

Shizhe He

+1 650 250 9154 | shizhehe@stanford.edu | [in linkedin.com/in/shizhehe](https://www.linkedin.com/in/shizhehe) | github.com/shizhehe | [personal website](#)

EDUCATION

Stanford University

Stanford, USA

B.Sc. in Computer Science (AI Track), Minor in International Relations; GPA: 3.98/4.0

Sep 2022 – Jun 2026

- *Societies:* European Student Association *Head of Internal Affairs*, Club Lacrosse, Sigma Phi Epsilon *VP of Communications*, Stanford Neurotech *Team Lead*; formerly BASES *VP of Core*, Club Swim *Captain of Spirit*
- *Relevant Coursework:* Systems for Machine Learning, Mining Massive Datasets, Deep Learning for Computer Vision

WORK EXPERIENCE

Beacon Text

San Francisco, USA

Co-Founder, [website](#)

Mar 2024 – Oct 2024

- Led development of RCS messaging advertising SAAS product, deploying AWS Cloudfront, S3, and Elastic Beanstalk.
- Scaled company to **three** corporate design retail partners with a total addressable audience of **300.000** contacts.
- Part of **Keel Cohort 1.0, Pear Garage 2023/24**

QuantCo

San Francisco, USA

Data Engineering - Machine Learning Intern

Jun 2024 – Aug 2024

- Designed custom text embedding models and fine-tuning pipeline for fundraiser descriptions on one of the **largest crowdfunding-platforms**. Fine-tuned **transformer-based embeddings** to encode fundraiser content and quality. Used as fundraiser-level features in different workstreams.
- Built large parts of training & analyses pipeline for new workstream to suggest optimal fundraiser goal during creation. Improved production model using text embeddings by **150%**. [Impact numbers protected under non-disclosure agreement.]
- Applied feature selection methods on gradient-boosted trees to reduce complexity and training time by **80%** with no significant reduction in model performance across different workstreams

Theros LLC

Remote

Software Engineering Intern

Jun 2023 – Sep 2023

- Co-developed **Flask API** for “Jot it Down”, a ChatGPT plugin for cross-chat and multi-user memory
- Implemented a comprehensive **Retrieval-Augmented Generation (RAG)** system and modular workflow structures to index and digest live internal wiki content. Used in customer support, internal digest tools, and automated emails
- Led **robust logging** pipeline and processing on **AWS** from the ground up, along with metric computation and visualization
- Built a mock language model toolbox for tests, reducing testing operation costs by over **90%**

Check24 Factory GmbH

Munich, Germany

Data Science Working Student

Jun 2022 – Aug 2022

- Developed end-to-end machine learning solution for product **popularity ranking**, enhancing sales for energy business unit
- Improved ranking and response times by developing fine-tuned machine learning models and real-time **fastAPI**
- Streamlined business unit operations by deploying ranking service on **AWS** and conducting **A/B testing**

Infineon Technologies AG

Neubiberg, Germany

Applied Machine Learning Intern

Jul 2021 – Aug 2021

- Developed internal CRM using **Flask**, improving how we match requirements and verifications in microchip design
- Enhanced language capabilities using **BERT**, automating requirement-verification matching process based on language

RESEARCH/PROJECT EXPERIENCE

Brains in Silicon, Stanford University

Stanford, USA

Student Researcher; Advised by Saarthak Sarup, Kwabena Boahen

Oct 2024 – present

- Investigating **neuromorphic computing**—dendritic computation for knowledge systems (e.g. RAG)

Translational AI Lab, Stanford University

Stanford, USA

Student Researcher; Advised by Magda Paschali, Ehsan Adeli

Apr 2023 – Sep 2024

- Investigating contrastive **self-supervised** learning with SO3-equivariance pretext task on 3D brain MRIs using **PyTorch**
- Designed model to understand underlying brain structures, involved in brain foundational model, [MLCN 2024 publication](#)

The Legend of Shizhelda for the Rand Entertainment System

Stanford, USA

CS107E Class Project; Taught by Julie Zelenski, Pat Hanrahan

Jan 2023 – Apr 2023

- Developed homage of 2D Zelda games from scratch with **bare-metal C** on a Raspberry Pi (ARM)

Lab for AI in Medicine, Technical University of Munich

Munich, Germany

Student Researcher; Advised by Kerstin Hammernik, Daniel Rueckert

Apr 2021 – Jun 2022

- Investigated **data distribution shift** in MRI reconstruction. Adapted methods to novel dynamic **7T MRI** reconstruction
- **Youngest** presenter at the 2022 ISMRM-ESMRMB Joint Annual Meeting/Conference in London, [abstract](#), [publication](#)

SKILLS

Languages: German (Fluent), Mandarin Chinese (Fluent), English (Fluent), French (Conversational)

Programming: Python, C, C++, React, Assembly, SQL, AWS, Git, Microservices Architectures, CUDA, NLP, CV

Frameworks & Libraries: PyTorch, Tensorflow, LightGBM, Huggingface, fastAPI, Flask, Docker, Pandas/Polars, Snowflake

Interest: Artificial Intelligence, Water Sports, Tennis, Lacrosse, Photography, Automotive Racing