

Daniel Berenberg

 [djberenberg](#) —  [daniel-j-berenberg](#)

 danielberenberg@gmail.com

 +1 609 670 4667

EDUCATION

- **Ph.D., Computer Science**, New York University, New York, NY, USA, **2020–2025**
Thesis: *Modern Machine Learning Methods for Protein Design*
- **M.S., Computer Science**, University of Vermont, Burlington, VT, USA **2017–2018**
- **B.S., Mathematics**, University of Vermont, Burlington, VT, USA **2013–2017**

EXPERIENCE

Mythos Scientific, *New York, NY* 2025 –
Technical Staff/Founder Dec 2025 –

- Co-leading all aspects of company operations including strategic direction, AI research, technology development, and investment management.

Center For Data Science, New York University, *New York, NY* 2024 – 2025
Undergraduate Data Science Tutor Aug 2024 – May 2025

- Facilitated targeted small-group sessions for students needing additional support.
- Developed course material including problem sets, exams, and supplementary information.

Teaching Assistant Jan 2024 – May 2024

- Led recitations, held office hours, and graded assessments for an entry-level undergraduate data science course on fundamental machine learning and statistics.

Prescient Design, Genentech, *New York, NY* 2021 – 2024
Machine Learning Scientist Aug 2023 – Jan 2024

- Co-invented a state-of-the-art generative model for antibody design, achieving over 95% sequence novelty and experimental validation with 100% expression and 30–75% binding rates to therapeutically relevant targets.

Machine Learning Engineer Apr 2021 – Aug 2023

- Founding member of Prescient Design. Supported acquisition by Genentech; helped oversee the integration and landing process of team, tools, and workflows.
- Developed and maintained core infrastructure and workflows for lab-in-the-loop.
- Participated in recruiting and technical interviews for engineering and research staff.

Center For Computational Biology, Flatiron Institute, *New York, NY* 2019 – 2020
Research Analyst Mar 2019 – Aug 2020

- Helped develop a structure-based protein function predictor using graph neural networks.
- Contributed a dataset of computational protein structure models and function annotations that has since become a standard benchmark in structure-based protein function prediction.
- Collaborated on building an Apache Spark-based pipeline to perform function annotation at scale; deployed on *Trichomonas* and *Plasmodium* genomes.

- Studied complex networks of crowdsourced opinions on causal attribution.
- Analyzed text data using NLP tools and language model representations.
- Developed a graph alignment method to fuse causal attribution networks.

- Designed a computer vision system to predict the heart rate and respiratory rate of a subject from a video of their thumb pressed against the camera.
- Developed high performance 3D convolutional network using Keras and Tensorflow.
- Implemented video preprocessing and manipulation software in python libraries including OpenCV, PIL, NumPy, and SciPy.

PUBLICATIONS

- Barot, M., **Berenberg, D.**, & Khajehabdollahi, S. (2026, August). *Architecture generalization with metanca* (Vol. ALIFE 2026: Proceedings of the 2026 Artificial Life Conference). <https://doi.org/10.1162/ISAL.a.951>
- Berenberg, D.**, Gruver, N., Amin, A., Mørch Groth, P., Chen, L., Srivastava, H., Notin, P., Marks, D., Gordon Wilson, A., Cho, K., & Bonneau, R. (2025). Residue-level text conditioning for protein language model mutation effect prediction. *ICLR 2025 Workshop on Generative and Experimental Perspectives for Biomolecular Design*.
- Frey, N. C., Hötzel, I., Stanton, S. D., Kelly, R., Alberstein, R. G., Makowski, E., Martinkus, K., **Berenberg, D.**, Bevers, J., Bryson, T., Chan, P., Czuby, A., D'Souza, T., Dwyer, H., Dziewulska, A., Fairman, J. W., Goodman, A., Hofmann, J., Isaacson, H., ... Gligorijević, V. (2025). Lab-in-the-loop therapeutic antibody design with deep learning. <https://doi.org/10.1101/2025.02.19.639050>
- Srivastava, H. R., **Berenberg, D.**, Qassab, O., Wang, Z., Bonneau, R., & Carlton, J. M. (2025). Plasmofp: Leveraging deep learning to predict protein function of uncharacterized proteins across the malaria parasite genus. *BioRxiv*. <https://doi.org/10.1101/2025.09.12.675843>
- Sullivan, S. A., Orosco, J. C., Callejas-Hernández, F., Blow, F., Lee, H., Ranallo-Benavidez, T. R., Peters, A., Raidal, S. R., Girard, Y. A., Johnson, C. K., Rogers, K. H., Gerhold, R., Mangelson, H., Liachko, I., Srivastava, H., Chandler, C., **Berenberg, D.**, Bonneau, R. A., Huang, P.-J., ... Carlton, J. M. (2025). Comparative genomics of the parasite trichomonas vaginalis reveals genes involved in spillover from birds to humans. *Nature Communications*, 16(1). <https://doi.org/10.1038/s41467-025-61483-w>
- Ahdritz, G., Bouatta, N., Floristean, C., Kadyan, S., Xia, Q., Gerecke, W., O'Donnell, T. J., **Berenberg, D.**, Fisk, I., Zanichelli, N., Zhang, B., Nowaczynski, A., Wang, B., Stepniewska-Dziubinska, M. M., Zhang, S., Ojewole, A., Guney, M. E., Biderman, S., Watkins, A. M., ... AlQuraishi, M. (2024). Openfold: Retraining alphafold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods*, 21(8), 1514–1524. <https://doi.org/10.1038/s41592-024-02272-z>
- Frey, N. C., **Berenberg, D.**, Zadorozhny, K., Kleinhenz, J., Lafrance-Vanasse, J., Hotzel, I., Wu, Y., Ra, S., Bonneau, R., Cho, K., Loukas, A., Gligorijevic, V., & Saremi, S. (2024). Protein discovery with discrete walk-jump sampling. <https://arxiv.org/abs/2306.12360>
- Hamamsy, T., Morton, J. T., Blackwell, R., **Berenberg, D.**, Carrierio, N., Gligorijevic, V., Strauss, C. E., Leman, J. K., Cho, K., & Bonneau, R. (2024). Protein remote homology detection and structural alignment using deep learning. *Nature biotechnology*, 42(6), 975–985.
- Ahdritz, G., Bouatta, N., Kadyan, S., Jarosch, L., **Berenberg, D.**, Fisk, I., Watkins, A., Ra, S., Bonneau, R., & AlQuraishi, M. (2023). Openproteinset: Training data for structural biology at scale. *Advances in Neural Information Processing Systems*, 36, 4597–4609.

- Frey, N. C., Berenberg, D., Kleinhenz, J., Hotzel, I., Lafrance-Vanasse, J., Kelly, R. L., Wu, Y., Rajpal, A., Ra, S., Bonneau, R., Cho, K., Loukas, A., Gligorijevic, V., & Saremi, S. (2023). Learning protein family manifolds with smoothed energy-based models. *ICLR 2023 Workshop on Physics for Machine Learning*. <https://openreview.net/forum?id=IilnB8jfoP9>
- Koehler Leman, J., Szczerbiak, P., Renfrew, P. D., Gligorijevic, V., **Berenberg, D.**, Vatanen, T., Taylor, B. C., Chandler, C., Janssen, S., Pataki, A., Carriero, N., Fisk, I., Xavier, R. J., Knight, R., Bonneau, R., & Kosciolk, T. (2023). Sequence-structure-function relationships in the microbial protein universe. *Nature Communications*, 14(1). <https://doi.org/10.1038/s41467-023-37896-w>
- Mahajan, S. P., Frey, N. C., **Berenberg, D.**, Kleinhenz, J., Bonneau, R., Gligorijevic, V., Watkins, A., & Saremi, S. (2023). Exploiting language models for protein discovery with latent walk-jump sampling. *Machine Learning for Structural Biology Workshop, NeurIPS*.
- Berenberg, D.**, Lee, J. H., Kelow, S., Park, J. W., Watkins, A., Gligorijević, V., Bonneau, R., Ra, S., & Cho, K. (2022). Multi-segment preserving sampling for deep manifold sampler. *arXiv preprint arXiv:2205.04259*.
- Hamamsy, T., Morton, J. T., **Berenberg, D.**, Carriero, N., Gligorijevic, V., Blackwell, R., Strauss, C. E., Leman, J. K., Cho, K., & Bonneau, R. (2022). Tm-vec: Template modeling vectors for fast homology detection and alignment. *bioRxiv*, 2022–07.
- Gligorijević, V., **Berenberg, D.**, Ra, S., Watkins, A., Kelow, S., Cho, K., & Bonneau, R. (2021). Function-guided protein design by deep manifold sampling. *bioRxiv*, 2021–12.
- Gligorijević, V., Renfrew, P. D., Kosciolk, T., Leman, J. K., **Berenberg, D.**, Vatanen, T., Chandler, C., Taylor, B. C., Fisk, I. M., Vlamakis, H., Xavier, R. J., Knight, R., Cho, K., & Bonneau, R. (2021). Structure-based protein function prediction using graph convolutional networks. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-23303-9>
- Morton, J. T., Strauss, C. E., Blackwell, R., **Berenberg, D.**, Gligorijevic, V., & Bonneau, R. (2020). Protein structural alignments from sequence. *BioRxiv*, 2020–11.
- Bagrow, J. P., **Berenberg, D.**, & Bongard, J. (2018). Neural language representations predict outcomes of scientific research. *arXiv preprint arXiv:1805.06879*.
- Berenberg, D.**, & Bagrow, J. P. (2018a). Efficient crowd exploration of large networks: The case of causal attribution. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 1–25.
- Berenberg, D.**, & Bagrow, J. P. (2018b). Inferring the size of the causal universe: Features and fusion of causal attribution networks. *arXiv preprint arXiv:1812.06038*.

PATENTS

- Mahajan, S. P., Saremi, S., **Berenberg, D.**, Bonneau, R. A., Cho, K., Frey, N. C., & Gligorijević, V. (2026, February). Generative protein design with smoothed energy-based models [US Patent App. 19/288,834].
- Gligorijevic, V., Kelow, S. P., Lee, J. H., Park, J. W., Ra, S. R., Watkins, A. M., **Berenberg, D.**, Bonneau, R. A., & Cho, K. (2025, June). Protein design with segment preservation [US Patent App. 18/842,226].
- Kelow, S. P., **Berenberg, D.**, Nerli, S., Leaver-Fay, A. P., Maguire, J. B., Chungyoun, M. F., Watkins, A. M., Lee, J. H., Ra, S. R., Dwyer, H. V., Lee, M. M., Cho, K., Bonneau, R. A., & Gligorijevic, V. (2025, August). Hybrid protein design [US Patent App. 19/012,402].

HONORS & AWARDS

- ICLR Outstanding Paper Award, *Protein Discovery with Discrete Walk-Jump Sampling*, 2024
- ICLR MLDD Oral Presentation Spotlight, *Multi-segment preserving sampling for deep manifold sampler*, 2023
- ISCB 3DSig Oral Presentation Spotlight, *Graph Embeddings for Protein Structural Comparison*, 2021

- CSCW Best Paper honorable mention, *Efficient crowd exploration of large networks: The case of causal attribution*, 2018
- Russian Language Award, Outstanding Achievement in First-Year Russian, University of Vermont, 2015

INVITED TALKS

- *Multi-segment preserving sampling for deep manifold sampler*, ML for Protein Engineering Seminar Series, Nov 2022

PROFESSIONAL ACTIVITIES & SERVICE

- Peer reviewer for NeurIPS, ICML, and ICLR
- Undergraduate student mentor at New York University

SKILLS

Programming	Python, BASH
Tools & Frameworks	Git, Standard ML libraries (e.g., PyTorch), Docker, LaTeX, PyMOL
Languages	English (fluent), Spanish (basic), Russian (basic)

Last updated: August 25, 2026