

# Jonathan Deaton

[Website](#) | [LinkedIn](#) | [GitHub](#) | [Scholar](#)

999 Green Street, San Francisco, California  
[jonpauldeaton@gmail.com](mailto:jonpauldeaton@gmail.com) | (415) 531-1568

## EXPERIENCE

---

### Member of Technical Staff

*EvolutionaryScale*

December 2023 – Present

*San Francisco, California*

- Researched development of LLMs (up to 98B parameters) trained on large-scale protein and genomic sequence data
- Studied scaling laws for unsupervised learning of protein structure, highlighted in [ESMc model release](#)
- Delivered technical solution for multi-modal supervision of protein function annotations in [ESM3](#)
- Enhanced learning of protein interactions in pre-training through novel data and evaluation innovation
- Invented novel architecture for bidirectional supervision superior to BERT MLM in perplexity and representation learning, and enabling efficient inference with key-value caching

### Research Scientist (previously Research Engineer)

*X Development (Google X)*

November 2021 – November 2023

*Mountain View, California*

- Delivered diverse machine learning approaches for real-world biomolecular sequence design problem including:
  - \* Fine-tuned BERT-like model which significantly outperformed our team's existing methods in molecular design by increasing experimental hit-rate from 58% to 96%
  - \* Self-supervised pre-training of BERT on large biological nucleotide sequence datasets, eliminating dependency on data collection from an unreliable experimental assay through transfer learning
  - \* Decoder-only LM with [RLHF](#) fine-tuning to accelerate generation of biologically plausible candidates
  - \* Investigated zero-shot capabilities of RNA language model in screening loss-of-function mutations
- Demonstrated robust RNA structure representations emerge in self-supervised pre-training of RNA language models
- Designed and built scalable distributed training system in JAX supporting language models. Accelerated rapid prototyping of ML algorithms for molecular design by integrating with [T5 architectures](#), speeding up training by 2-5x, and increasing algorithmic expressiveness and data-scaling

### Software Engineer

*Google Health / Google Research*

September 2019 – November 2021

*Palo Alto, California*

- Led development of software to validate computer-vision models for dermatological condition classification
- Contributed to productionization of the [DermAssist](#) ML model thereby achieving regulatory approval as a CE-marked Class I medical device
- Conducted research advancing state-of-the-art computer-vision models and published research findings on long-tail class-imbalance

### Undergraduate Bioengineering Researcher

*Quake Lab*

June 2014 – June 2017

*Stanford, CA*

- Identified novel bacteriophages in metagenomic sequence data by developing data-analysis, visualization, and ML software in Python
- Performed molecular biology laboratory experiments including cell culture, qPCR, DNA sequencing, microfluidic device fabrication
- Communicated scientific findings in academic conferences, lab meetings, and publications

## EDUCATION

---

### Stanford University

*MS in Computer Science, Artificial Intelligence*

January 2018 – June 2019

*Stanford, California*

- GPA: 3.957 / 4.0
- Course Assistant: CS 140: Operating Systems, CS 161: Design and Analysis of Algorithms
- Coursework: Probabilistic Graphical Models, Machine Learning Theory, Deep Learning, Reinforcement Learning

### Stanford University

*BS in Bioengineering with Honors*

September 2012 – March 2018

*Stanford, California*

- GPA: 3.839 / 4.0
- [Research thesis](#): Identified Bacteriophage genomes in metagenomic sequencing data with machine learning

## PUBLICAITONS

---

- [Simulating 500 million years of evolution with a language model](#) . Hayes T, Rao R, Akin H, Sofroniew N, Oktay D, Lin Z, Verkuil R, Tran V, **Deaton J**, Wiggert M, Badkundri R, Shafkat I, Gong J, Derry A, Molina R, Thomas N, Khan Y, Mishra C, Kim C, Bartie L, Nemeth M, Hsu P, Sercu T, Candido S, Rives A
- [Big Self-Supervised Models Advance Medical Image Classification](#). Azizi S, Mustafa B, Ryan F, Beaver Z, Freyberg F, **Deaton J**, Loh A, Karthikesalingam A, Kornblith S, Chen T, Natarajan V, Norouzi M. International Conference on Computer Vision 2021.
- [Underspecification Presents Challenges for Credibility in Modern Machine Learning](#). D'Amour A, Heller K, et al. Journal of Machine Learning Research, 2020.
- [Addressing the Real-world Class Imbalance Problem in Dermatology](#). Weng WH, **Deaton J**, Natarajan V, Elsayed GF, Liu Y. NeurIPS Machine Learning for Health Workshop 2020.
- [Mini-Metagenomics and Nucleotide Composition Aid the Identification and Host Association of Novel Bacteriophage Sequences](#). **Deaton J**, Yu F, Quake S. Advanced biosystems, 2019.

## PROJECTS

---

### **Ring Attention for GPU in JAX/Pallas**

Personal/FOSS Project

- Implemented [Ring Attention](#) in JAX with Flash Attention Pallas kernels for GPU. Eanbles arbitrary user-defined attention biases to support common attention variants.

### **Transformer Pointer-Generator Networks for Summarization**

Stanford CS 224N Course Project

- Implemented fusion of Transformer Encoder/Decoder and pointer-generator network for text summarization
- Successfully trained novel deep network architecture with low compute resources

### **AgarLE: Competitive Multi-Agent RL Environment**

Personal/FOSS Project

- Build high-performance multi-agent reinforcement learning environment in C++
- Benchmarked state-of-the-art RL methods such as Proximal Policy Optimization and Deep-Q Learning