Herbie Bradley

✉ mail@herbiebradley.com   🔗 herbiebradley.com
🐦 @herbiebradley   ☺ github.com/herbiebradley

EDUCATION

2021 -	**PhD Candidate**	**University of Cambridge**
	Supervisor: Prof. Samuel Albanie, Department of Engineering	
2021	**MRes AI for Environmental Risk**	**University of Cambridge**
	Dissertation: *Neural network emulation of the chemistry-climate system*	
	Supervisors: Professor Alex Archibald, Dr Luke Abraham, Dr Doug McNeall	
2020	**MSc Data Analytics**	**University of Warwick**
	Distinction │ Dissertation: *Dataset Distillation with Deep Learning*	
2019	**BSc (Hons) Computer Science and Mathematics**	**Keele University**
	First Class │ Dissertation: *An Exploration of Style Transfer with Deep Learning*	

PUBLICATIONS & PREPRINTS

- **H. Bradley**, H. Fan, T. Galanos, R. Zhou, D. Scott, J. Lehman. 'The OpenELM Library: Leveraging Progress in Language Models for Novel Evolutionary Algorithms'., 2023. To appear in: *Genetic Programming Theory and Practice XX*.

- **Under review.** J. Suarez, D. Bloomin, K. Choe, H. Li, R. Sullivan, N. Kanna, D. Scott, R. Shuman, **H. Bradley**, L. Castricato, P. Isola, K. You, Y. Jiang, Q. Li, J. Chen, X. Zhu. 'Neural MMO 2.0: A Massively Multi-task Addition to Massively Multi-agent Learning'., 2023.

- **Under review.** L. Ruis, A. Findeis, **H. Bradley**, H. Rahmani, K. Choe, E. Grefenstette, T. Rocktäschel. 'Do LLMs selectively encode the goal of an agent's reach?'., 2023.

- A. Chan, **H. Bradley**, N. Rajkumar. 'Reclaiming the Digital Commons: A Public Data Trust for Training Data'., 2023. In: *AIES '23: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*.

- **arXiv preprint.** E. Meyerson, M. J. Nelson, **H. Bradley**, A. Moradi, A. K. Hoover, J. Lehman. 'Language Model Crossover: Variation through Few-Shot Prompting'., 2023.

- S. Biderman, H. Schoelkopf, Q. Anthony, **H. Bradley**, K. O'Brien, E. Hallahan, M. A. Khan, S. Purohit, U. S. Prashanth, E. Raff, A. Skowron, L. Sutawika, O. van der Wal. 'Pythia: A Suite for Analyzing Large Language Models'., 2023. In: *Proceedings of the 40th International Conference on Machine Learning* (Oral).

- **H. Bradley**, L. Abraham, P. Nowack, D. McNeall. 'Transformers for Fast Emulation of Atmospheric Chemistry Box Models'., 2022. In: *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*.

- J. Phang, **H. Bradley**, L. Gao, L. Castricato, S. Biderman. 'EleutherAI: Going Beyond "Open Science" to "Science in the Open"'., 2022. In: *NeurIPS 2022 Workshop on Broadening Research Collaborations*. (Nominated for Best Paper).

Experience

June 2023 -	**Research Scientist Intern**	**Stability AI**
	Working on reinforcement learning from human feedback (RLHF) for large language models.	
Nov. 2022 -	**Research Lead**	**Cambridge AI Safety Hub**
	Supervising multiple research projects on interpretability of large language models and reward modelling, each consisting of teams of 3-5 students from varied backgrounds.	
Sept. 2022 -	**Research Scientist**	**EleutherAI & CarperAI**
	Managing a team of 5-10 volunteer researchers from undergrads to postdocs working on research in the intersection of evolutionary algorithms & large language models. Additionally, worked on AI safety projects across interpretability and AI governance.	
2021 - 2022	**Programme Lead**	**Cambridge Existential Risks Initiative**
	Organised climate change portion of summer research internship, including application design, candidate selection, interviews, project selection, and mentorship.	

Teaching

Guest Lecturer, University of Cambridge
R255 Advanced Topics In Machine Learning—AI Safety (Spring '23)

Supervisions, University of Cambridge
Advanced Machine Learning & Deep Neural Networks (Spring '22 & '23)
Climate Change Mitigation (Autumn '21 & '22)

Effective Climate Change Mitigation Autumn '21 & Spring '22
Founder & Organiser, student-led 8 week seminar series with 60 participants in collaboration with Cambridge Zero and the Cambridge Existential Risks Initiative.

Awards

2020 - 2024	EPSRC CDT (MRes + PhD) Studentship (£120,000)
2019 - 2020	University of Warwick Taught Masters Scholarship (£10,000)

British national and UK resident.