

Alexander Hägele

Website: haeggee.github.io
Email: alexanderhagele@gmail.com
LinkedIn: [haeggee](https://www.linkedin.com/in/haeggee)
GitHub: github.com/haeggee
Google Scholar: [Link](#)

Education

EPFL (current)

Ph.D. in Computer and Communication Science

Lausanne, Switzerland

Oct. 2023–Current

- Supervisor: Prof. Martin Jaggi (Machine Learning & Optimization, [MLO](#))
- Focus: Training language models and understanding scaling behavior. This includes optimization algorithms, data, efficiency, as well as adaptive computing in deep learning.
- Core contributor to the [SwissAI](#) Initiative LLM project, working on the entire stack of fully open AI models
- Teaching, supervising, and mentoring students in courses, [hackathons](#), and research projects

ETH Zürich

B.S. and M.S. in Computer Science (with distinction)

Zürich, Switzerland

Sept. 2017–Oct. 2023

- M.S. Thesis: *Sequential Decision Making for Population Dynamics*.
Supervised by C. Bunne, Prof. A. Krause ([LAS Group](#)) & Prof. M. Cuturi ([Apple](#))
- B.S. Thesis: *Certifying Neural Network Robustness using Generative Models*.
Supervised by M. Mirman, Prof. Martin Vechev ([SRI Lab](#)).

École Polytechnique Paris

Graduate Exchange Semester

Palaiseau, France

Sept. 2022–Dec. 2022

University of Toronto

Undergraduate Exchange Semester

Toronto, Canada

Sept. 2019–Dec. 2019

Publications

- [1] **Alexander Hägele**, Elie Bakouch, Atli Kosson, Loubna Ben Allal, Leandro Von Werra, and Martin Jaggi. “Scaling Laws and Compute-Optimal Training Beyond Fixed Training Durations”. In: *Advances in Neural Information Processing Systems* (2024). [Spotlight Paper](#).
Also presented as a Spotlight Presentation at the NGSM Workshop and Best Poster Award at ES-FOMO @ ICML’24. [\[PDF\]](#).
- [2] **Alexander Hägele**, Jonas Rothfuss, Lars Lorch, Vignesh Ram Somnath, Bernhard Schölkopf, and Andreas Krause. “BaCaDI: Bayesian Causal Discovery with Unknown Interventions”. In: *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics* (2023).
[Oral presentation at AISTATS \(32/1689 submissions, top 1.9%\)](#).
Also presented at the 1st Workshop on Causal Representation Learning (CRL) @ UAI’22. [\[PDF\]](#).
- [3] Matthew Mirman, **Alexander Hägele**, Timon Gehr, Pavol Bielik, and Martin Vechev. “Robustness Certification with Generative Models”. In: *Proceedings of the 42nd ACM SIGPLAN Conference on Programming Language Design and Implementation*. [\[PDF\]](#). ACM. 2021.

Research & Work Experience

Apple

Student Researcher in the Machine Learning Research ([MLR](#)) team.

Supervised by Marco Cuturi ([homepage](#)) through a collaboration for the MSc thesis.

Paris, France

Spring 2023

École Polytechnique Fédérale de Lausanne (EPFL)

Part of the [Summer@EPFL](#) program. Competitive application, $\approx 2\%$ acceptance rate.

Supervised by Prof. Thiran and S. Salehkaleybar ([INDY Lab](#)).

Lausanne, Switzerland

Summer 2022

ETH Zürich

Research Project. Resulting Publication: [2]

Supervised by L. Lorch, J. Rothfuss, Prof. Krause ([LAS Group](#)).

Zürich, Switzerland

Spring 2022

ETH Zürich

Bachelor Thesis & Research Project. Resulting Publication: [3]
Supervised by M. Mirman and Prof. Vechev (SRI Lab).

Zürich, Switzerland

2020/2021

Spacemaker AI

Software Engineering Intern. [Interview](#).

Oslo, Norway

Winter 2020

Teaching Experience

EPFL

Lausanne, Switzerland

- Doctoral TA: *Machine Learning (CS-433, Dept. Computer Science)* Fall 2024
- Doctoral TA: *Introduction to Object Oriented Programming (CS-112, Dept. Life Sciences)* Spring 2024

ETH Zürich

Zürich, Switzerland

- Student TA: *Formal methods and Functional Programming (Dept. Computer Science)* Spring 2020

Skills

- **Technical:** Machine Learning, Probabilistic Modeling, Analytical Thinking, Programming Methodology
- **Programming:** Python, C/C++, JavaScript, Go, Haskell
- **Frameworks:** PyTorch, JAX, NumPy, Docker
- **Languages:** *German* (Mother Tongue), *English* (native), *French* (very good)

Professional Services

- **Reviewing:** *Conferences:* ICLR (2025), AISTATS (2023);
Workshops: Efficient Systems for Foundation Models (ES-FOMO) at ICML'24, Frontiers in Learning, Control, and Dynamical Systems (F4LCD) at ICML'23, Neuro Causal and Symbolic AI (nCSI) at NeurIPS'22, Causal Representation Learning (CRL) UAI'22;

Talks

Conferences

- **ICML'24, Workshop on Next Gen. of Seq. Models (Vienna):** Scaling Laws and Compute-Optimal Training Beyond Fixed Training Durations (Spotlight Presentation). 26.07.2024
- **AISTATS'23 (Valencia):** Bayesian Causal Discovery with Unknown Interventions (Oral Presentation). 26.04.2023

Invited

- **Stanford ML Lunch:** Scaling Laws and Compute-Optimal Training Beyond Fixed Training Durations. 04.12.2024
- **LTS4 (Prof. Frossard, EPFL):** ML for Science Beyond i.i.d: Bayesian Causal Discovery and Decision-Making for Population Dynamics. 16.08.2023
- **Causality Discussion Group (CDG, TU Darmstadt, online):** Bayesian Causal Discovery with Unknown Interventions. 27.07.2022

Scholarships and Awards

- **Spotlight** Paper Award, NeurIPS'24 [1]. 2024
- **Best Poster** Award at ES-FOMO and **Spotlight** Presentation at NGSM, ICML'24 Workshops [1]. 2024
- **Distinction** for the Master's degree in Computer Science, ETH Zürich. 2023
- **Distinction** for the Bachelor's degree in Computer Science, ETH Zürich. 2021
- **Book Price** in Theoretical Computer Science, ETH Zürich. 2018
- **Valedictorian**, Abitur Schubart Gymnasium Aalen. **Awards:** Mathematics, Physics, Scheffel, Ferry-Porsche. 2016

Extracurricular Activities

- **Tennis** 2019 –2021
Team Captain, National League C Switzerland, TC Waidberg.
- **Voluntary Social Year** 2016 –2017
Administrative Assistant, Red Cross & Fire Department Aalen.

Last updated: *January 14, 2025*