Keynote:
N. Katherine Hayles

Speakers:
Deanna Cachoian-Schanz
Alex Campolo
Christoph Engemann
Moritz Hiller
Rainer Mühlhoff
Florian Sprenger
Katia Schwerzmann

New Reading Scenes: On Machine Reading and Reading Machine Learning Research

27.-28.2.2025

Hybrid Workshop

Conception and organization by **Katia Schwerzmann**

Workshop of the Institute for Advanced Study in the Humanities (KWI) Essen
In cooperation with the SFB Virtuelle Lebenswelten, Ruhr-Universität Bochum

INSTITUTE FOR ADVANCED STUDY IN THE HUMANITIES
UNIVERSITY ALLIANCE RUHR

Sonderforschungsbereich
VIRTUELLE Lebenswelten

Live: Collaboration Space GB 8
129 Ruhr-Universität Bochum
Virtual: Zoom
Registration Link

Schedule

Thursday, February 27

13:00 – 13:30 **Katia Schwerzmann (KWI Essen):** *Opening Remarks*

13:30 – 14:30 **Moritz Hiller (Bauhaus Universität Weimar):** *The Subject of Reading: On Machine Philology*

14:30 – 15:30 **Florian Sprenger (Ruhr Universität Bochum):** *Machines Reading. Chains of Translation Between Worlds and Computers*

15:30 – 16:00 **Pause**

16:00 – 17:15 **Deanna Cachoian-Schanz (University of Pennsylvania, US), Katia Schwerzmann:** *On Human and Machine Translation, Normative Gaps, and Trans Possibilities*

18:00 – 19:30 **Keynote by N. Katherine Hayles (Duke University, US):** TBA

Friday, February 28

11:00 – 12:00 **Christoph Engemann (Ruhr Universität Bochum):** *Machine Reading & Human Speaking: Platform-Orality*

12:00 – 13:00 **Pause**

13:00 – 14:00 **Alex Campolo (Durham University, UK):** *What Type of Scientific Literature Does the Machine Learning Community Generate? And How Can It Be Read Critically?*

14:00 – 15:00 **Rainer Mühlhoff (Universität Osnabrück):** *Reading as Assessing: Automated Homework Grading With LLMs*

15:00 – 15:30 **Katia Schwerzmann:** *Closing Remarks*

Zoom Link

https://ruhr-uni-bochum.zoom-x.de/j/67557387410?pwd=EE6FFOcuSo9mtkbfDQDtapPcxqyPoK.1

Description

Reading has undergone dramatic transformations over the past few decades. Media and literary theorist N. Katherine Hayles has discussed how forms of reading, modes of attention, and even neurological architecture are heavily influenced by the medium of reading—on screen vs. on print—and its media-specific features such as layout, typography, and the presence of hyperlinks (Hayles 2012; 2021; 2023). Under "machine reading," Hayles refers to machines' ability to process vast amounts of text and uncover patterns that would be imperceptible to a human reader. Additionally, the ability to search for keywords in digital texts facilitates a form of "distant reading," enabling readers to engage with texts in new ways by adopting abstract, visual, quantifying approaches (Moretti 2013; Jänicke et al. 2015).

Recently, literary scholar Julika Griem has proposed to analyze what she calls "reading scenes," where the practice of reading is explicitly thematized in literary texts and visual media. This media reflexivity enables us to analyze the changing forms, valuations, and norms assigned to reading as a cultural practice (Griem 2021). Griem's approach asks us to attend to the technical, social, and cultural contexts of the practice of reading in addition to its cognitive dimensions. What new reading scenes emerge with large language models (LLMs) and the research practices surrounding them? And what are the consequences for the "University as a service" (Kirschenbaum and Raley 2024)?

The emergence of large language models has transformed modes of reading and introduced new forms of attention and valuations. Large language models not only automate the production of texts across various genres—such as emails, blog posts, essays—but can even generate texts "in the style of" a specific author, as long as their writing is sufficiently present in the training dataset. Even more intriguing, LLMs allow users to generate digests of the key elements of a text and its broader significance. More specialized tools such as NotebookLM are fine-tuned to automate text analysis: they can summarize any text with reference to specific passages, render its structure, generate quizzes, and propose "essay questions." Traditional reading methods such as the "explication de texte" or "close reading" that rely on an inquisitive and cautious analysis of a short passage—a reading that pays attention to the formal and rhetorical dimension of a piece of text—competes with automated tools that establish the relevance of the components of a text through the statistical weighing of its constitutive elements: the tokens or words. This, in turn, is made possible by the model's prior learning of statistical patterns present in the training data.

Machine writing consists of generating an output that predicts the sequence of words that has the statistically highest likelihood to follow the words in the input. Critical questions about reading arise as reading practices become increasingly exteriorized and automatized, and as newer models like GPT-o1 become fine-tuned through the mechanism of "chain-of-thought" (Wei et al. 2023; OpenAI 2024) that simulates "close-reading" by dissecting (that is, literally analyzing) the user's prompt in its simplest elements, enabling the model to tackle each of these successively.

The shift in reading raises a series of critical questions that we will explore by reflection on emerging machine reading scenes. Lines of inquiry may include but are not limited to:

- What forms of reading are automatized through machine processing? What cultural, technical, ethical, and economic valuations are encoded into these machine reading scenes? Do close reading and the reading of longer texts, both of which require sustained attention, lose their status as foundational skills to be learned in educational settings? What are the normative implications behind machines' "interpretation" of what counts in a text and the reduction of texts to containers for information?
- While the output of AI tools carries the "statistical normativity" of the dataset and the "ethicotechnical normativity" of the model's fine-tuning (Schwerzmann and Campolo 2024), this output is also always tailored to the individual user's prompting. Is reading bound to lose the socializing function of modulating access to a public common discourse?
- As researchers process an ever-increasing quantity of texts, new reading strategies, sometimes machine-aided, become necessary. What reading strategies and formats are being developed? For instance, is the format of the more-or-less detailed, automated summary bound to become the dominant approach to texts? What reading method could researchers of the humanities and social sciences develop to closely read the outputs of LLMs (such as the screenshot below) while attending to their material and medial conditions as well as their specific, technological and ethical, forms of normativity?
- Louise Amoore and her team have shown that computer science literature is a scene on which AI gives an account of its "paradigmatic worldview" (Amoore et al. 2023). The reflexive character of computer science literature constitutes a specific, yet still understudied reading scene to understand today's algorithmic rationality. What forms of reading might researchers of the humanities and social sciences develop in order to engage with computer science research, which is often outside their traditional fields of expertise?
- Acknowledging the structuralist distinction between "énoncé" and "énonciation" (Benveniste 1971), who speaks in texts generated by LLMs (Schwerzmann 2024)? What type of subject position is asserted in machine learning generated texts? And what kind of subjectivity do these texts constitute in their address to a reader? Finally, how can these questions help us to critically engage with LLMs' forms of knowledge (re)production?
- How does machine translation as the foundational problem of natural language processing at the inception of today's large language models (Cho, Courville, and Bengio 2015) differ from translation understood as a cultural-linguistic practice? Under the umbrella of "sequence modeling," machine translation consists in "aligning" a sequence of words in the source language to a sequence of words in the target language, based on a parallelization of the word sequences. Machine translation is made possible through the probabilistic, asemantic quantification of languages and the presupposition of their equivalence. By contrast, human translation as a reading-writing practice is characterized by its accounting for specific temporal, geographical, and affective contexts that are not reducible to the norm of the "highest likelihood" but are instead highly dependent on the embodied experience of the translator (Tymoczko 2002; Cachoian-Schanz 2021). How can translation help us to think both about human and machine reading?

References

Amoore, Louise, Alexander Campolo, Benjamin Jacobsen, and Ludovico Rella. 2023. “Machine Learning, Meaning Making: On Reading Computer Science Texts.” Big Data & Society, 1–13.

Benveniste, Emile. 1971. Problems in General Linguistics. Coral Gables, FL: University of Miami Press.

Cho, Kyunghyun, Aaron Courville, and Yoshua Bengio. 2015. “Describing Multimedia Content Using Attention-Based Encoder-Decoder Networks.” IEEE Transactions on Multimedia 17 (11): 1875–86. https://doi.org/10.1109/TMM.2015.2477044.

Griem, Julika. 2021. Szenen des Lesens: Schauplätze einer gesellschaftlichen Selbstverständigung. Wie wir lesen - Zur Geschichte, Praxis und Zukunft einer Kulturtechnik, Band 3. Bielefeld: transcript.

Hayles, N. Katherine. 2012. How We Think: Digital Media and Contemporary Technogenesis. Chicago: The University of Chicago Press.
———. 2021. Postprint: Books and Becoming Computational. New York, NY: Columbia University Press.
———. 2023. “Subversion of the Human Aura: A Crisis in Representation.” American Literature 95 (2): 255–79. https://doi.org/10.1215/00029831-10575063.

Jänicke, S, G Franzini, M F Cheema, and G Scheuermann. 2015. “On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges.” The Eurographics Association, 1–21.

Kirschenbaum, Matthew, and Rita Raley. 2024. “AI and the University as a Service.” PMLA/Publications of the Modern Language Association of America 139 (3): 504–15. https://doi.org/10.1632/S003081292400052X.

Moretti, Franco. 2013. Distant Reading. London: Verso.

OpenAI. 2024. “OpenAI O1 System Card.” https://cdn.openai.com/o1-system-card.pdf.

Schwerzmann, Katia. 2024. “From Enclosure to Foreclosure and Beyond: Opening AI’s Totalizing Logic.” Under Review at AI & Society, 1–20.

Schwerzmann, Katia, and Alexander Campolo. 2025. “‘Desired Behaviors’: Alignment and the Emergence of a Machine Learning Ethics.”

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. “Attention Is All You Need.” arXiv. http://arxiv.org/abs/1706.03762.

Wei, Jason, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models.” arXiv. http://arxiv.org/abs/2201.11903.