Yana Veitsman

🌐 Website   **G** Google Scholar   ○ Github   **in** LinkedIn   ✉ yanav@lst.uni-saarland.de

Education

Universität Göttingen   Dec 2025 — now
PhD in Computer Science   Göttingen, Germany
- Topic: Cognitively Plausible Strategies for Sample-Efficient Modeling

Universität des Saarlandes   Oct 2023 — Oct 2025
MSc Language Science and Technology, CGPA: 1.4/1.0 (German scale)   Saarbrücken, Germany
- Master Thesis: *Why Cross-Lingual Alignment Fails for Better Cross-Lingual Transfer?*
- Organizer of LangLunches — informal inter-department research talks
- Student Council President

American University in Bulgaria   Jan 2019 — Dec 2022
BA Computer Science (Summa Cum Laude), CGPA: 3.93/4.0   Blagoevgrad, Bulgaria
- Bachelor Thesis: *NLP-based Social Media Post Generator*
- Major: Computer Science
- Minors: Mathematics, Modern Languages and Cultures (Spanish language focus)

Universitat de València   Feb 2021 — Jun 2021
Erasmus+ exchange, Department of Computer Engineering   València, Spain
- 1st class grade in the course Data Mining and Machine Learning (out of 30 students)

Research Experience

Doctoral Researcher, Universität Göttingen   Dec 2025 — now
Supervisor: Prof. Dr. Lisa Beinborn   Göttingen, Germany
- Working on sample-efficient modeling approaches for multilingual NLP inspired by human cognition

Research Assistant, Universität des Saarlandes   Feb 2024 — Aug 2025
Supervisors: Prof. Michael Hahn, Prof. Vera Demberg   Saarbrücken, Germany
- Co-led a project on exploring persistent theoretical limitations in pre-trained LLMs
- Co-authored a paper on the mechanisms behind in-context learning in Gemma model family
- Co-authored a paper on expressivity of state space models (SSMs) using formal languages framework

Research Assistant, German Research Center for Artificial Intelligence (DFKI)   Feb 2024 — Oct 2024
Supervisors: Dr. Simon Ostermann, Dr. Cennet Oguz   Saarbrücken, Germany
- Streamlined annotation, validation, and pre-processing of data for multimodal anaphora resolution

Industry Experience

Software Developer   Dec 2021 — Dec 2022
VMware Bulgaria   Sofia, Bulgaria
- Integrated an LLM-based search system with a documentation portal using RabbitMQ, Python, and .NET
- Developed a graphic design requests portal in .NET and Angular, used by 90+ technical writers

Junior Backend Developer   Jun 2021 — Dec 2021
Social Paragraph, schip.gent startup   Brussels, Belgium
- Developed a Python app for generating tweets from web-scraped content under tight computational limits

Honors, Awards, and Scholarships

Professional Development Grant   Dec 2024
American Councils for International Education
- Awarded a grant for conference travel among 8000 FLEX program alumni

Best Group Project Award, 3rd place (out of 14)   Jul 2024
Universität des Saarlandes, Data Science course
- Led the project titled *Predicting Processing Times in Resource-Constrained Manufacturing Settings with*

Transformer-Based and Regression Analysis Models

Outstanding Achievement in Computer Science   Apr 2023
American University in Bulgaria

- Awarded in the Foundations of Computing concentration among the graduating class of 200 students

Outstanding Achievement in Modern Languages and Cultures   Apr 2023
American University in Bulgaria

- Awarded in Spanish language studies among the graduating class of 200 students

Conferences and Schools

The 31st International Conference on Computational Linguistics   Jan 2025
The 1st Workshop on Language Models for Low-Resource Languages   Abu Dhabi, UAE
Poster Presenter, Session Chair

                                                                      Dec 2024
The 38th Annual Conference on Neural Information Processing Systems   Vancouver, Canada
Poster Presenter

                                                                      Sep 2024
The 2nd Athens Natural Language Processing Summer School   Athens, Greece
Participant

                                                                      May 2024
The 33rd Computational Linguistics Student Conference   Saarbrücken, Germany
Speaker

Talks and Workshops

LLM Prompting Strategies for Everyone   May 2025
IT-Hub Terricon Valley   Karaganda, Kazakhstan

                                                                      Mar 2025
Hands-On AI for Students   Remote
FLEX Alumni Community Grant Project

                                                                      May 2024
Transformers or RNNs or SSMs: Who's More Sensitive?   Saarbrücken, Germany
TaCoS, The 33rd Computational Linguistics Student Conference

                                                                      May 2023
Introduction to Natural Language Processing   Karaganda, Kazakhstan
IT-Hub Terricon Valley

                                                                      Apr 2022
Comparing Modeling of Russian and English Languages   Blagoevgrad, Bulgaria
from the Point of View of NLP Potential
The 17th Fellowship of the Mind Student Conference

Teaching and Mentorship

ABLE Mentor Program Mentor   Sep 2022 — Dec 2022
American University in Bulgaria

- Advised a 10-th grade local high school student in writing and publishing a book on car tuning

C++ Subject Tutor   Sep 2020 — Dec 2020
American University in Bulgaria

- Conducted tutoring sessions for 20+ students, helping them grasp the concepts of OOP in C++

Skills

Spoken Languages: Russian (Native), English (C2), Spanish (C1), German, Bulgarian, Kazakh (A2)
Programming Languages: Python, C# (.NET), C++, SQL
Machine and Deep Learning: numpy, pandas, PyTorch, HuggingFace transformers, TransformerLens
Tools: Git, Bash, Linux, Docker, AWS/Azure, Weights & Biases, Hydra

Coursework Highlights

Linguistics: Computational Linguistics, Speech Science, Semantic Theory
Mathematics: Real Analysis, Automata, Formal Languages and Applications, Statistics with R
Machine and Deep Learning: Data Mining and Machine Learning, Elements of Machine Learning, Neural Networks: Implementation and Application
Research Seminars: Aligning Language Models with Human Preferences: Methods and Challenges, NLP for Low-Resource Languages, Robust and Efficient NLP