

FANFEI LI

fanfei.fli@gmail.com · Website · Google Scholar · LinkedIn

EDUCATION

Max Planck Institute for Intelligent Systems PhD in Machine Learning, University of Tübingen Advisor: Prof. Wieland Brendel	Jun 2025 – Expected 2028
University of Tübingen M.S. in Quantitative Data Science Methods (Grade: 1.2, German scale; 1.0 best)	Sep 2022 – May 2025
UCLA Anderson School of Management M.S. in Business Analytics (GPA: 3.8/4.0)	Sep 2019 – Dec 2020
University of California, Los Angeles B.S. in Applied Mathematics, Minor in Music Industry College Honors (UCLA's highest undergraduate academic distinction); Dean's Honors List	Sep 2015 – Jun 2019

SELECTED PUBLICATIONS

- [1] **LittleLearner: Language Models Under Pedagogically Controlled Knowledge Exposure** *Oral: NeurIPS 2026*
F Li, J Zeller*, M Prada-Corral, T Wiedemer, P Mayilvahanan, R Cotterell, W Brendel*
- [2] **MentisOculi: Revealing the Limits of Reasoning with Mental Imagery** *ICML 2026*
J Zeller, T Wiedemer, F Li, T Klein, P Mayilvahanan, M Bethge, F Wichmann, R Cotterell, W Brendel
- [3] **LAION-C: An Out-of-Distribution Benchmark for Web-Scale Vision Models** *ICML 2025*
F Li, T Klein, W Brendel, R Geirhos, R S Zimmermann

* Equal contribution

EXPERIENCE

Max Planck Institute for Intelligent Systems Research Intern <ul style="list-style-type: none">Developed LAION-C, an out-of-distribution robustness benchmark with six novel synthetic distortions designed to challenge vision models trained on web-scale data.Evaluated 50+ vision models across different architectures and training regimes, comparing their performance with human observers to investigate discrepancies in out-of-distribution robustness.	Oct 2023 – May 2025 Tübingen, Germany
TikTok Data Scientist (full-time from March 2021) <ul style="list-style-type: none">Applied time-series clustering to analyze creator growth trajectories, identify behavioral patterns, and inform data-driven product strategies.Used causal inference methods (PSM and difference-in-differences) to quantify the impact of individual and organizational factors on creator growth and monetization.Built and maintained large-scale data pipelines in SQL and Python, providing reliable datasets and metrics for downstream analysis and modeling.	Nov 2020 – Oct 2022 Beijing, China
UCLA Anderson School of Management & Riot Games Research Project <ul style="list-style-type: none">Ran experiments with XGBoost and LSTM models on minute-level League of Legends gameplay data to estimate in-game win probabilities.Used Python to preprocess 300GB of data, performed feature selection, feature engineering, and hyperparameter tuning.	Jul 2020 – Aug 2020 Los Angeles, CA
Amazon Web Services Data Engineer Intern <ul style="list-style-type: none">Built and deployed production ETL pipelines to ingest data from Amazon S3 into Redshift, including modifications to Java and Ruby code to migrate storage endpoints between AWS regions.Developed automated data processing and reporting workflows in SQL and Python to support analysis of subscription, revenue, and customer behavior.	May 2020 – Sep 2020 Beijing, China