Anej Svete

anej.svete@protonmail.com | LinkedIn | GitHub | Google Scholar | Personal Website

Education

ETH Zürich, ETH AI Center, *PhD in Computer Science* *2023–Present*
- **Research interests**: Development and analysis of neural architectures for language modeling

University of Washington, *Research Visit* *April 2026–August 2026*
- Research visit during my PhD, working with Prof. Noah Smith

ETH Zürich, *Master of Data Science* *2020–2023*
- **GPA:** $5.84/6$; highest in the Master's program cohort. **Willi Studer Prize** for the highest GPA.
- Relevant Courses: Deep Learning, Advanced Machine Learning, Optimization for Data Science, Natural Language Processing, Fundamentals of Mathematical Statistics, Statistical Modeling, Causality

University of Ljubljana, *Bachelor of Computer Science and Mathematics* *2017–2020*
- **GPA:** $10/10$; top student in the department each year; **Rector's** and **Dean's** Prizes for exceptional study results.
- Completed two Stanford courses: Machine Learning with Graphs and Mining Massive Datasets.

Work and Teaching Experience

Meta FAIR, Research Intern, *Menlo Park, California* *Nov 2026–Present*
- Designing new language model architectures in the FAIR Alignment team.

Allen Institute for AI (Ai2), Research Assistant, *Seattle, Washington* *2025–2026*
- Led research on the efficiency and reasoning abilities of diffusion language models; **oral** at ICLR 2026 (**top 1.12 %**).
- Core contributor to **OLMo Hybrid**, Ai2's open hybrid language model, working on architecture design and scaling laws.

ETH Zürich, Head Teaching Assistant, *Zürich, Switzerland* *2021–Present*
- Taught **Large Language Models**, **Natural Language Processing**, **Advanced Formal Language Theory**
- Designed from scratch, reaching 650+ students (LLMs); managed materials, logistics, and a team of 30+ teaching assistants.

XLAB d.o.o., Software Developer, *Ljubljana, Slovenia* *2017–2020*
- Android application software developer.

Scholarships and Awards

Named among 22 **notable alumni** of the first 30 years of the CS department at the University of Ljubljana. 2026

ETH AI Center Scholarship, a graduate scholarship for exceptional PhD candidates in AI at ETH Zürich. 2023–Present

Willi Studer Prize for the highest GPA in the study program at the end of my master's studies. 2023

Ad Futura Scholarship, awarded to excellent students studying abroad by the Slovenian government. 2020–2022

Rector's and **Dean's** Prizes for exceptional study results at the end of my bachelor's studies. 2018–2021

Intern Scholarship by XLAB during my bachelor's studies. 2017–2020

Golden Owl Award by the President of Slovenia for a perfect score on the national matriculation exam (one of 16). 2017

Selected Talks

Expressivity-Maxxing: What To Do When Your Transformer Is Giving Mid, EPFL AI Center, Lausanne 2026

Diffusion Language Models: Problem Solving and Reasoning, AI@JSI Seminar, Jožef Stefan Institute, Ljubljana 2026

How Diffusion Language Models Reason, EurIPS 2025 Workshop on Principles of Generative Modeling 2025

Transformers Can Represent n-gram Language Models, FLaNN Seminar (online) 2024

Analysing Language Models with Formal Language Theory, University of Ljubljana 2024

Lower Bounds on the Expressivity of Recurrent Neural Language Models, University of Melbourne (online) 2024

Service

Organizer of the Seminar on Formal Languages and Neural Networks (FLaNN)	2025–Present
ICML Tutorial on Language Models and Formal Language Theory (Organizer)	2025
ACL Tutorial on Language Models and Formal Language Theory (Leader)	2024
ESSLLI Course on Language Models and Formal Language Theory (Lecturer)	2023
NLP Tutorial at the International Committee of the Red Cross (Organizer)	2023
Reviewer: ICLR, ICML, CoLM, ACL, EMNLP, NAACL	2023–Present

Publications

Context-free recognition with transformers (ICML 2026).
Selim Jerad, **Anej Svete**, Sophie Hao, Ryan Cotterell, and William Merrill.

Disentangling the expressivity of RoPE (CoLM 2026).
Selim Jerad, **Anej Svete**, Jiaoda Li, and Ryan Cotterell.

A framework for understanding learnability in transformers (ICML 2026).
Blanka Kövér, Alexandra Butoi, **Anej Svete**, Michael Hahn, and Ryan Cotterell.

Olmo hybrid: From theory to practice and back (CoLM 2026).
William Merrill, Yanhong Li, Tyler Romero, **Anej Svete**, Caia Costello, Pradeep Dasigi, Dirk Groeneveld, David Heineman, Bailey Kuehl, Nathan Lambert, Chuan Li, Kyle Lo, Saumya Malik, DJ Matusz, Benjamin Minixhofer, Jacob Morrison, Luca Soldaini, Finbarr Timbers, Pete Walsh, Noah A. Smith, Hannaneh Hajishirzi, and Ashish Sabharwal.

Causally evaluating the learnability of formal language tasks (ICML 2026).
Vésteinn Snæbjarnarson, **Anej Svete**, Josef Valvoda, Reda Boumasmoud, Brian DuSell, and Ryan Cotterell.

Revisiting padded transformer expressivity: Which architectural choices matter and which don't (ICML 2026).
Anej Svete, William Merrill, Ryan Cotterell, and Ashish Sabharwal.

On the reasoning abilities of masked diffusion language models (ICLR 2026 (**Oral**)).
Anej Svete and Ashish Sabharwal.

Probability distributions computed by autoregressive transformers (ICLR 2026).
Andy Yang, **Anej Svete**, Jiaoda Li, Anthony Widjaja Lin, Jonathan Rawski, Ryan Cotterell, and David Chiang.

The transformer cookbook (TMLR 2026).
Andy Yang, Christopher Watson, Anton Xue, Satwik Bhattamishra, Jose Llarena, William Merrill, Emile Dos Santos Ferreira, **Anej Svete**, and David Chiang.

Training neural networks as recognizers of formal languages (ICLR 2025).
Alexandra Butoi, Ghazal Khalighinejad, **Anej Svete**, Josef Valvoda, Ryan Cotterell, and Brian DuSell.

Unique hard attention: A tale of two sides (ACL 2025 (**Oral**)).
Selim Jerad, **Anej Svete**, Jiaoda Li, and Ryan Cotterell.

Gumbel counterfactual generation from language models (ICLR 2025).
Shauli Ravfogel[*], **Anej Svete**[*], Vésteinn Snæbjarnarson, and Ryan Cotterell.

Information locality as an inductive bias for neural language models (ACL 2025 (**Oral**; **Selected as one of the five papers for a panel discussion on computational linguistics and machine learning**)).
Taiga Someya, **Anej Svete**, Brian DuSell, Timothy J. O'Donnell, Mario Giulianelli, and Ryan Cotterell.

What languages are easy to language-model? A perspective from learning probabilistic regular languages (ACL 2024).
Nadav Borenstein, **Anej Svete**, Robin Chan, Josef Valvoda, Franz Nowak, Isabelle Augenstein, Eleanor Chodroff, and Ryan Cotterell.

Computational expressivity of neural language models (ACL 2024).
Alexandra Butoi, Robin SM Chan, Ryan Cotterell, William Merrill, Franz Nowak, Clemente Pasti, Lena Strobl, and **Anej Svete**.

On affine homotopy between language encoders (NeurIPS 2024).
Robin SM Chan, Reda Boumasmoud, **Anej Svete**, Yuxin Ren, Qipeng Guo, Zhijing Jin, Shauli Ravfogel, Mrinmaya Sachan, Bernhard Schölkopf, Mennatallah El-Assady, and Ryan Cotterell.

On the representational capacity of neural language models with chain-of-thought reasoning (ACL 2024).
Franz Nowak, **Anej Svete**, Alexandra Butoi, and Ryan Cotterell.

An L* algorithm for deterministic weighted regular languages (EMNLP 2024).
Clemente Pasti, Talu Karagöz, Franz Nowak, **Anej Svete**, Reda Boumasmoud, and Ryan Cotterell.

A probability–quality trade-off in aligned language models and its relation to sampling adaptors (EMNLP 2024).
Naaman Tan, Josef Valvoda, Tianyu Liu, **Anej Svete**, Yanxia Qin, Min-Yen Kan, and Ryan Cotterell.

Can transformers learn n-gram language models? (EMNLP 2024).
Anej Svete, Nadav Borenstein, Mike Zhou, Isabelle Augenstein, and Ryan Cotterell.

On efficiently representing regular languages as RNNs (ACL 2024).
Anej Svete, Robin Chan, and Ryan Cotterell.

Transformers can represent n-gram language models (NAACL 2024).
Anej Svete and Ryan Cotterell.

Lower bounds on the expressivity of recurrent neural language models (NAACL 2024).
Anej Svete, Franz Nowak, Anisha Sahabdeen, and Ryan Cotterell.

Formal aspects of language modeling.
Ryan Cotterell, **Anej Svete**, Clara Meister, Tianyu Liu, and Li Du.

A geometric notion of causal probing (arXiv preprint 2023).
Clément Guerner, **Anej Svete**, Tianyu Liu, Alexander Warstadt, and Ryan Cotterell.

On the representational capacity of recurrent neural language models (EMNLP 2023).
Franz Nowak, **Anej Svete**, Li Du, and Ryan Cotterell.

Efficiently representing finite-state automata with recurrent neural networks (arXiv preprint 2023).
Anej Svete and Ryan Cotterell.

Recurrent neural language models as probabilistic finite-state automata (EMNLP 2023).
Anej Svete and Ryan Cotterell.

Algorithms for acyclic weighted finite-state automata with failure arcs (EMNLP 2022).
Anej Svete, Benjamin Dayan, Ryan Cotterell, Tim Vieira, and Jason Eisner.

Ne gre le za melodijo: Kako Evropa glasuje za svoje najljubše skladbe (Uporabna informatika 2020).
Anej Svete, Jakob Hostnik, and Lovro Šubelj.