SoTL-Netzwerktreffen 2026

Beitragsabstract

Human-AI Agency: Was AI Literacy über Bedienkompetenz hinaus braucht

Volker Göhler, TU Bergakademie Freiberg

AI Literacy wird in vielen Studiengängen als Lehrziel verfolgt und überwiegend über Selbstauskunft erhoben. Bei dreizehn von 16 durch Lintner (2024) geprüften Skalen wird die *Selbsteinschätzung* erfasst, bei drei die *Leistung*. Werden beide gemeinsam erhoben, wird eine geringe Korrelation beobachtet (Zhang et al. 2026). Im Vortrag wird AI Literacy deshalb eingegrenzt auf nachweisbares Wissen darüber, wie generative KI ihre Ausgaben erzeugt, wo ihre Grenzen liegen und wie eine gegebene Ausgabe geprüft wird. *Sicherheit in der Bedienung* und Zutrauen in die eigene Fähigkeit werden ausgeschlossen.

In einer Lehrveranstaltung zur Softwareentwicklung wurden vier KI-Agenten ein Semester lang als Teammitglieder in die GitHub-Repositorien der Studierenden eingebunden. Vom Junior-Entwickler-Agenten wird nur umgesetzt, was Studierende selbst spezifiziert haben. Sein Code muss anschließend von ihnen geprüft werden. Die Anfragen des Berater-Agenten sind bewusst vage formuliert und müssen priorisiert, präzisiert oder zurückgewiesen werden. Verlangt wird damit weniger Wissen über KI als die Fähigkeit, eigene Ziele gegenüber den Agenten zu setzen und zu halten.

Für diese Fähigkeit wird der Begriff *human-AI agency* vorgeschlagen (Göhler, in Vorbereitung), definiert als die Handlungsfähigkeit, die in der Interaktion mit generativer KI tatsächlich ausgeübt wird. Unterschieden werden Zielhoheit, Widerstand gegen Zustimmung und Überredung durch das System (Sharma et al. 2024; Salvi et al. 2025), die korrekte Zuschreibung von Urheberschaft und die zutreffende Einschätzung der eigenen Handlungsfähigkeit. Human-AI agency kann nur im Arbeitsprozess beobachtet werden, nicht im Ergebnis.

Literatur

Göhler, V. (in Vorbereitung). *Human-AI Agency as a Catalyst for Innovative Teaching in Higher Education*. Dissertation, TU Bergakademie Freiberg.

Lintner, T. (2024). A systematic review of AI literacy scales. *npj Science of Learning*, 9(1). https://doi.org/10.1038/s41539-024-00264-4

Salvi, F., Horta Ribeiro, M., Gallotti, R. & West, R. (2025). On the conversational persuasiveness of GPT-4. *Nature Human Behaviour*, 9(8), 1645–1653. https://doi.org/10.1038/s41562-025-02194-6

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R. et al. (2024). Towards Understanding Sycophancy in Language Models. In *Proceedings of the Twelfth International Conference on Learning Representations (ICLR 2024)*. https://doi.org/10.48550/arXiv.2310.13548

Zhang, S., Xiao, R., Botelho, A. F., Liao, G., Chiu, T. K. F., Stamper, J. & Koedinger, K. R. (2026). How to Assess AI Literacy: Misalignment Between Self-Reported and Objective-Based Measures. In *Proceedings of the 16th International Learning Analytics and Knowledge Conference (LAK 2026)* (S. 405–414). ACM. https://doi.org/10.1145/3785022.3785088