

Call for Paper: 1. Tagung der AG Mediendidaktik

Pädagogische und inhaltliche Kompetenzen (PCK) einer Künstlichen Intelligenz? Erfassung der Bewertungsgenauigkeit von ChatGPT-4 am Beispiel prüfungsrelevanter Lerntagebücher

Gentile et al. (2023) prognostizieren einen „genuine paradigm shift“ im Bezug auf die Rolle der Lehrkraft im Zeitalter Künstlicher Intelligenz (KI). Eine hybride Zusammenarbeit von Lehrkraft und KI ist erwartbar (Molenaar 2021); die Diskrepanz zwischen theoretisch möglichen und tatsächlichen KI-Tätigkeiten in Bildungskontexten ist jedoch relativ groß (Zhang und Aslan 2021).

Die vorliegende quantitative Studie adressiert die tatsächliche Interaktion zwischen Lehrkraft und KI unter der Fragestellung, inwiefern eine KI die pädagogischen als auch inhaltlichen Kompetenzen (PCK) einer Lehrkraft im Bewertungsprozess von prüfungsrelevanten Lerntagebüchern¹ simulieren kann (Schütze et al. 2018; Glogger et al. 2012; Naujoks und Händel 2020; Nückles et al. 2020; Celik 2023; Alers et al. 2024; Jukiewicz 2024). Hierzu überprüft die Lehrkraft aktiv den Bewertungsvorschlag von ChatGPT-4 pro Lerntagebuchkategorie, bestätigt oder verändert diesen. Zur Bestimmung der Bewertungsgenauigkeit von ChatGPT-4 wird innerhalb von sechs Monaten und zu unterschiedlichen Zeitpunkten eine Confusion Matrix 3x3 herangezogen (Bernius et al. 2022). Ein Datensatz von mehr als 500 Interaktionen wird ausgewertet.

Es ist erwartbar, dass sich die Bewertungsgenauigkeit in Abhängigkeit von den zu bewertenden Inhalten (CK) und den zu verwendeten Punktzahlen (PK) unterscheidet; daraus ergeben sich konkrete Handlungsoptionen für Lehrkräfte in einem KI-gestützten Bewertungsprozess.

Literaturverzeichnis:

Alers, H., Malinoska, A., Meghoe, G., Apfel, E. (2024). Using ChatGPT-4 to Grade Open Question Exams. *Advances in Information and Communication, Proceedings of the 2024 Future of Information and Communication Conference (FICC)*, Volume 1, 1–9.

Bernius, J., Krusche, S., Bruegge, B. (2022). Machine learning based feedback on textual student answers in large courses. *Computers and Education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caeai.2022.100081>.

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138. <https://doi.org/10.1016/j.chb.2022.107468>.

Gentile, M., Città, G., Perna, S., Allegra, M. (2023). Do we still need teachers? Navigating the paradigm shift of the teacher's role in the AI era. *Frontiers in Education*, 8. doi:10.3389/feduc.2023.1161777.

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology*, 104 (2), 452–468. doi.org/10.1037/a0026683.

Jukiewicz, M. (2024). The future of grading programming assignments in education: The role of ChatGPT in automating the assessment and feedback process. *Thinking Skills and Creativity*, 52. <https://doi.org/10.1016/j.tsc.2024.101522>.

¹ Lerntagebücher werden als formatives und Lernstrategien integrierendes Prüfungsformat im Lehrangebot eines Weiterbildungsinstitutes eingesetzt. Diese gliedern sich in vier Kategorien: In der ersten Kategorie benennen, gliedern und erklären die Lernenden die zentralen Inhalte. Anschließend verknüpfen die Lernenden in der zweiten Kategorie die Lerninhalte mit eigenen Erfahrungen und Beispielen. In der dritten Kategorie arbeiten die Lernenden die Relevanz der Lerninhalte für die berufliche Praxis heraus. Zuletzt formulieren sie in der vierten Kategorie Verständnisschwierigkeiten, weiterführende Fragen oder ein Fazit. Mittels Leitfragen werden die Lernenden bei der Erstellung unterstützt (Nückles et al. 2020).

Molenaar, I. (2021). Personalisation of learning: Towards hybrid human-AI learning technologies. *Blockchain, and Robots*, 57–77. doi.org/10.1111/ejed.12527.

Naujoks, N., Händel, M. (2020). Nur vertiefen oder auch wiederholen? Differenzielle Verläufe kognitiver Lernstrategien im Semester. *Unterrichtswissenschaft*, 48, 221–241. doi.org/10.1007/s42010-019-00062-7.

Nückles, M., Roelle, J., Glogger-Frey, I., Waldeyer, J., Renkl, A. (2020). The self-regulation-view in writing-to-learn: Using journal writing to optimize cognitive load in self-regulated learning. *Educational Psychology Review*, 32, 1089–1126. doi.org/10.1007/s10648-020-09541-1.

Schütze, B., Souvignier, E., Hasselhorn, M. (2018). Stichwort – Formatives Assessment. *Erziehungswissenschaft*, 21, 697–715. doi.org/10.1007/s11618-018-0838-7.

Zhang, K., Aslan, A. (2021). AI technologies for education: Recent research & future directions, *Computers and Education: Artificial Intelligence*, Vol. 2. Verfügbar unter <https://www.sciencedirect.com/science/article/pii/S2666920X21000199> [12.04.2024].