ARTIFICIAL INTELLIGENCE AND DEEPFAKE LEARNING IN HIGHER EDUCATION

Branko Anđić
Johannes Kepler University, Austria

Since the second half of 2022, the hype surrounding artificial intelligence and its impact on nearly all human activities—ranging from research, science, education, industry, and economy to everyday life—has remained unabated. The beginning of 2025 has also been marked by significant AI-related plans, indicating that this hype will only continue to grow. One of the strategic initiatives of the new U.S. government is the Stargate Project, which aims to bring together major technology giants such as OpenAI and Oracle from the United States, SoftBank from Japan, and MGX from the United Arab Emirates. With a budget of $500 billion over the next four years, this initiative will focus on developing AI infrastructure in the U.S. However, other countries have also formulated investment plans for AI infrastructure and capacity. For example, France's AI Investment Plan foresees an investment of approximately €109 billion in AI over a five-year period. Meanwhile, in the United Kingdom, the AI Safety Institute (AISI) has been established, receiving £100 million in funding to evaluate and mitigate AI risks. These are just some examples illustrating that the era of intensive AI development and advancement is only beginning. Educational systems must keep pace with these developments and be prepared to harness the benefits of AI technology while also formulating strategies to mitigate its negative impact on learning and teaching processes.

Higher education appears to be undergoing the most significant transformations due to the integration of artificial intelligence into teaching, learning, feedback provision, assessment, writing, and the growing role of AI in student research. Several aspects of AI in education have already been studied, such as institutional adoption policies and guidelines (Jin et al., 2025), Intelligent Tutoring Systems, and Automated Assessment and Feedback (Ramadhan et al., 2024). On the other hand, certain topics, such as deepfakes and their consequences for learning, remain underexplored. Deepfakes refer to the use of AI to create media—such as photos, audio, and video content—that appears authentic but is, in reality, artificially generated.

Charles B. Hodges from Georgia Southern University has introduced the term deepfake learning, which directly pertains to education. According to Hodges (2024), "Deepfake learning is the phenomenon of learners using artificial intelligence to complete assessments in such a way that it misrepresents that they have done the hard work of learning, when in fact they have not done the work themselves." To some extent, the above definition can be compared to the cognitive phenomenon of illusion of competence. The illusion of competence refers to a cognitive bias in which individuals overestimate their abilities, knowledge, understanding of a specific topic, and capacity to complete a task within a given domain (Dunning et al., 2003; Metcalfe, 1998). However, a search across prestigious databases and academic search engines—including Web of Science (WoS), Scopus, Education Resources Information Center (ERIC), PsycINFO, and Google Scholar—revealed that, apart from the definition proposed by Hodges (2024), no other manuscripts have specifically investigated deepfake learning or the illusion of competence as a phenomenon emerging from AI use in teaching, learning, or academic tasks at universities. Nevertheless, nearly all universities are facing ethical challenges regarding AI use in the preparation of bachelor's, master's, and doctoral dissertations (Jin et al., 2025). Some of the

most common challenges university students encounter include: difficulties in distinguishing ethical from unethical AI use in writing and assignments; lack of knowledge on how to cite AI-generated content; plagiarism and over-reliance on AI; limited understanding of AI capabilities and limitations. However, when comparing the key characteristics of the illusion of competence with some of the impacts AI has on the learning process, a significant similarity can be observed (see Table 1).

Table 1
Key Characteristics of the Illusion of Competence and the Impact of AI on Learning

Characteristics of the Illusion of Competence	The Impact of AI on Learning
Over-reliance on easily accessible information	Students can quickly obtain relevant information using AI assistants that search vast databases and provide concise, real-time answers. AI can summarize extensive texts and key concepts, enabling students to efficiently absorb essential information from the literature.
Passive learning – Learning based on reading, recognition, and repetition without active information processing can create a false sense of confidence.	The use of AI may lead to a reduction in students' cognitive engagement with learning materials, resulting in lower cognitive effort and causing metacognitive laziness (Fan et al., 2024).
Familiarity effect – If information seems familiar (e.g., a student recognizes it in a text), they may mistakenly assume they have mastered the knowledge and can apply it in practice.	Students frequently revise and correct texts using AI tools under the false belief that they could perform the task independently without AI assistance.
Dependence on various sources of knowledge when solving tasks	The integration of AI in education may lead to learners' dependence on technology, potentially reinforcing metacognitive laziness (Fan et al., 2024).

As shown in Table 1, the application of AI in the learning process can potentially contribute to the development of deepfake learning or the illusion of competence. However, this remains a significant knowledge-theoretical and evidence-data gap, as well as a practical gap. Given the rapid advancement of AI and ambitious investment plans in this field, research on deepfake learning and the illusion of competence should be intensified. Future studies should first define deepfake learning and describe its characteristics, followed by the development of instruments to assess deepfake learning. Based on data from such research, it may be possible to develop pedagogical interventions aimed at improving metacognition in AI-supported learning and preventing the emergence of deepfake learning or the illusion of competence. One motivational factor for further research in this field could be the perspective of ChatGPT, one of the most well-known chatbots, on this topic (Figure 1).

Figure 1
Deepfake Learning as Perceived by ChatGPT

Source: OpenAI (2025). ChatGPT version 4.0. Prompt: Based on the text provided below (the two pages of this manuscript), generate a clear illustration of deepfake learning or the development of the illusion of competence).

In light of the rapid advancement of AI and its growing presence in higher education, a crucial question arises – how can we ensure that technology enhances rather than undermines authentic learning? The phenomena of deepfake learning and the illusion of competence are merely the starting points of a much deeper discussion on AI's role in

education. If we fail to establish clear methodological frameworks for researching these issues and do not design appropriate pedagogical interventions, we risk that students will increasingly replace understanding and critical thinking with passive reliance on AI tools. Instead, future research and practice should aim to create a balance—where AI serves as a support system rather than a substitute for the learning process. Educational systems have both the opportunity and responsibility to lead this transition—not only by defining new rules but also by developing innovative strategies that ensure students do not merely use AI but also understand its implications, limitations, and ethical challenges. The future of higher education depends not only on technological progress but also on the academic community's ability to direct this technology toward genuine learning and knowledge development.

References

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13544

Hodges, C. B. (2024). What is deepfake learning? [Preprint]. https://doi.org/10.13140/RG.2.2.25153.57440

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence, 8*, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348

Metcalfe, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality and Social Psychology Review, 2*(2), 100–110. https://doi.org/10.1207/s15327957pspr0202_3

Ramadhan, A., Warnars, H. L. H. S., & Razak, F. H. A. (2024). Combining intelligent tutoring systems and gamification: A systematic literature review. *Education and Information Technologies, 29*(6), 6753–6789. https://doi.org/10.1007/s10639-023-12092-x

Branko Anđić   PhD, Researcher, Center for Higher Education, School of Education, Johannes Kepler University, Altenberger Straße 69, Standort. Science Park 5, 4040 Linz, Austria.
E-mail: brankoan01@gmail.com
ORCID: https://orcid.org/0000-0003-2691-8357