Andrew Dickson

612-257-5651 | amdickson128@gmail.com | github.com/amdson

Summary

- Machine learning for protein biology and bioinformatics, including novel model architectures and data pipelines, gene ontology models, and protein sequence/function prediction at large scale.
- Strong coding and systems background, including HPC workflows, Linux command-line, web app development, and computational biology tools.
- Research in protein biophysics, biofilm modeling, agent-based simulations, and Bayesian Monte Carlo methods.

Education

University of California - Berkeley Berkeley, CA

Ph.D. in Applied Science and Technology, 3.91 GPA *Aug. 2020 – Dec. 2025*

Concentrations in Machine Learning, Bioinformatics, Biophysics

University of California - Berkeley Berkeley, CA

Bachelor of Science in Electrical Engineering and Computer Science, 3.86 GPA *Aug. 2016 – May 2020*

Minor in Mechanical Engineering

Research Experience

Postdoc January 2026 – Present

United States Department of Agriculture *Albany, CA*

- Built a tractable method for efficient continued training of genomic sequence models, on highly redundant variant data

Graduate Student Researcher August 2020 – August 2025

University of California - Berkeley *Berkeley, CA*

- Architected and deployed the core infrastructure for the Go Bench universal benchmarking hub (gobench.org).
- Developed and implemented novel machine learning algorithms for protein function annotation through embedding search, as well as for unsupervised annotation attribution for interpretability.
- Gigabyte-scale computational analysis of the FG-Nucleoporin protein family
- Mentored 4 Fung Masters students towards a graduate thesis on interpretable protein function annotation.
- Designed and implemented GPU accelerated physics and reaction-diffusion dynamics for agent-based model of biofilm growth within the human gut microbiome.

Student Intern August 2023 – November 2023

United States Department of Agriculture *Albany, CA*

- Developed interactive website for design and validation of STARP PCR primers using Flask API and HTML/Javascript frontend.
- Designed and developed BLAST backend for selective filtration of multiply associative PCR primers.
- Designed custom Redis/Celery based backend for asynchronous handling of computationally intensive requests.

Industry & Teaching Experience

Software Developer May 2020 – August 2020

Stottler Henke Associates, Inc. *San Mateo, CA*

- Designed and implemented A* variant for multiple entity routing in real-world environments
- Developed machine learning protocol for categorizing bacteria from hyperspectral images using ResNet architecture

Educational Writer May 2016 – August 2019

Brilliant.org *Berkeley, CA*

- Wrote nine chapters in the Brilliant.org machine learning curriculum including those on Support Vector Machines, Linear and Tree Regression, PCA, ICA, Expectation-Maximization, and Cross-Validation, with over 100,000 readers
- Designed cryptography inspired curriculum for novice Python students
- Rewrote computer science course on algorithms and data structures for correctness and clarity

Graduate Student Instructor January 2020 – June 2025

University of California - Berkeley *Berkeley, CA*

- Developed homework, exam, and review material for Bioengineering 102 materials science course, receiving Department of Bioengineering Outstanding GSI Award
- Six semesters teaching communicative writing for engineers with the Berkeley Fung Masters program

Publications

- Peyro M, Soheilypour M, Nibber VS, **Dickson AM**, Mofrad MRK. "FG nucleoporins feature unique patterns that distinguish them from other IDPs." *Biophys J.* 2021; http://doi.org/10.1016/j.bpj.2021.06.031
- Peyro M, **Dickson A.**, Mofrad MRK. "Nucleoporins' exclusive amino acid sequence features regulate their transient interaction with and selectivity of cargo complexes in the nuclear pore." *Mol. Bio. of the Cell*, 2021; http://doi.org/10.1091/mbc.E21-04-0161
- **Dickson A.**, Asgari E, et. al. "Go Bench: Shared-Hub for Universal Benchmarking of Machine Learning-Based Protein Functional Annotations." *Bioinformatics*, 2022; http://doi.org/10.1093/bioinformatics/btad081
- **Dickson A**, Mofrad MRK. "Fine-tuning protein embeddings for functional similarity evaluation." *Bioinformatics*, 2023; http://doi.org/10.1093/bioinformatics/btae445
- Valiei A, **Dickson A.**, Mofrad MRK, Amanian-D J. "Metabolic interactions shape emergent biofilm structures in a conceptual model of gut mucosal bacterial communities." *npj Biofilms and Microbiomes*, 2024; http://doi.org/10.1038/s41522-024-00572-y
- Parsa M., Aminian-D. J. **Dickson A.**, Mofrad MRK "SIMBA-GNN: Simulation-augmented Microbiome Abundance Graph Neural Network" *npj Systems Biology and Applications*, 2025; http://doi.org/10.1101/2025.05.27.656377
- Amanian-D J., **Dickson A.***, et. al. "MetaBiome: a multiscale model integrating agent-based and metabolic networks to reveal spatial regulation in gut mucosal microbial communities". mSystems, 2025; http://doi.org/10.1128/msystems.01652-24 [*Co-First Author]
- **Dickson A.**, Mouline, S, Mofrad MRK "Residue-level Protein Function Prediction with Interpretability Methods" Bioinformatics, (in press)
- **Dickson A.**, Mofrad MRK, Amanian-D J. "Bacteria-host interactions determine ecological outcomes in gut mucosal community models" *Manuscript submitted for publication*, 2026
- Long Y., **Dickson A.***, et. al. "STARP Primer: a web-based program of primer design for flexibly genotyping individual single nucleotide polymorphisms" *Manuscript submitted for publication*, 2026 [*Co-First Author]
- **Dickson A**, Mofrad MRK. "Bayesian Calibration of Agent-Based Models for Rare Statistical Events". PNAS, (in press);
- **Dickson A**, Mofrad MRK. "Coupling Genome-Scale Representations with Generative Modeling for the Gut Microbiome" *Springer*, 2025

Conference Presentations

- Asgari E., **Dickson A.**, et. al "Fine-tuning of Language Model-Based Representation for Protein Functional Annotation" *Intelligent Systems for Molecular Biology*, 2020
- **Dickson A.**, Asgari E., et. al "GO Bench: Shared-hub for Universal Benchmarking of Machine Learning-Based Protein Functional Annotations" *Intelligent Systems for Molecular Biology*, 2022
- **Dickson A.**, Mofrad MRK "Fine-tuning Protein Embeddings for Generalizable Protein Annotation" *Bioengineering Institute of California*, 2023
- **Dickson A.**, Valiei A, Mofrad MRK, "Agent Based Modeling of Invasion Scenarios in the Gut Microbiome" *Biomedical Engineering Society*, 2024
- Mouline, S, **Dickson A.**, Mofrad MRK, "Identifying Functionally Important Residues: BERT Entropy and Interpretability Applied to Protein Function Prediction" *Biomedical Engineering Society*, 2024

Technical Skills

Languages: Python, C/C++, Java, JavaScript, R, Matlab, Taichi, SQL, HTML/CSS
ML & HPC: PyTorch, JAX, Lightning, HuggingFace, CUDA, SLURM, Weights & Biases, ONNX, TensorFlow, AWS
Scientific Computing: NumPy, SciPy, Scikit-Learn, Pandas, Polars, FAISS, PyMC, BioPython
Bioinformatics: MMseqs2, BLAST, HMMER, MUSCLE/MAFFT, BEAST, MEGA, PyMOL, VMD, UniProt, Pfam, InterPro, Gene Ontology
Web & Infrastructure: Flask, Django, Celery, Redis, Nginx, React, Node.js, Docker, Git
Other: Inkscape, Adobe Illustrator, COMSOL, SolidWorks