Matthew Shu

☎ 301-803-0096 • ✉ matthew.shu@yale.edu • 🖥 www.matthewshu.com

⊙ Nonameentered • in linkedin.com/in/mattshu/ • Olney, MD

Education

Yale University: Statistics & Data Science (Joint Bachelor's and Master's), Education Studies (Minor)   *2021 - May 2025*
Courses: Measuring Culture, Trustworthy Deep Learning, Digital Platforms, Algorithms & Societal Implications, AI Foundation Models, Data Analysis. *Thesis (Planned)*: How Artificial Intelligence Summaries of Historical Events Influence Political Opinions
Awards: Angels Wings Scholar ($10,000), Hadden Scholar ($16,000), Linck Fellowship ($2500), Cohere Research Grant ($1000)

Experiences

Yale University   **Survey Design, LLMs**
Sociology of Machine Learning Research Intern, under Prof. Daniel Karell   *December 2023 - Present*
- Researching how LLM-generated summary texts affect people's learning experiences on the internet (Pubs. 4, thesis)
- Scraped and evaluated 1000+ academic peer reviews to track LLM use in writing and its effects on scientific innovation

Uber   **Tensorflow, PyTorch, Optuna, SparkSQL, LLMs**
Machine Learning Engineer Intern   *May 2023 - August 2023, June 2024 - August 2024*
- 2024 (Eats Home feed): Improved home page recommendations for Eats users with new aggregate and embedding features. Introduced region-based features to feature store available to all UberEats ML models, changes deployed to production homefeed model
- 2023 (Eats Search): Built image-text grocery classifier, piloted LLMs for data quality, optimized hyperparameter tuning on search model

Podsie (Now Part of Teaching Lab)   **React.JS, Ruby on Rails, PostgreSQL, EdTech**
Full-Stack Software Engineer   *September 2022 - August 2024*
- Podsie: First employee, build features to increase engagement and accessibility like points, leaderboards, and GPT-powered translation
- LLM Flashcards (part of $1.5 million NSF study): Architected system to create/evaluate/edit flashcards with AI-human supervision

Human-Computer Interaction Institute at Carnegie Mellon University   **React.JS, Figma**
Research Intern, under Prof. Brad Myers and Amber Horvath   *June 2022 - August 2022*
- Using React.JS, introduced filtering and sorting of relevant notes in Visual Studio Code extension for programmers

University of Maryland at College Park   **NLP, Python, Vue.JS, SQL, Docker**
Research Intern, under Prof. Jordan Boyd-Graber and Shi Feng   *January 2019 - Present*
- Developing KAR³L Flashcards AI app with 500+ users to study how NLP can support long-term learning (Pubs. 2, 3)
- Subreviewer for ACL Rolling Review, the centralized review system for computational linguistics conferences (e.g., EMNLP)
- Investigated bias in question answering systems (Pub. 2). Coordinated 2021 QANTA question answering dataset update

Technaplex Software   **Swift, Fastlane**
iOS Developer, Co-Founder (1 of 2)   *June 2018 - Present*
- Built two iOS apps to support learning: Flashcard Adder (~2.7k downloads) and QuizDB for iOS (~8.5K downloads)

Publications

4. Daniel Karell, **Matthew Shu**, Keitaro Okura, and Thomas Davidson. Artificial Intelligence Summaries of Historical Events Improve Knowledge Compared to Human-Written Summaries. 2024 (*Preprint*).

3. **Matthew Shu**, Nishant Balepur, Shi Feng, and Jordan Boyd-Graber. KARL: Knowledge-Aware Retrieval and Representations Aid Retention and Learning in Students. *Empirical Methods in Natural Language Processing*, 2024.

2. Nishant Balepur, **Matthew Shu**,..., and Jordan Boyd-Graber. A SMART Mnemonic Sounds like 'Glue Tonic': Mixing LLMs with Student Feedback to Make Mnemonic Learning Stick. *Empirical Methods in Natural Language Processing*, 2024.

1. Mao, Andrew, Naveen Raman, **Matthew Shu**,..., and Jordan Boyd-Graber. Eliciting Bias in Question Answering Models through Ambiguity. In Proceedings of the 3rd Workshop on Machine Reading for Question Answering, 92–99, 2021.

Activities & Leadership

Yale Statistics & Data Science Student Advisory Committee   **Logistics**
Committee Member   *April 2022 - Present*
- 24-25: Leading committee of 13 students; formed subgroups for advising resources, community events, and merchandise
- 22-24: Organized events such as game and advising nights to connect the Statistics & Data Science community at Yale

Yale Poorvu Center for Teaching and Learning   **Teaching**
Student Leader, Pedagogical Partner for Statistics & Data Science Senior Capstone Seminar   *August 2022 - June 2024*
- 23-24: Co-led program of 10 pedagogical partners to support faculty in developing and teaching courses across the sciences & humanities
- 22-23: Worked with professor and students in two data science classes to improve learning through classroom observation and interaction

CS50: Intro to Computer Science at Yale   **Python, C, Teaching**
Teaching Assistant   *August 2022 - December 2022*
- Co-taught weekly supplementary lecture to small cohort of students and hosted office hours to help students with assignments

Splash at Yale   **Teaching, Logistics**
Splash Spring 2022 Director, Admin Team, Class Teacher   *September 2021 - Fall 2023*
- One of two directors for Splash Spring 2022, where Yale students teach a one-day class about their interests to 7-12th grade students

Skills

Languages: Human (English/Chinese/Thai), Swift, Python, SQL, R, C, Java, HTML/CSS/Javascript, Typescript, Lisp, Bash

Tools: PostgreSQL, FastAPI, SQLAlchemy, Vue, React, Docker, Traefik, Tidyverse, SpaCy, Pandas, PyTorch, Fastlane, Git