Yifan Zhu

Doctoral Researcher | Aalto University | ELLIS Institute Finland
yifan.zhu@aalto.fi | yifan-zhu.com | Google Scholar

Driven by the vision of empowering human decision-making with AI, my research aims to align AI agents with users' bounded rationality to support human agency. By developing computational and generative user models, I build AI agents that actively infer users' characteristics at test-time, such as latent goals and cognitive bounds, and thereby provide adaptive assistance, ultimately contributing to more trustworthy, human-centric, and efficient collaborative AI systems.

Education

Aalto University & ELLIS Institute Finland, Finland — Sep 2023 –
Doctoral Researcher, Department of Computer Science
ELLIS PhD Student, ELLIS PhD & Postdoc Program

- Research: Computational Rationality User Modelling for Adaptive AI Assistants.
- Supervisor: Prof. Samuel Kaski; ELLIS co-supervisor: Dr. Mingfei Sun (University of Manchester); advisor: Dr. Sammie Katt.

Aalto University, Finland — Aug 2021 – Jul 2023
Master of Science (Technology) with Honours, Department of Computer Science

- GPA: 4.92/5.0.
- Thesis: Interactive personalization for explainability via human-in-the-loop multi-objective Bayesian optimization.
- Supervisor: Prof. Antti Oulasvirta; advisor: Dr. Suyog Chandramouli.

Nanjing University of Aeronautics and Astronautics, China — Sep 2017 – Jun 2021
Bachelor of Engineering in Computer Science and Technology

- GPA: 4.0/5.0 (average mark: 90/100); rank: 7/137.

University of Edinburgh, U.K. — Sep – Dec 2019
Nominated undergraduate exchange, School of Informatics

- Average mark: 73/100 (First Class).

Peer-reviewed Publications

[1] More Than Irrational: Modeling Belief-Biased Agents
Yifan Zhu, Sammie Katt, Samuel Kaski.
In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2026). [Paper] [Code] [Page]

[2] Interactive Personalization of Classifiers for Explainability using Multi-Objective Bayesian Optimization
Suyog Chandramouli*, **Yifan Zhu***, Antti Oulasvirta.
In Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP 2023). [Paper]

[3] DSDCS: Detection of Safe Driving via Crowd Sensing
Yun Du, Xin Guo, Chenyang Shi, **Yifan Zhu**, Bohan Li.
In Proceedings of the International Conference on Advanced Data Mining and Applications (ADMA 2018). [Paper]

* Equal contribution.

Preprints and Manuscripts

[4] Propose to Learn, Learn to Propose: Evaluability-Aware Assistance under Bounded Rationality
Yifan Zhu, Sammie Katt, Samuel Kaski.
Preprint, 2026; Under review. [Paper]

[5] Model and Predict Time-Inconsistent Agents in Contextual Choice Tasks
Jaeyoung Lee*, **Yifan Zhu***, Thomas Quilter, Samuel Kaski.
Manuscript, 2025; Under review.

[6] Generative-Model Predictive Planning for Navigation in Partially Observable Environments
Thomas Quilter*, **Yifan Zhu***, Guorui Quan*, Mingfei Sun, Samuel Kaski.
Preprint, 2026. [Paper]

* Equal contribution.

Research and Professional Experience

Aalto University, *Computational Behavior Lab* Feb – May 2023
Research Assistant

- Supervisor: Antti Oulasvirta; advisor: Suyog Chandramouli.
- Designed and implemented a personalised text-to-image generation prototype based on stable diffusion with Multi-objective Bayesian optimisation.

Aalto University, *Computational Behavior Lab* Jun – Dec 2022
Master's Thesis Worker

- Supervisor: Antti Oulasvirta; advisor: Suyog Chandramouli.
- Built the interactive personalisation system and conducted the experiments and user study for multi-objective Bayesian optimisation of explainable classifiers.
- Published a UMAP paper. [2]

Accenture, Shanghai, China Oct 2020 – Jan 2021
Consulting SAP Intern

- Developed an integration between a capital platform and SAP using ABAP, and implemented income-statement reporting.
- Developed Core Data Services (CDS) views for SAP Analysis.

Teaching Experience

Teaching Assistant – CS-C3240 **Machine Learning** Autumn 2025
Department of Computer Science, Aalto University

- Instructor: Pekka Marttinen.

Teaching Assistant – CS-C3240 **Machine Learning** Autumn 2024
Department of Computer Science, Aalto University

- Instructor: Pekka Marttinen.

Mentoring

Nghi Vo Dong (Master's student, Aalto University) Jun 2026 –
Summer research internship

- Supervised research on user modelling for adaptive coding assistance, including human coding datasets, coding task simulation and user simulation.

Academic Service

Organising Committee

ELLIS Summer School on AI for Research (with €5K from Helsinki Institute for Information Technology) 2026

ELLIS Summer School on Collaborative and Generative AI (CoGenAI) 2024

Academic Reviews

Program Committee, *AAAI 2027* 2026

Reviewer, *NeurIPS 2026 UserSim Workshop* 2026

Reviewer, *NeurIPS 2026 Trustworthy AI for Good Workshop* 2026

Program Committee, *AAAI 2026* 2025

Reviewer, *NeurIPS 2025 Reliable ML from Unreliable Data Workshop* 2025

Awards and Grants

- Doctoral Incentive Scholarship, *Aalto University*, Espoo, Finland. 2026
- ELIAS Mobility Grant, *ELLIS Doctoral Symposium*, Paris, France. 2024
- ELISE Mobility Grant, *ELLIS Doctoral Symposium*, Helsinki, Finland. 2023
- Aalto Incentive Scholarship, *Aalto University*, Espoo, Finland. 2022 – 2023
- M.Sc. Thesis Grant (€16K), *Foundation for Aalto University Science and Technology*. 2022
- Outstanding Graduate, *NUAA*, Nanjing, China. 2021
- Meritorious Winner (top 6%), *Mathematical Contest in Modeling*, U.S.A. 2020
- Academic Scholarship, First Prize (top 10%), *NUAA*, Nanjing, China. 2018 – 2020
- Excellent Student Scholarship, First Prize (top 2%), *NUAA*, Nanjing, China. 2018 – 2020
- National Scholarship (top 0.2%), *Ministry of Education*, China. 2018
- Excellent summer project on Privacy Leak and Preservation (¥5K), *National Internet Plus Education Program, China Telecom*, China. 2018

Selected Presentations

- More Than Irrational: Modeling Belief-Biased Agents, AAAI 2026, Singapore (Poster).
- More Than Irrational: Modeling Belief-Biased Agents, FCAI Computational Rationality Team (Talk).
- Interactive personalization for explainability via human-in-the-loop multi-objective Bayesian optimization, ELLIS Doctoral Symposium 2023, Helsinki, Finland (Poster).
- Interactive Personalization of Classifiers for Explainability using Multi-Objective Bayesian Optimization, UMAP 2023, Limassol, Cyprus (Oral).
- Interactive Personalization of Classifiers for Explainability using Multi-Objective Bayesian Optimization, HelsinCHI 2023, Helsinki, Finland (Poster).