Zachary Fisher

Staff Software Engineer focused on post-training and reasoning improvements for Gemini.

zachfisher.com linkedin.com/in/zkf me@zachfisher.com

EMPLOYMENT

Google DeepMind (formerly Google Research)

Staff Software Engineer (L6)

Mountain View, CA April 2020 –

- Trained and deployed auto-rater models critical for the launch of LLM use of map tools in AI Search Overviews; launched for all users of Google Search AI Overviews in 2024.
- Led a 5-engineer team to create attribute extraction models for Google Shopping and Google Maps, with double-digit million dollar revenue impact, including demo at Google I/O 2023.
- Developed and submitted novel large-scale training and evaluation data for Google Gemini for data science agents, puzzles, and mathematics.
- Distilled and launched first transformer model to run at scale over 1B+ shopping webpages indexed by Google.

Perimeter Institute for Theoretical Physics

Waterloo, ON, Canada

Postdoctoral Researcher

September 2017 – February 2020

It from Qubit postdoctoral fellow (Simons Collaboration on Quantum Fields, Gravity and Information).

EDUCATION

University of California, Berkeley

August 2012 – *May* 2017

Ph.D. in Physics. Thesis title: "Entropy Bounds and Entanglement." Formulated and proved the quantum null energy condition (500+ citations).

Massachusetts Institute of Technology

September 2008 – June 2012

S.B. in Physics. Thesis title: "Shuttling of ions for characterization of a microfabricated ion trap."

SELECTED PUBLICATIONS

Author of seven publications in the field of **machine learning** and seven publications in the field of **high-energy physics**, including:

- Y. Sheng, S. Gandhe, B. Kanagal, N. Edmonds, Z. Fisher, S. Tata, A. Selvan.
 "Measuring an LLM's Proficiency at Using APIs: A Query Generation Strategy." KDD-ADS (2024).
- R. Aksitov, S. Miryoosefi, Z. Li, D. Li, S. Babayan, K. Kopparapu, **Z. Fisher**, R. Guo, S. Prakash, P. Srinivasan, M. Zaheer, S. Kumar.

"ReST meets ReAct: Self-Improvement for Multi-Step Reasoning LLM Agent." cs.CL/2312.10003

• J. Ainslie, S. Ontanon, C. Alberti, V. Cvicek, **Z. Fisher**, P. Pham, A. Ravula, S. Sanghai, Q. Wang, L. Yang. "ETC: Encoding Long and Structured Inputs in Transformers." EMNLP (2020). stat.ML/2004.08483

Full list of pulications is available at zachfisher.com/pubs.

SKILLS

Languages & Frameworks: Python (JAX, PyTorch, TensorFlow), C++, JavaScript

Data & Infra: Apache Beam, BigQuery, Kubernetes, GCP

Scientific: NumPy, SciPy, SymPy, Mathematica