XIAOYAN(ELENA) BAI

Chicago, Illinois United States

OBJECTIVE ✉ smallyan@uchicago.edu   🖥 elena-baixy.github.io   ⌨ Elena-Baixy

I am a second year PhD student in CS from University of Chicago, seeking research intern in Summer 2026. My research interest is in understanding language models and responsible AI. Availability: June 2026 - September 2026.

EDUCATION

University of Chicago   **September 2024 – Present**
PhD student in Computer Science   *Illinois, United States*

University of Michigan, Ann Arbor   **September 2022 – April 2024**
B.S.E in Computer Science, Minor in Art and Design   *Michigan, United States*

Shanghai Jiao Tong University   **September 2020 – August 2024**
B.S. in Electronic and Computer Engineering   *Shanghai,China*

SKILLS

Programming Skill: Python, PyTorch, HTML, React, Unity, Java, Matlab, C/C++

PUBLICATIONS

Know Thyself? On the Incapability and Implications of AI Self-Recognition *(preprint)*
Xiaoyan Bai, Aryan Shrivastava, Ari Holtzman, Chenhao Tan

Why Can't Transformers Learn Multiplication? Reverse-Engineering Reveals Long-Range Dependency Pitfalls *(preprint)*
Xiaoyan Bai*, Itamar Pres*, Yuntian Deng, Chenhao Tan, Stuart Shieber, Fernanda Viégas, Martin Wattenberg, Andrew Lee

Concept Incongruence: An Exploration of Time and Death in Role Playing. *(NeurIPS 2025)*
Xiaoyan Bai, Ike Peng*, Aditya Singh*, Chenhao Tan

Learn To be Efficient: Build Structured Sparsity in Large Language Models. *(NeurIPS 2024)*, **Spotlight**
Haizhong Zheng, **Xiaoyan Bai**, Xueshen Liu, Z. Morley Mao, Beidi Chen, Fan Lai, Atul Prakash

A Mechanistic Understanding of Alignment Algorithms: A Case Study on DPO and Toxicity. *(ICML 2024)*, **Oral (Top 1.5% of submissions)**
Andrew Lee, **Xiaoyan Bai**, Itamar Pres, Martin Wattenberg, Jonathan K. Kummerfeld, Rada Mihalcea

WORKING EXPERIENCE

Teaching Assistant | *University of Chicago*   **October 2024 – December 2024**

Grader for Foundations of CS | *University of Michigan*   **September 2023 – December 2023**

Teaching Assistant | *Massachusetts Institute of Technology Lincoln Lab*   **June 2023 – August 2023**
- Create course materials for combining modern methods in machine learning and game-like modeling to quantitatively analyze socially relevant technology and policy questions

AI Training Assistant Internship | *Emogent*   **December 2021 – January 2022**
- Preprocess and analyze the training data by Python for the interactive AI Irene, which serves in the museum as a guide.

SERVICE

NeurIPS | *reviewer*   **2025**

ICLR DeLTa and SCOPE workshop | *reviewer*   **2025**

ACL ARR | *reviewer*   **2024**

HONORS

University of Michigan Dean's Honor List   **Winter 2023 - Winter 2024**
China Collegiate Algorithm Design and Programming Challenge Contest | *Bronze Medal*   **March 2021**
Undergraduate Scholarship | *Awarded for excellent academic performances*   **September 2020 - August 2021**