

Hailey Schoelkopf

hailey.ruth.schoelkopf@gmail.com • [Website](#) • [Github](#) • [Google Scholar](#) • [Semantic Scholar](#) • [LinkedIn](#)

EDUCATION

Yale University | New Haven, CT

August 2019 – May 2023

B.S. Computer Science & Mathematics, Distinction in Major

Cumulative GPA: 3.80

EXPERIENCE

Research Scientist | EleutherAI

Sep 2022 - Present

- Co-lead maintainer of [LM Eval Harness](#)—widely adopted as the open standard for few-shot LM evaluation.
- Co-led creation and publication of [Pythia LLM Suite](#) (awarded ICML 2023 Oral presentation), pretrained the models on up to 256 accelerators.
- Informed new practitioners ([Transformer Math 101](#), [EleutherAI Cookbook](#), [FM Dev Cheatsheet](#)) and assisted in guiding [policy](#) and [journalism](#).
- Led large-scale training portions of LLM research projects like [Llemma](#) and contributed to EleutherAI's [GPT-NeoX](#) library for efficient large-scale distributed training.

Independent Researcher | BigScience Research Workshop

Mar 2022 - Nov 2022

- Contributed to research on multilingual instruction tuning ([BLOOMZ](#)), [evaluating diverse multilingual prompts](#) and exploring impact of adaptation to Prefix Language Modeling objective.
- Performed comparison of many then-SOTA parameter-efficient finetuning (PEFT) approaches for efficient multilingual adaptation ([BLOOM+1](#)), reimplementing and reproducing many methods from literature.

SKILLS & TECHNICAL TOOLS

- Python (proficient), R, C (light experience)
- PyTorch, OpenAI Triton
- Familiarity and experience with a majority of existing LLM distributed training frameworks, including **GPT-NeoX (contributor)**, Megatron-DeepSpeed, Megatron-LM, T5x, MosaicML Composer, NVIDIA NeMo-Megatron
- Experience performing distributed training of ML models on hundreds of machines using GPU clusters (SLURM)

SELECTED TALKS

- [MIT FutureTech, Workshop on AI Scaling and its Implications](#) (October 2023). Invited Talk on *Scaling, Parallelism, and Hardware: What is the future of LLMs?*.
- [ODSC East 2024](#) (April 2024). Talk on *Practical Challenges in LM Evaluation*.
- [Mastering LLMs](#) (June 2024). Invited Talk on *A Deep Dive on LLM Evaluation*.
- [ICML 24 Tutorial](#) (July 2024). Tutorial on *Challenges in Language Model Evaluations*.
- [ICML 24, Workshop on Efficient Systems for Foundation Models](#) (July 2024). Panelist, *Data and Architecture Trends Across Industry and Open Communities*.
- [Princeton PLI, Workshop on Useful and Reliable AI Agents](#) (August 2024). Panelist, *Evaluating agents for real-world use*.

SELECTED PUBLICATIONS

For a full list of publications, see [my Google Scholar](#) or [Semantic Scholar](#) pages.

- Biderman, S.*, **Schoelkopf, H.***, Anthony, Q., Bradley, H., O'Brien, K., Hallahan, E., Aflah Khan, M., Purohit, S., Sai Prashanth, U., Raff, E., et al. 2023. Pythia: A suite for analyzing large language models across training and scaling. *arXiv preprint arXiv:2304.01373*. (Oral ICML 2023)
- Biderman, S.*, **Schoelkopf, H.***, Sutawika, L.*, et al. 2024. Lessons from the Trenches on Reproducible Evaluation of Language Models. *arXiv preprint arXiv:2405.14782*.

- Azerbayev, Z., **Schoelkopf, H.** et al. 2023. Llemma: an Open Language Model for Mathematics. *arXiv preprint arXiv:2310.10631* (2023). **(ICLR 2024)**
- Biderman, S.*, **Schoelkopf, H.**, Miranda, B., Mukobi, G., Madan, V., Ibrahim, A., Bradley, H., Biderman, S., Koyejo, S. 2024. Why Has Predicting Downstream Capabilities of Frontier AI Models with Scale Remained Elusive?. *arXiv preprint arXiv:2406.04391*.
- Anthony, Q., Biderman, S., and **Schoelkopf, H.** 2023. Transformer Math 101. Blog post, <https://blog.eleuther.ai/transformer-math/> .
- Biderman, S., Sai Prashanth, U., Sutawika, L., **Schoelkopf, H.**, Anthony, Q., Purohit, S., and Raff, E. 2023. Emergent and Predictable Memorization in Large Language Models. *arXiv preprint arXiv:2304.11158*. **(NeurIPS 2023)**
- Anthony, Q., Biderman, S., and **Schoelkopf, H.** 2023. Transformer Math 101. Blog post, <https://blog.eleuther.ai/transformer-math/> .
- Biderman, S., Sai Prashanth, U., Sutawika, L., **Schoelkopf, H.**, Anthony, Q., Purohit, S., and Raff, E. 2023. Emergent and Predictable Memorization in Large Language Models. *arXiv preprint arXiv:2304.11158*. **(NeurIPS 2023)**