Gustaf Ahdritz

gahdritz@g.harvard.edu · +1 914-506-0204

Google Scholar, GitHub

Education

Graduate School of Arts and Sciences, Harvard University
Ph.D., Computer Science. Sept. 2022 – Present
Advisors: Prof. Boaz Barak and Jonathan Frankle

School of Engineering and Applied Sciences, Columbia University
M.Sc., Computer Science. GPA: 4.08 Sept. 2020 – May 2021
Advisor: Prof. Mohammed AlQuraishi

Columbia College, Columbia University
B.A. (*summa cum laude*), Computer Science & History. GPA: 4.07 Sept. 2016 – May 2020
Advisors: Profs. Kathleen McKeown and Matthew Connelly

Work Experience

AIML @ Apple
Ph.D. Intern April 2024 – September 2024

Worked on uncertainty quantification, especially for language models. Produced a preprint ("Provable Uncertainty Decomposition via Higher-Order Calibration").

AlQuraishi Lab @ Columbia University
Researcher May 2021 – Aug. 2022

Continued work on OpenFold as a predoctoral researcher.

Apple
Software Engineering Intern Summer 2020

Designed, developed, and tested a new macOS application for the Screen Time device management team.

Publications and Preprints

[1] **G. Ahdritz**, A. Gollakota, P. Gopalan, C. Peale, U. Wieder. Provable Uncertainty Decomposition via Higher-Order Calibration. *Under review*, 2024.

[2] *G. Tanzer, ***G. Ahdritz**, L. Melas-Kyriazi. Modeling Real-Time Interactive Conversations as Timed Diarized Transcripts. *arXiv*, 2024.

[3] ***G. Ahdritz**, *T. Qin, N. Vyas, B. Barak, B.L. Edelman. Distinguishing the Knowable from the Unknowable with Language Models. *ICML 2024*, 2024.

[4] **G. Ahdritz**, N. Bouatta, S. Kadyan, L. Jarosch, D. Berenberg, I. Fisk, A. M. Watkins, R. Bonneau, M. AlQuraishi. OpenProteinSet: Training data for structural biology at scale. *NeurIPS 2023 (Datasets and Benchmarks track)*, 2023.

[5] S. Song, B. Kruft, M. Zhang, C. Li, S. Chen, C. Zhang, M. Tanaka, X. Wu, M. AlQuraishi, **G. Ahdritz**, C. Floristean, R. Stevens, V. Vishwanath, A. Ramanathan, S. Foreman, K. Hippe, P. Balaprakash, Y. He. DeepSpeed4Science Initiative: Enabling Large-Scale Scientific Discovery through Sophisticated AI System Technologies. *4th AI for Science Workshop at NeurIPS 2023*, 2023.

[6] *L. Bailey, ***G. Ahdritz**, *A. Kleiman, S. Swaroop, F. Doshi-Velez, W. Pan. Soft prompting might be a bug, not a feature. *Workshop on Challenges in Deployable Generative AI at ICML 2023,* 2023.

[7] ***G. Ahdritz**, *N. Bouatta, S. Kadyan, Q. Xia, W. Gerecke, T. J. O'Donnell, D. Berenberg, I. Fisk, N. Zanichelli, B. Zhang, A. Nowaczynski, B. Wang, M. M. Stepniewska-Dziubinska, S. Zhang, A. Ojewole, M. Efe Guney, S. Biderman, A. M. Watkins, S. Ra, P. Ribalta Lorenzo, L. Nivon, B. Weitzner, Y. A. Ban, S. Chen, M. Zhang, C. Li, S. L. Song, Y. He, P. K. Sorger, E. Mostaque, Z. Zhang, R. Bonneau, M. AlQuraishi. OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods* 21:1514-1524, 2024.

[8] R. Chowdhury, N. Bouatta, S. Biswas, C. Floristean, A. Kharkar, K. Roy, C. Rochereau, **G. Ahdritz**, J. Zhang, G. M. Church, P. K. Sorger, M. AlQuraishi. Single-sequence protein structure prediction using a language model and deep learning. *Nature Biotechnology* 40:1617–1623, 2022.

Invited Talks

SBGrid Consortium April 2023
OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization

OpenBioML Jan. 2023
OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization

Mosaic NLP Workshop Oct. 2022
OpenFold

NSF ICICLE AI Institute Sept. 2022
OpenFold: Technical Overview

Open-Source Software and Datasets

OpenFold Aug. 2021 – Present

Wrote and currently maintain the first trainable open-source reproduction of AlphaFold2. Includes original AlphaFold2 parameters under an open license. Work featured at NVIDIA GTC and by AWS, TechCrunch, and Weights and Biases.

OpenProteinSet June 2022

Produced the largest public database of protein alignment data, including multiple sequence alignments and template hits for more than 4 million diverse protein sequences. Can be used to train or finetune machine learning models for protein-related tasks at the scale of AlphaFold2.

Academic Awards and Fellowships

Kempner Institute Graduate Fellowship Sept. 2023

Five-year fellowship awarded by Harvard's Kempner Institute to "pursue research central to the Kempner's core mission of understanding the basis of intelligence in natural and artificial systems."

Ashford Fellowship Sept. 2022

Five-year fellowship awarded to Harvard GSAS students "highly likely to make a substantial impact in their chosen field of study, as well as in society."

Andrew P. Kosoresow Memorial Award May 2020

Awarded to teacher's assistants in the Columbia SEAS Computer Science Department "for outstanding contributions to teaching in the Department of Computer Science and exemplary service to the Department and its mission."

Dean Hawkes Memorial Prize April 2019

Awarded to the Columbia College junior "judged to be the most deserving on the basis of work in the humanities."

Teaching

Harvard CS229br: Foundations of Deep Learning Spring 2023
Teaching Fellow Professor: Boaz Barak

Columbia COMS3157: Advanced Programming Spring 2019 – Spring 2021
Head Teaching Assistant (S20 – S21), Teaching Assistant (S19 – F19) Professor: Jae Woo Lee

Reviewing

ICLR 2025
ICLR, ICML, NeurIPS 2024
ICML, NeurIPS (top reviewer), Current Opinion in Structural Biology 2023