

Tamara Czinczoll

Berlin, Germany

+49 1775609295 | t.czinczoll@gmail.com | t-czinczoll | Tamara Czinczoll

EDUCATION

Hasso Plattner Institute / University of Potsdam

Potsdam, Germany

PhD Computer Science

Nov. 2021 – Present

- Language Modeling with Latent Text Representations
- Supervised by Gerard de Melo and Ralf Herbrich

Stanford University

Stanford, USA

Research Stay

Feb. 2026 – May 2026

- Working on human communication-inspired LLM training objectives
- Visiting Professor Monica Lam's group

University of Amsterdam

Amsterdam, Netherlands

MSc Artificial Intelligence, cum Laude

Sep. 2019 – Aug. 2021

- “All the World's a Stage”: Can Neural Approaches Understand Analogies in NLP?
- Supervised by Ekaterina Shutova, Helen Yannakoudakis and Pushkar Mishra

Boğaziçi University

Istanbul, Turkey

Computer Science (Erasmus Semester Abroad)

Aug. 2017 – Dec. 2017

University of Osnabrück

Osnabrück, Germany

BSc Cognitive Science

Oct. 2015 – March 2019

EXPERIENCE

Applied Scientist Intern

May 2025 – Present

Amazon EU Search

Luxembourg, Luxembourg

- Developed an inference-compute scaling framework for iteratively refining rankings with LLMs
- Designed a taxonomy for difficult search query types
- Actively engaged with the Amazon Search research teams by providing feedback to ongoing research projects, reviewing paper submissions and joining the paper reading group

AI Trainer

Nov. 2021 – May 2025

Mittelstand-Digitalzentrum Berlin

Berlin, Germany

- Developed AI demos to showcase to German companies, e.g., multimodal search, an LLM chatbot and a diffusion-based interior designer
- Designed and taught AI workshops on business use cases to SMEs in Germany
- Set up and taught courses for people without a computer science background on classical machine learning and modern neural models

Teaching Assistant

Oct. 2020 – Jan. 2021

University of Amsterdam

Amsterdam, Netherlands

- *Natural Language Processing 1*: Helped students with assignments, questions about the lecture and exam preparation

NLP Developer Intern

April 2019 – June 2019

Psychological AI

Heilbronn, Germany

- Explored and evaluated methods for lexical substitution and context-aware phrase suggestions

NLP Research Intern

July 2018 – January 2019

Fraunhofer IIS

Erlangen, Germany

- Engaged in collaborative research with the *Natural Language User Interface* group and developed an embedding-based approach for the topic of *Named Entity Recognition with external knowledge sources*.
- Compared the performance of different embedding methods

SELECTED PUBLICATIONS

NextLevelBERT: Masked Language Modeling with Higher-Level Representations for Long Documents | *long-context, masked language modeling, compression* Published at ACL 2024

- Goal: obtain long-document embeddings
- Method: Chunk input documents → embed the chunks → pretrain model with masked language modeling on sequence of chunk embeddings → obtain contextualized chunk embeddings → mean-pool into document embedding
- Results: beat NomicEmbed 8k context length model on a set of long-document tasks
- Advantages: Enables contextualized embeddings for sequences of up to 260k tokens with 130MB model size

Scientific and Creative Analogies in Pretrained Language Models

analogical reasoning, GPT-2, BERT, dataset

Published at EMNLP 2022

- Goal: Evaluate how well generative LMs can understand complex analogies
- Method: Collected dataset of conceptual analogies with multi-attribute mappings from scientific and creative domains
- Result: the LMs at the time failed to understand more complex analogies; good performance on the established BATS word analogy dataset does not translate to good performance on more human-like analogies

TECHNICAL SKILLS

Machine Learning: PyTorch, Lightning, scikit-learn, tensorflow, W&B, HuggingFace, numpy, pandas

Programming Languages: Python, Java, Matlab, Prolog/ASP

Developer Tools: Git, Docker, Linux, SLURM, VS Code, Cursor, PyCharm, LaTeX