

Hauptseminar „Philosophische Aspekte des Digitalen und der Digitalisierung“

TU Dresden | Institut für Philosophie | PD Dr. R. Hauswald | rico.hauswald@tu-dresden.de | SS 2023 | Di (5) 14.50-16.20

Seminarplan

4.4. **Einführung**

11.4. **Die Epistemologie der Wikipedia**

Textgrundlage: Fallis + Wray

18.4. **Die Epistemologie des Googelns**

Textgrundlage: Simpson

25.4. **Das Internet als Herausforderung und Chance für Journalismus und demokratischen Diskurs**

Textgrundlage: Goldman + Coady

2.5. **Algorithmen und die Zukunft der Demokratie**

Textgrundlage: Danaher

9.5. **Algorithmische Diskriminierung, Bias und Ungerechtigkeit**

Textgrundlage: Peters

16.5. **Deep Fakes**

Textgrundlage: Habgood-Coote

23.5. **Echokammern, moralische Empörung und die „Gamifizierung“ der Kommunikation am Beispiel von Twitter**

Textgrundlage: Nguyen

30.5. **Pfingsten**

6.6. **Grundprinzipien und blinde Flecken der KI-Ethik**

Textgrundlage: Hagendorff

13.6. **KI, Verantwortlichkeit und Erklärbarkeit**

Textgrundlage: Baum et al.

20.6. **Ausfall**

27.6. **Philosophische Aspekte von LLMs (ChatGPT usw.)**

Textgrundlage: Floridi + Chalmers [als Ergänzung: Schwitzgebel/Strasser]

4.7. **Ausfall**

11.7. **Abschlussdiskussion / Ausweichtermin**

Texte:

Baum, K./Mantel, S./Schmidt, E./Speith, T. (2022): From Responsibility to Reason-Giving Explainable Artificial Intelligence. In: *Philosophy & Technology* 35:12, S. 1-30.

Chalmers, D. (2023): Could a Large Language Model be Conscious?
<https://philarchive.org/rec/CHACAL-3>.

Coady, D. (2011): An Epistemic Defence of the Blogosphere. In: *Journal of Applied Philosophy* 28 (3), S. 277-294.

Danaher, J. (2016): The Threat of Algocracy: Reality, Resistance and Accommodation. In: *Philosophy & Technology* 29, S. 245-268.

Fallis, D. (2008): Toward an Epistemology of *Wikipedia*. In: *Journal of the Association for Information Science and Technology* 59 (10), S. 1662-1674.

Floridi, L. (2023): AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models. In: *Philosophy & Technology* 36:15, S. 1-7.

Goldman, A. (2008): The social epistemology of blogging. In: M. J. van den Joven & J. Weckert (Hg.): *Information Technology and Moral Philosophy*. Cambridge: Cambridge University Press, S. 111-122.

Habgood-Coote, J. (2023): Deepfakes and the epistemic apocalypse. In: *Synthese* 201:103, S. 1-23.

Hagendorff, T. (2022): Blind spots in AI ethics. In: *AI Ethics* 2, S. 851-867.

Nguyen, C. T. (2021): How Twitter gamifies communication. In: J. Lackey (Hg.): *Applied Epistemology*. Oxford: Oxford University Press, S. 410-436.

Peters, U. (2022): Algorithmic Political Bias in Artificial Intelligence Systems. In: *Philosophy & Technology* 35:25, S. 1-23.

Schwitzgebel, E./Schwitzgebel, D./Strasser, A. (2023): Creating a Large Language Model of a Philosopher. <https://arxiv.org/ftp/arxiv/papers/2302/2302.01339.pdf>.

Simpson, T. W. (2012): Evaluating Google as an Epistemic Tool. In: *Metaphilosophy* 43 (4), S. 426-445.

Wray, K. (2009): The Epistemic Cultures of Science and *Wikipedia*: A Comparison. In: *Episteme* 6 (1), S. 38-51.