Michael Hu

+1 (678) 908 2031
✉ michael.hu@nyu.edu
michahu.github.io

Education

Fall 2022– **PhD in Data Science**, *New York University*
- Research interests: NLP, training data, meta-learning.
- Advisors: Kyunghyun Cho, Tal Linzen.
- GPA: 4.0 / 4.0.

Fall 2017–
Spring 2021 **BSE in Computer Science**, *Princeton University*
- Minors: Statistics and Machine Learning, Robotics and Intelligent Systems.
- Advisors: Karthik Narasimhan, Tom Griffiths.
- GPA: 3.91 / 4.0. Summa Cum Laude.

Selected Publications

* indicates equal contribution

[1] Michael Y. Hu, Jackson Petty, Chuan Shi, William Merrill, and Tal Linzen. Between Circuits and Chomsky: Pre-pretraining on formal languages imparts linguistic biases. *Preprint*, 2025.

[2] Mayee F. Chen*, Michael Y. Hu*, Nicholas Lourie, Kyunghyun Cho, and Christopher Re. Aioli: A unified optimization framework for language model data mixing. *ICLR*, 2025.

[3] Ethan Wilcox, Michael Y. Hu, Aaron Mueller, Tal Linzen, Alex Warstadt, Leshem Choshen, Chengxu Zhuang, Ryan Cotterell, and Adina Williams. Bigger is not always better: The importance of human-scale language modeling for psycholinguistics. *Preprint*, 2024.

[4] Michael Y. Hu, Angelica Chen, Naomi Saphra, and Kyunghyun Cho. Latent State Models of Training Dynamics. *Transactions on Machine Learning Research*, 2023.

[5] Sreejan Kumar, Carlos G. Correa, Ishita Dasgupta, Raja Marjieh, Michael Y. Hu, Robert D. Hawkins, Nathaniel D. Daw, Jonathan D. Cohen, Karthik Narasimhan, and Thomas L. Griffiths. Using natural language and program abstractions to instill human inductive biases in machines. *NeurIPS*, 2022.

[6] Tsung-Yen Yang*, Michael Hu*, Yinlam Chow, Peter J. Ramadge, and Karthik Narasimhan. Safe Reinforcement Learning with Natural Language Constraints. *NeurIPS*, 2021.

Awards

2022–2027 **NSF Graduate Research Fellowship** ($45,000 per year for 3 years)

2022 **Outstanding Paper Award**, NeurIPS 2022 (top 0.15% of submissions)
Using Natural Language and Program Abstractions to Instill Human Inductive Biases in Machines

2021 **Spotlight**, NeurIPS 2021 (top 2.9% of submissions)
Safe Reinforcement Learning with Natural Language Constraints

2021 Outstanding Computer Science Senior Thesis Prize ($600)

2020 Princeton Center for Statistics and Machine Learning Summer Research Award ($4,000)

Experience

May 2025–
Aug 2025 **Incoming Research Intern**, *Microsoft Semantic Machines*.
- Emergent communication in multi-agent reinforcement learning.

Sept 2021–
Sept 2022 **Software Engineer**, *Yobi*. First employee.
- Implemented behavior prediction models and pipelines at scale (machine learning for advertising).
- Drafted consumption and revenue projections that helped Yobi raise 10+ million for Series A.