Hailey Schoelkopf

hailey.ruth.schoelkopf@gmail.com • 203-215-0079 • Github • Google Scholar • Semantic Scholar • LinkedIn

EDUCATION

Yale University | New Haven, CT August 2019 – May 2023

B.S. - Mathematics + Computer Science, Distinction in Major *Cumulative GPA: 3.80*

EXPERIENCE

Research Scientist | EleutherAI *Sep 2022 - Present*

- Co-lead on LM Evaluation Harness, a library that is widely adopted as the open standard for LLM evaluation.
- Co-led the creation and publication of the Pythia LLM Suite (awarded ICML 2023 Oral presentation), and pretrained the models on up to hundreds of accelerators.
- Communicated technical details, informing new practitioners (Transformer Math 101, EleutherAI Cookbook, FM Dev Cheatsheet) and guiding policy and journalism.
- Led or contributed to many research projects, on topics ranging from model interpretability, evaluation, and more.
- Led systems optimization and large-scale training portions of research projects including Llemma.

Independent Researcher | BigScience Research Workshop *Mar 2022 - Nov 2022*

- Worked part-time as an independent researcher at the BigScience collaboration on language model evaluation pipelines and new model architectures, and contributed to research in multi-task finetuning of language models and multilingual adaptation of language models.

Research Assistant | LILY Lab, Yale University *Aug 2020 - Aug 2022*

- Worked under Dragomir Radev, performing NLP research in topics including language modeling, code generation, text summarization, and information retrieval.
- Implemented deep learning research code, assisted with engineering and upkeep on a constrained GPU cluster, research paper writing, and mentored other undergraduate students.

SKILLS & TECHNICAL TOOLS

- Python (proficient), R, C/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 / Gen AI models on hundreds of machines using GPU clusters

SPEAKING OPPORTUNITIES

- **MIT FutureTech, Workshop on AI Scaling and its Implications.** (October 2023) Invited Talk on *Scaling, Parallelism, and Hardware: What is the future of LLMs?*
- **The AI Epiphany** (November 2023) Invited Talk on *Building Research Infrastructure for the LLM Era*
- **ODSC East 2024** (April 2024) **Tutorial** on *Practical Challenges in LM Evaluation*
- **Mastering LLMs** (June 2024) Invited Talk on *A Deep Dive on LLM Evaluation*
- **ICML 24** (July 2024) **Tutorial** on *Challenges in Language Model Evaluations*

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.*

- 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)**
- Azerbayev, Z., **Schoelkopf, H.** et al. "Llemma: An open language model for mathematics." *arXiv preprint arXiv:2310.10631* (2023). **(ICLR 2024)**
- Muennighoff, N., Wang, T., Sutawika, L., Roberts, A., Biderman, S., Scao, T.L., Bari, M.S., Shen, S., Yong, Z.X., **Schoelkopf, H**., Tang, X., et al. 2022. Crosslingual Generalization through Multitask Finetuning. *arXiv preprint arXiv:2211.01786*.