

Gustaf Ahdriz

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

[gahdriz.github.io](https://github.com/gahdriz)

Education

Graduate School of Arts and Sciences, Harvard University

Ph.D., Computer Science.

Sept. 2022 – Present

Advisors: Profs. Jonathan Frankle and Boaz Barak

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

Publications and Preprints

- [1] G. Ahdriz, 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. *arXiv*, 2023. doi.org/10.48550/arXiv.2308.05326.

To appear at NeurIPS 2023 (Datasets and Benchmarks track).

- [2] *L. Bailey, *G. Ahdriz, *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.

- [3] *G. Ahdriz, *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, 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. *bioRxiv*, 2022. doi.org/10.1101/2022.11.20.517210.

Presented at the *Critical Assessment of Structure Prediction (CASP) 15*, 2022.

- [4] R. Chowdhury, N. Bouatta, S. Biswas, C. Floristean, A. Kharkar, K. Roy, C. Rochereau, G. Ahdriz, 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.

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.

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

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

Work Experience

- AlQuraishi Lab, Columbia University**
Researcher May 2021 – Aug. 2022
Continued work on OpenFold as a predoctoral researcher.
- Apple, Inc.**
Software Engineering Intern Summer 2020
Designed, developed, and tested a new macOS application for the Screen Time device management team.

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.”

Reviewing

ICLR

2024

ICML, NeurIPS, Current Opinion in Structural Biology

2023