.

Dr. May is an active member of the national and international scientific community in Engineering Education Research. He has also organized several international conferences himself – such as the annual "International Conference on Smart Technologies & Education (STE)" – and serves as a board member for further conferences in this domain and for several Divisions within the American Society for Engineering Education.

Doha Meslem, Bergische Universität Wuppertal

WIP: AI in Online Laboratory Teaching - A Systematic Literature Review

Introduction

The presence of ChatGPT has recently, and in a short period of time, become increasingly prevalent in the day-to-day life. Education, being a part and a reflection of the day-to-day life, has therefore also been affected by this change. The fast spread of this technology within this context has however come with its challenges. These include the lack of an adequate understanding of it, of how to use it, and how to integrate it in an efficient way in the daily life (Gill & Kaur, 2023). Many students across disciplines try to make use of such technologies within their education (Farhi et al., 2023). Although this is encouraged, the correct tools are sometimes missing. Additionally, the access to and prevalence of such a technology has eradicated certain aspects and introduced new ones in education, meaning, reshaping education. Aspects that were influenced include memorization, grammar and writing skills, coding, language learning, STEM concepts visualization and understanding to name a few. Consequently, and given that the use of this technology is encouraged, the educational approach has evolved to emphasize different aspects than in the past, as this project will demonstrate. This project focuses on engineering education as a case study to explore this transformation and its integration into the curriculum, preparing well-equipped and versatile engineers A case study of laboratory within this context was chosen to narrow the focus and get a better understanding of it. To further engage within the process of laboratory-based instruction, feedback; interactive and adaptive, plays the major role. To further this project however and set solid grounds for educating about and within this topic, a thorough study of the current state in research needs to be done. This necessitates the presence of this paper as an initial step towards building better understanding of what is lacking and what is expected in this field. This paper studies which aspects were present in research, which present a gap, and what conclusions can be drawn.

This study is part of the KICK 4.0 project, which explores the integration of AI-based Natural Language Processing (NLP) systems in engineering education, particularly within laboratory environments. Specifically, the goal is to evaluate how NLP-powered real-time feedback can enhance laboratory-based education, while also identifying the limitations and challenges of such AI-human collaboration.

While artificial intelligence in education is well studied as shown by Chen et al. (2020) and later in this paper, the application of NLP and NLP-similar approaches in laboratory-based learning focused on real time feedback for analyzing and enhancing the approaches remains understudied.

Natural Language Processing (NLP), a subfield of artificial intelligence, focuses on enabling machines to interpret, understand, and generate human language. It involves tasks such as text classification, sentiment analysis, language translation, and text generation, all aimed at making human-computer interaction more natural. The presence of NLP can be observed in virtual assistants, search engines, transcription tools, and educational technologies, reflecting its widespread integration in both daily life and learning environments. (Nadkarni et al., 2011; Chopra, Prashar, & Sain, 2013; Joseph et al., 2016)

The broader relevance of this study is hence the enhancing personalized feedback and student learning outcomes. The results and outcomes of this study would abridge the gap in research. The presented framework can additionally act as fundamental for further research in this field and help educators and researchers in related STEM-fields develop similar strategies, ensuring reproducibility and interdisciplinarity.

The main goal of this project is to provide a comprehensive overview of NLP applications in laboratory-based engineering education. To achieve this, the study explores the following key questions: How are AI and NLP systems utilized in laboratory-based instruction? In what ways do these systems influence the quality of feedback? Additionally, how can users be effectively trained to make the most of NLP tools?

This paper is divided into a theoretical and an applicational aspect. It tackles the above-mentioned aspects by first introducing the background, the necessity, and the purpose of the paper. The applicational aspect is divided into methodology and the discussion of the outcomes.

Methodology

To conduct a systematic and thorough literature review, a standardized nine-step process was followed to ensure all necessary aspects were covered. This is chosen as to allow for a standardized structured approach as shown by Heil (2021). Each step in the process builds upon the previous one, forming a coherent and comprehensive workflow.

It begins with determining the research principle, which involves defining the overall objective to ensure alignment with the research question. With alignment to the focus of the project as stated earlier, NLP in engineering education focused on laboratory-based scenarios, the principle here is a systematic compilation of the current state of research regarding the use of AI in engineering education, or sensitive literature research.

This is followed by the definition of the search components. This involves breaking down the research question into key elements. In this study, the search components are defined by 11 main keywords. The keywords are lecturers and students, NLP and AI, higher education, potential, risk and limitation, feedback, competencies, and laboratory teaching. Their synonyms were also used, including ChatGPT, generative AI, engineering pedagogy, technical pedagogy, advantages, changes, possibilities, critic, disadvantages, experiences, digital laboratories and online laboratories to name some.

After this, databases are selected based on specific criteria. The main databases considered include google scholar, scopus, science direct, web of science, nautos and espacnet. The studied databases so far were however scopus, science direct, and web of science based on relevance, scope, and reproducibility within our context.

The next step is identifying keywords, which involves selecting terms that reflect the core ideas of the research. This ensures that the search does not miss studies due to different terminologies. This is followed by the identification of descriptors, which are subject-specific terms from indexing systems or academic taxonomies that enhance the precision of the search.

Developing the search string follows, which involves combining the keywords and descriptors into structured queries using Boolean operators such as “AND,” “OR,” and

“NOT.” The same search strings were used for all databases; however, the subjects or research areas differ as will be shown in the upcoming sections. This is followed and intertwined with the next step which is validating the search string.

Validating the search string involves conducting initial test searches to check and refine the queries, ensuring they capture only the relevant results. The search components are applied in the following sequence: initially, all components are used. The focus then shifts to boundaries, followed by a combination of boundaries and laboratory. These components and their combinations are adjusted based on the results by modifying Boolean operators and evaluating the outcomes as part of the validation process.

Once validated, the research is conducted by running the finalized search strings across the selected databases. The search strings, including both keywords and synonyms, were documented, improved and validated, repeated for all other databases and checked for reproducibility.

The next step involves documenting and exporting the search results, which includes systematically recording all key steps and outcomes, such as the search parameters, the number of results obtained for each database, the used search string, and the dates. This step ensures transparency and enables the process to be replicated. Additionally for this paper, the exported search results for each database are authors, title, paper specifications, DOI, source, abstract and keywords.

The final step is reviewing and refining search results by screening studies and excluding irrelevant sources. The inclusion criteria for "education as a whole" focus on higher education teaching, engineering education, university education, higher education pedagogy, and engineering pedagogy. For "education in universities," it includes laboratory environments, digital, online, cross, mixed reality, and computer labs. The target group consists of students, teachers, and educators. The results included work in progress papers, as well as done studies. The time was limited from 2020 to present, and the type included specifically only scientific articles. By following this nine-step process, the literature review was designed to be systematic, transparent, and aligned with the research objectives, forming a robust foundation for the research presented.

In the process of systematically reviewing articles retrieved from the database, a final manual filtering step was conducted, involving a detailed examination of keywords, titles, and abstracts. This approach led to the development of a structured Relevance Categorization System to better assess and rank the relevance of the articles within our research context.

The priority framework consists of four key priority levels. The first is NLP and ChatGPT-related research, which is the highest priority. The second is engineering education and teaching in higher education. The third is digital and online laboratory contexts. Lastly, the fourth includes competencies, limitations, boundaries, and feedback mechanisms.

The relevance levels are divided into three categories based on their alignment with key research priorities. These categories indicate the extent to which an article aligns with these priorities, ranging from high to low relevance. The table 1 below provides an overview of these levels and their criteria.

Relevance Level	Criteria
G **(High Relevance)**	Assigned to articles that comprehensively address priorities 1, 2, and 3 together, or include all four priorities (1, 2, 3, and 4).
R **(Low Relevance)**	Assigned to articles that either lack all key priorities or focus solely on one aspect (Priority 2, 3, or 4 individually). Also includes articles on general machine learning (ML) that do not specifically cover NLP.
Y **(Moderate Relevance)**	sub-levels: Y: Articles covering combinations: (1 + 2), (1 + 3), and (2 + 3), emphasizing that engineering education (2) and laboratory contexts (3) are relevant only when paired with NLP (1). Combinations: (1 + 2 + 4) and (1 + 3 + 4) are treated as Y, reflecting that the inclusion of Priority 4 enhances relevance only when accompanied by Priority 1 and other critical priorities. YY: Articles focusing solely on Priority 1 (NLP and ChatGPT). Articles combining (1 + 4) are categorized at the same level as YY due to the shifted contextual focus.

Table 1

Additional cases include papers on AI applications in education (Priority 2) and labs (Priority 3) without NLP-specific discussions. These were labeled as YAI, indicating partial but secondary relevance in the context of our NLP-focused analysis.

This system ensures an evaluation of the articles, aligning their content with the research objectives and refining the focus based on their thematic fit within the established priorities. This framework is reproducible for similar studies.

Discussion of Results and Learning Outcomes

The science direct database research included 3070 examined articles, where the used subject areas are: Computer science, Engineering, and Social Sciences.

Out of these articles, more than 86% were classified as irrelevant, falling under the lowest relevance rank. A significant number of these articles misinterpreted the term "NLP" to mean unrelated concepts, such as "Noise-Like Pulses" as seen in articles Wang et al. (2025) and Tao et al. (2023) for example or discussed applications entirely unrelated to the research focus.

Additionally, over 350 articles studied NLP in different contexts, thereby reducing their relevance. For instance, Jiang and Wang (2024) in their article Railway accident causation prediction, apply NLP for railway accident analysis using transformer-based architecture. Similarly, Ren et al. (2020), in their paper use NLP for sarcasm detection, while Predices et al. (2021) in their study apply it for web browsing analytics. These illustrate how NLP is often applied in diverse, non-educational contexts, emphasizing the need to identify and separate relevant research from peripheral studies.

The second relevance level is represented by a much smaller proportion of papers, averaging around 20 per 1,000 articles. Many general studies exist in literature, such as Gill et al. (2024), Shorey et al. (2024), and Hsu & Silalahi (2024) to name a few, with the focus broadly on ChatGPT, bots, and their societal effects without specific ties to education or laboratory contexts.

Considerable amount of literature aligns more closely with educational applications from the educators' perspective. Du et al. (2024), explore using NLP and large language models (LLMs) to automatically evaluate student project reports. Similarly, Caccavale et al. (2024) in their article towards education 4.0, investigate the potential of LLMs as virtual tutors in chemical engineering. Tate et al. (2024)'s study examines the extent to which AI provides holistic essay scoring, while White et al. (2023) research focuses on assessing chemistry knowledge. These studies highlight LLMs as tools supporting educators, either by assisting with instructional tasks, as in the former papers, or by enhancing grading processes, as in the latter mentioned papers.

NLP's role from the students' perspectives is also present in literature. For instance, Li et al. (2020) examine the potential of text summarization using NLP, while articles such as Gayed (2022) and Zhai & Wibowo (2023) assess its impact on English language learners. These examples demonstrate how NLP is being integrated into educational processes for the advantage of students. Often, however, without addressing the full context relevant to our research goals, further emphasizing the identified research gap.

Less than 20 papers were categorized within the AI without NLP category, comprising studies on general AI applications in education. Examples include a study done by Bernius et al. (2022), which examines machine learning (ML)-based feedback on student answers in large courses, Cunha et al. (2024) which explores ML-enhanced geometry instruction in basic education, and Xing et al. (2023), which discusses the use of AI for thermodynamics teaching.

Finally, the highest relevance level was absent in the reviewed papers, highlighting a critical research gap and further underscoring the need for this study.

In comparison to studies focusing on education, studies focused on laboratory-based education are notably scarce, despite the context of this research's case study, with an overall ratio of five papers in education to one in laboratory research. Many papers in this category remain general, focusing on digitalization or AI in education rather than specific instructional contexts. For instance, Udugama et al. (2023) discusses digital tools in chemical engineering education, Appels et al. (2024) reimagines educational quality in the context of digital transformation, and Khosravi et al. (2022) examines explainable AI in education. In the context of laboratories, AI-based studies are similarly minimal, with article published by Hysmith et al. (2024) being one of the few examples.

A broader trend observed is that chemical engineering appears more frequently in studies within engineering education compared to other disciplines. Additionally, a general trend that's observed suggests that even within the most relevant articles, the dominant research appears concentrated around model development and improvements in language-based tasks. There's relatively less emphasis on context-specific applications like laboratory education or engineering pedagogy, reinforcing the idea that much of the research remains abstract or tool-focused rather than domain-integrated.

Moreover, the data suggests that research on AI/NLP applications in engineering contexts, particularly related to education and lab work, is underrepresented compared to studies in non-engineering disciplines.

Conclusion and Future Outlook

This paper sheds light on the crucial role that AI, particularly NLP, plays in reshaping engineering education, with an emphasis on laboratory-based learning. The rapid integration of tools like ChatGPT has redefined educational dynamics, affecting feedback, assessment, and the overall learning process. However, while these tools bring potential, they also highlight a pressing challenge how to implement them within engineering labs.

The systematic review conducted here shows a distinct research gap. Although studies on NLP in education are well founded, their application in lab environments is minimal. Most studies are focused on improving NLP models or addressing narrow use cases without considering how they fit within engineering education’s practical framework. Engineering-specific studies are outnumbered by broader research in other disciplines.

Future research should expand the number of analyzed databanks and prioritize interdisciplinary collaboration to refine feedback mechanisms. Future research should not only explore AI’s integration into laboratories but also define structured frameworks for its application. One possible direction is the development of an NLP-based tutoring system, which could assist students in real-time laboratory work by providing structured feedback, explanations, and recommendations. Such a system could be tested within the fluid mechanics laboratory setting, where it could guide students through experimental design, data analysis, and troubleshooting processes. Implementing and evaluating such frameworks would bridge the existing research gap and provide practical solutions for integrating AI into laboratory-based education.

References

A. Chopra, A. Prashar, and C. Sain, "Natural language processing," *International Journal of Technology Enhancements and Emerging Engineering Research*, vol. 1, no. 4, p. 131, 2013. [Online]. Available: https://doi.org/ISSN2347-4289

A. D. White, G. M. Hocky, H. A. Gandhi, M. Ansari, S. Cox, G. P. Wellawatte, S. Sasmal, Z. Yang, K. Liu, Y. Singh, and W. J. Peña Ccoa, "Assessment of chemistry knowledge in large language models that generate code," *Digital Discovery*, vol. 2, no. 2, pp. 368–376, 2023. [Online]. Available: https://doi.org/10.1039/d2dd00087c

B. Jiang and K. Wang, "Railway accident causation prediction with improved transformer model based on lexical information and contextual relationships," *Knowledge-Based Systems*, vol. 296, 111897, 2024. [Online]. Available: https://doi.org/10.1016/j.knosys.2024.111897

C. R. Cunha, A. Moreira, S. Coelho, V. Mendonça, and J. P. Gomes, "Converging extended reality and machine learning to improve the lecturing of geometry in basic education," *Journal of Engineering Research*, 2024. [Online]. Available: https://doi.org/10.1016/j.jer.2024.10.016

C. Zhai and S. Wibowo, "A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university," *Computers and Education: Artificial Intelligence*, vol. 4, 100134, 2023. [Online]. Available: https://doi.org/10.1016/j.caeai.2023.100134

D. Perdices, J. Ramos, J. L. García-Dorado, I. González, and J. E. López de Vergara, "Natural language processing for web browsing analytics: Challenges, lessons learned, and opportunities," *Computer Networks*, vol. 198, 108357, 2021. [Online]. Available: https://doi.org/10.1016/j.comnet.2021.108357

E. A. Heil, *Methode der systematischen Literaturrecherche*. Justus-Liebig-Universität Gießen, 2021. [Online]. Available: http://dx.doi.org/10.22029/jlupub-17189

F. Farhi, R. Jeljeli, I. Aburezeq, F. F. Dweikat, S. A. Al-shami, and R. Slamene, "Analyzing the students' views, concerns, and perceived ethics about ChatGPT usage," *Computers and Education: Artificial Intelligence*, vol. 4, 100180, 2023. [Online]. Available: https://doi.org/10.1016/j.caeai.2023.100180

H. Du, Q. Jia, E. Gehringer, and X. Wang, "Harnessing large language models to auto-evaluate the student project reports," *Computers and Education: Artificial Intelligence*, vol. 7, 100268, 2024. [Online]. Available: https://doi.org/10.1016/j.caeai.2024.100268

H. Hysmith, E. Foadian, S. P. Padhy, S. V. Kalinin, R. G. Moore, O. S. Ovchinnikova, and M. Ahmadi, "The future of self-driving laboratories: From human-in-the-loop interactive AI to gamification," *Digital Discovery*, vol. 3, no. 4, pp. 621–636, 2024. [Online]. Available: https://doi.org/10.1039/d4dd00040d

H. Khosravi et al., "Explainable artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 3, 100074, 2022. [Online]. Available: https://doi.org/10.1016/j.caeai.2022.100074

I. A. Udugama et al., "Digital tools in chemical engineering education: The needs and the desires," *Education for Chemical Engineers*, vol. 44, pp. 63–70, 2023. [Online]. Available: https://doi.org/10.1016/j.ece.2023.05.002

J. M. Gayed, M. K. Carlon, A. M. Oriola, and J. S. Cross, "Exploring an AI-based writing assistant's impact on English language learners," *Computers and Education: Artificial Intelligence*, vol. 3, 100055, 2022. [Online]. Available: https://doi.org/10.1016/j.caeai.2022.100055

J. P. Bernius, S. Krusche, and B. Bruegge, "Machine learning based feedback on textual student answers in large courses," *Computers and Education: Artificial Intelligence*, vol. 3, 100081, 2022. [Online]. Available: https://doi.org/10.1016/j.caeai.2022.100081

L. Appels, S. De Maeyer, and P. Van Petegem, "Re-imagining educational quality. The need for a multidimensional approach in evaluating educational quality through TIMSS Data," *Studies in Educational Evaluation*, vol. 83, 101409, 2024. [Online]. Available: https://doi.org/10.1016/j.stueduc.2024.101409

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

L. Li, J. Geissinger, W. A. Ingram, and E. A. Fox, "Teaching natural language processing through big data text summarization with problem-based learning," *Data and Information Management*, vol. 4, no. 1, pp. 18–43, 2020. [Online]. Available: https://doi.org/10.2478/dim-2020-0003

L. Ren, B. Xu, H. Lin, X. Liu, and L. Yang, "Sarcasm detection with sentiment semantics enhanced multi-level memory network," *Neurocomputing*, vol. 401, pp. 320–326, 2020. [Online]. Available: https://doi.org/10.1016/j.neucom.2020.03.081

P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman, "Natural language processing: An introduction," *Journal of the American Medical Informatics Association*, vol. 18, no. 5, pp. 544–551, 2011. [Online]. Available: https://doi.org/10.1136/amiajnl-2011-000464

S. R. Joseph, H. Hlomani, K. Letsholo, F. Kaniwa, and K. Sedimo, "Natural language processing: A review," *International Journal of Research in Engineering and Applied Sciences*, vol. 6, no. 3, p. 207, 2016. [Online]. Available: http://www.euroasiapub.org

S. Shorey, C. Mattar, T. L.-B. Pereira, and M. Choolani, "A scoping review of ChatGPT's role in healthcare education and research," *Nurse Education Today*, vol. 135, 106121, 2024. [Online]. Available: https://doi.org/10.1016/j.nedt.2024.106121

S. S. Gill and R. Kaur, "ChatGPT: Vision and challenges," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 262–271, 2023. [Online]. Available: https://doi.org/10.1016/j.iotcps.2023.05.004

S. S. Gill et al., "Transformative effects of ChatGPT on modern education: Emerging era of AI chatbots," *Internet of Things and Cyber-Physical Systems*, vol. 4, pp. 19–23, 2024. [Online]. Available: https://doi.org/10.1016/j.iotcps.2023.06.002

T. P. Tate, J. Steiss, D. Bailey, S. Graham, Y. Moon, D. Ritchie, W. Tseng, and M. Warschauer, "Can AI provide useful holistic essay scoring?" *Computers and Education: Artificial Intelligence*, vol. 7, 100255, 2024. [Online]. Available: https://doi.org/10.1016/j.caeai.2024.100255

W.-L. Hsu and A. D. Silalahi, "Exploring the paradoxical use of ChatGPT in education: Analyzing benefits, risks, and coping strategies through integrated UTAUT and PMT theories using a hybrid approach of SEM and fsQCA," *Computers and Education: Artificial Intelligence*, vol. 7, 100329, 2024. [Online]. Available: https://doi.org/10.1016/j.caeai.2024.100329

W. Xing, X. Huang, C. Li, and C. Xie, "Teaching thermodynamics with augmented interaction and learning analytics," *Computers & Education*, vol. 196, 104726, 2023. [Online]. Available: https://doi.org/10.1016/j.compedu.2023.104726

Z. Wang et al., "Ultra-flat NIR-to-MIR supercontinuum source with all-silica fibers based on rectangular NLP pumping," *Optics & Laser Technology*, vol. 181, 111791, 2025. [Online]. Available: https://doi.org/10.1016/j.optlastec.2024.111791