

Céline Budding

[Homepage](#) | [Google Scholar](#) | [LinkedIn](#)

Location: Rotterdam, The Netherlands

Email: c.e.budding@tue.nl

PHD CANDIDATE PHILOSOPHY OF AI

I am a fourth-year PhD candidate in Philosophy of AI. In my project, I investigate what large language models like ChatGPT "know" about language. Specifically, I take inspiration from philosophical debates in AI in the 80s and 90s and argue that large language models can be considered to have tacit knowledge of language.

More generally, I am interested in what AI systems learn from the data, how we can evaluate and explain their performance, and how we can translate insights from various academic fields—machine learning, cognitive science, and philosophy—to effective regulation and policy for AI.

EXPERIENCE

PhD Candidate

Philosophy & Ethics group, Eindhoven University of Technology

Eindhoven, The Netherlands

Aug 2021 - current

- Project: What do large language models know? Tacit knowledge as a potential causal-explanatory structure, supervised by dr. C.A. Zednik and prof. dr. V.C. Müller.

EDUCATION

Master of Science in Computational Neuroscience

Bernstein Center for Computational Neuroscience Berlin

Berlin, Germany

Oct 2018 - May 2021

- Thesis: Evaluating interpretability methods on structural brain MRI data with synthetic lesions, supervised by dr.rer.nat. S. Haufe and prof.dr. K. Ritter.

Bachelor of Science in Liberal Arts & Sciences

University College Roosevelt/Utrecht University

Middelburg, The Netherlands

Sep 2015 - June 2018

- Thesis: Classification of MRI scans from the ADNI database using voxel-based morphometry, principal component analysis, and support vector machine, supervised by dr. R. van den Doel and dr. G. Andringa.

SELECTED CONFERENCE & WORKSHOP PRESENTATIONS

- Assessing the performance of LLMs: the problem of emergent abilities. Poster presentation at the 1st Workshop on Standardizing Tasks, meAsures and NLP Datasets. Jan 29, 2024, Paris, France.
- What do large language models know? Tacit knowledge as a potential causal-explanatory structure. Presentation at PhAI 2023, Dec 15-16, 2023, Erlangen, Germany.
- Toward Cognitive Models in Explainable AI: Mechanistic Interpretability as a Missing Link?. Presentation together with dr. C.A. Zednik at the Biannual Conference of the European Society for Philosophy of Science. September 20-23 2023, Belgrade, Serbia.
- Tacit knowledge and how-explanations for large language models. Presentation at the 49th meeting of the Society for Philosophy of Psychology. June 20-23 2023, Pittsburgh, US.
- Tacit Knowledge and 'How'-Explanations for Large Language Models. Presentation at the Graduate Workshop of PhilML22. October 19 2022, Tübingen, Germany.

PUBLICATIONS

- Wilming, R., Budding, C., Müller, K. R., & Haufe, S. (2022). Scrutinizing XAI using linear ground-truth data with suppressor variables. *Machine learning*, 111(5), 1903-1923.
- Budding, C., Eitel, F., Ritter, K., & Haufe, S. (2021). Evaluating saliency methods on artificial data with different background types. *arXiv preprint arXiv:2112.04882*.

TEACHING EXPERIENCE

Teaching assistant

Eindhoven University of Technology

2021-current

- Q2 2023-2024 & Q2 2022-2023, Philosophy and ethics of AI (MSc level).
- Q4 2021-2022, USE-BASE (introductory ethics, first-year BSc students).

ACADEMIC ACTIVITIES

Participant Summer School on the Law, Ethics, and Policy of AI

KU Leuven

July 1-10 2024

- A ten-day intensive summer school on the law, ethics, and policy of AI, specifically focusing on developments surrounding the new AI Act in the EU.

Participant Summer School Science Communication

Leiden University

July 3-7 2023

- One-week science communication summer school consisting of lectures and a group project, in which we developed a magazine about visual science communication.

Board member WISE network

Eindhoven University of Technology

May 2023-current

- PhD representative in the Women in Science & Engineering (WISE) board.

Organizer workshop on the Philosophy of Large Language Models

Hybrid/Eindhoven University of Technology

Jan 26 + 27 2023

- Two-day hybrid international workshop about communicating with chatbots and explaining large language models.

Representative PhD Council

4TU.Ethics Graduate School

Jan-Dec 2022

- Organized social events and a writing retreat and attended management team meetings of the 4TU.Ethics network.

NON-ACADEMIC ACTIVITIES

Committee member 'AI & big data'

Stichting Koninklijk Nederlands Normalisatie Instituut (NEN)

March 2023-current

- Member of the Dutch working group on AI & big data, as well as Joint Working Group 5 (ISO and CEN/CENELEC) on natural language processing methods.

Studentdocent/Student Teacher

Stichting Oud Geleerd Jong Gedaan

Jan 2021 - current

- Developing and teaching low-threshold and interactive lectures about Artificial Intelligence for an audience of elderly people (5-100+ participants).