Benjamin Feuer

Education

New York University

New York Plant Control of the Control of the

M.S. / Ph.D. in Computer Science Sept. 2021–May 2025

Email: bf996@nyu.edu

Website: penfever.github.io

Advised by Dr. Chinmay Hegde

Columbia UniversityNew York, NYM.F.A. in Film ProductionSept. 2008–Dec. 2012

Wesleyan University

B.A. in Film History

Middletown, CT

Sept. 2002–May 2006

Selected Publications

• B Feuer, C Hegde. WildChat-50M: A Deep Dive into the Role of Synthetic Data in Post-Training. ICML, 2025.

- B Feuer, M Goldblum, T Datta, S Nambiar, R Besaleli, S Dooley, M Cembalest, J P Dickerson. <u>Style Outweighs Substance:</u> Failure Modes of LLM Judges in Alignment Benchmarking. *ICLR*, 2025.
- C White, S Dooley, M Roberts, A Pal, B Feuer, S Jain, R Shwartz-Ziv, N Jain, K Saifullah, S Naidu, C Hegde, Y LeCun, T Goldstein, W Neiswanger, M Goldblum. <u>LiveBench: A Challenging, Contamination-Free LLM Benchmark.</u> *ICLR (Spotlight)*, 2025. In OpenAI o3-mini T. Report, Phi-4 T. Report, Gemini 2.0 T. Report, Owen 3 T. Report, etc.
- B Feuer, C Yang, Z Jubery, Z K. Deng, A Nakkab, M Z Hasan, S Chiranjeevi, K O Marshall, N Baishnab, A K Singh, A Singh, S Sarkar, N Merchant, C Hegde, B Ganapathysubramanian. <u>BioTrove: A Large Multimodal Dataset Enabling AI for Biodiversity</u>. NeurIPS (Spotlight), 2024
- B Feuer, R T Schirrmeister, V Cherepanova, C Hegde, F Hutter, M Goldblum, N Cohen, C White. <u>TuneTables: Context Optimization for Scalable Prior-Data Fitted Networks</u>. *NeurIPS*, 2024

Recent Publications

- K Arabi, B Feuer, R T Witter, C Hegde, N Cohen. <u>Hidden in the Noise: Two-Stage Robust Watermarking for Images</u>. *ICLR*, 2025
- B Feuer, J Xu, N Cohen, P Yubeaton, G Mittal, C Hegde. <u>SELECT: Large-Scale Benchmark of Data Curation Strategies for Image Recognition</u>. NeurIPS, 2024
- B Feuer, Y Liu, C Hegde, J Freire. <u>ArcheType: A Novel Framework for Open-Source Column Type Annotation using Large Language Models</u>. VLDB, 2024
- D McElfresh, S Khandagale, J Valverde, V Prasad C, B Feuer, C Hegde, G Ramakrishnan, M Goldblum, C White. When Do Neural Nets Outperform Boosted Trees on Tabular Data? NeurIPS, 2023
- B Feuer, A Joshi, M Pham, C Hegde. <u>Distributionally Robust Classification on a Data Budget.</u> *Transactions in Machine Learning Research*, 2023

• B Feuer, A Joshi, M Cho, K Jani, S Chiranjeevi, ZK Deng, A Balu, A K Singh, S Sarkar, N Merchant, A Singh, B Ganapathysubramanian, C Hegde. Zero-Shot Insect Detection via Weak Language Supervision, The Plant Phenome Journal (2024), AAAI Workshop on AI for Agriculture and Food Sciences (AIAFS), 2023 (Oral)

Awards, Honors, Invited Talks

Annus Mirabilis: A Year of Growth In LLM Research, David Bau Lab, Northeastern University	2025
Annus Mirabilis: A Year of Growth In LLM Research, Jetbrains Research	2025
Annus Mirabilis: A Year of Growth In LLM Research, Tom Goldstein Lab, UMD	2025
Pearl Brownstein Award (Best Doctoral Research), New York University	2025
Style and Substance: Failure Modes of LLM Judges, NLP, Penn State University	2024
Mitigating Biodiversity Loss with Frontier AI, Machine Learning, Rensselaer Polytechnic Institute	2024
Top Reviewer Award, NeurIPS	2024
Style and Substance: Failure Modes of LLM Judges, AI In Research Sem., New York University	2024
The Deborah M. Rosenthal Award (Best CS Qualifying Exam), New York University	2024
Big LLMs for Big Data, Big Data, New York University	2024
SOE Scholarship, Summer Scholars, New York University	2021

Teaching

• Recitation Leader at New York University

Discrete Mathematics

Fall 2021, Spring 2022

- Lead lectures for groups of 20-25 students, prepared and distributed course materials, graded homework and exams
 Number Theory

 Spring 2022
- Adjunct Faculty at Ramapo College, Montclair State University

World Cinema Analysis

Fall 2018 - Spring 2021

- Created syllabus, taught undergraduates, judged student film competitions, career and academic advising

Service

Workshop Organizer

Lead Organizer: Dataworld @ ICML 2025, DCVLR @ NeurIPS 2025

Reviewer

ICML, ICLR, NeurIPS, TMLR: 2024, 2025

Professional Experience

Oumi.AI

Research Fellow

Jan - Sept 2025

- Series A Startup; building the 'Linux of artificial intelligence'; #1 trending GitHub repo, 6500+ stars
- Conducted original research, developed code for model post-training

State of Place
Research Consultant
Boston, MA (remote)
Sept 2024 - Apr 2025

- Lead technical team on a Phase 1 SBIR grant with US DOT
- Developed novel, scalable computer vision algorithms for built environment features and smart cities

Arthur
New York, NY (hybrid)
Research Fellow
May - Aug 2024

- Series B Startup; AI observability platform, including the world's first LLM firewall
- Conducted original research into failure modes of LLM-judges for alignment benchmarking

Covalent San Francisco, CA (remote)
Research Intern May - Aug 2021

- Y-Combinator & Wharton-backed startup focused on group matching, similar to <u>Lunchelub</u>
- Improved core app functionality for group and individual matching using BERT transformer embeddings

Arizona State University
Research Consultant
Tempe, AZ
May - Aug 2020

- Implemented an algorithm to compute the NDVI index of seedlings based on infrared imagery from a Raspberry Pi
- Project was awarded an \$1.6 Mn SBIR grant from the Dept. of Energy in 2021