

Bang Nguyen

Ph.D. Candidate, Computer Science & Engineering, University of Notre Dame

bnguyen5@nd.edu · bnguyen5.github.io · [LinkedIn](#) · [Google Scholar](#)

RESEARCH INTERESTS

Natural language processing and generative AI. I build LLM-based applications and evaluate them end to end with simulated interactions and targeted benchmarks, and I use what the evaluation reveals to refine the systems. Applications include education (question generation and assessment) and science (agents that replicate research).

EDUCATION

Ph.D., Computer Science and Engineering

expected May 2027

University of Notre Dame

- Dissertation: *Evaluating and Refining AI Assistants in Education and Research with Language Model Simulation*. Advisor: Dr. Meng Jiang
- Candidacy exam passed Nov 2025. GPA 4.0/4.0

M.S., Computer Science and Engineering

Jan 2026

University of Notre Dame

B.A., Computer Science

2022

The College of Wooster

- Minors: Statistical and Data Sciences, Communication Studies. Summa cum laude, Departmental Honors, Phi Beta Kappa. GPA 4.0/4.0

TEACHING EXPERIENCE

Instructor of Record, CSE 40657 Natural Language Processing

Fall 2026

University of Notre Dame

- Sole instructor for an upper-level undergraduate course (about 15 students; three lectures a week). Designed the full 16-week curriculum, from n-gram language models and neural sequence models to transformers, LLM pre- and post-training, evaluation, interpretability, and responsible and multilingual NLP
- Built five implement-from-scratch Python/PyTorch assignments and a four-milestone team project (proposal, baseline report and presentation, poster, final report); organized a joint poster session with the graduate Large Language Models course
- Ran an anonymous mid-semester survey and responded to it by adding a primer, posting slides and scheduling an extra tutorial; students vote on which four of twelve advanced topics (e.g., retrieval and RAG, LLM agents, responsible LLMs) the course covers in weeks 11–14
- Wrote clear course policies: an AI-use statement required with every project milestone, a structured late-day policy, and accessibility and wellbeing resources; released lecture materials under a CC BY 4.0 license at bnguyen5.github.io/teaching/nlp-fall-2026

Graduate Teaching Assistant, Computer Science and Engineering

2022–2025

University of Notre Dame

- Large Language Models (Fall 2025): designed about 50 quiz questions and organized the end-of-semester project poster session, including poster printing and logistics
- Algorithm Analysis (Fall 2022): held office hours, graded assignments and exams with detailed feedback, wrote and presented assignment solutions
- Secure Software Engineering (Spring 2023): graded assignments with detailed feedback
- Gave instructors feedback on course structure and materials

Teaching Assistant and Course Design Assistant, Introduction to Statistics 2020 – 2022
The College of Wooster (with Dr. Marian Frazier)

- Adapted and revised in-class activities and designed quiz questions; supported students during in-class activities and held independent office hours

Teaching Assistant, Introduction to Data Science; Data Visualization 2020 – 2022
The College of Wooster (with Dr. Jillian Morrison)

- Helped students troubleshoot R and Tableau in in-class activities, assignments and projects; held independent office hours; led a class session independently

RESEARCH EXPERIENCE

Research Assistant, Data Mining Lab 2022 – present
University of Notre Dame

- *AI for education* (June 2022–Nov 2024): NSF RETTTL-funded collaboration with UIUC and Gallaudet University on AI for accessible STEM learning for Deaf and Hard of Hearing students
 - Demonstrated that standard reference-based metrics fail to rank question-generation systems (Findings of EMNLP 2024, first author)
 - Built LLM student simulators that predict how learners respond to test items, enabling large-scale offline evaluation of generated content (QG-SMS, ACL 2025, first author)
- *AI agents for scientific research* (June 2025–present): multi-institution collaboration with the Center for Open Science, Old Dominion University and Penn State, funded by Coefficient Giving
 - Lead developer of an agentic framework that replicates published social-science studies end to end; developed the agent tool layer and sandboxed environments in which agents locate data, write and run analysis code, execute experiments and interpret findings
 - Converted the expert replication workflow into an end-to-end benchmark (ReplicatorBench, KDD 2026, co-first author)
 - On the team’s follow-up robustness benchmark, lead evaluation design, converting human analysts’ peer reviews into stage-wise labels for sub-agent evaluation, and a human annotation study with 15 social scientists; transitioned the agent infrastructure to the new project lead
- *RL fine-tuning with AI feedback* (2025–present): fine-tuning LLM question generators with reinforcement learning using multidimensional LLM-as-judge rewards; investigating how to align inexpensive judge rewards with more faithful simulated-user feedback

INDUSTRY EXPERIENCE

Applied Research Consultant, AMRE May – Jul 2021
The College of Wooster; industry partner: The Goodyear Tire & Rubber Company

- Proposed and built deep learning image-segmentation models for automated quality control, improving accuracy by 48% over the existing rule-based system; Goodyear adopted the results and renewed the project for 2022
- Co-led the full ML cycle with another student: annotated data, fine-tuned and evaluated segmentation models, and built a working prototype; met weekly with Goodyear engineers to align requirements and translate results into actionable insights

UNDERGRADUATE RESEARCH AND DATA EXPERIENCE

Senior Independent Study Thesis, Computer Science Aug 2021 – May 2022
The College of Wooster

- *Queering NLP: A Non-Heteronormative Approach to Quantifying and Investigating Sentiment Bias against LGBTQ+ Identities in Word Embeddings*

- Quantified sentiment bias against LGBTQ+ identities in word embeddings; received the Critical Digital Engagement Award (top 3 of 175 theses) and presented a poster at oSTEM 2021

Statistical and Data Sciences Research Consultant, AMRE

May – Aug 2020

The College of Wooster; partner: STEM Success Initiative

- Analyzed institutional course and student data for the STEM Success Initiative to study retention and persistence in STEM fields; presented findings to the initiative
- Worked with several college departments to gather and clean institutional data; ran R and Tableau workshops for the project team

Computer Science Research Assistant, *Anglophone Chile* digital archive

Sep 2019 – May 2020

The College of Wooster

- Co-developed two websites for [AnglophoneChile](#), an international digital archive of 19th-century English-language newspapers from Chile (about 460 issues, 3,900+ pages)
- Evaluated and edited Python NLP pipelines to extract Chile-related news items

Web and New Media Intern, Office of Marketing and Communication

Mar 2019 – May 2022

The College of Wooster

- Maintained the college website (CMS, HTML/CSS); supported redesigns and user research

PUBLICATIONS

* equal contribution

Conference papers

1. **B. Nguyen***, D. Soós*, et al. ReplicatorBench: Benchmarking LLM Agents for Replicability in Social and Behavioral Sciences. *KDD 2026* (AI for Sciences Track). Poster; also an oral presentation at the KDD 2026 Workshop on SciSoc Agents & LLMs.
2. M. Yu, **B. Nguyen**, O. Zino, M. Jiang. Context Selection and Rewriting for Video-based Educational Question Generation. *AAAI 2026* (EAAI Symposium: AI for Education).
3. **B. Nguyen**, T. Du, M. Yu, L. Angrave, M. Jiang. QG-SMS: Enhancing Test Item Analysis via Student Modeling and Simulation. *ACL 2025* (Main). Poster.
4. **B. Nguyen**, M. Yu, Y. Huang, M. Jiang. Reference-based Metrics Disprove Themselves in Question Generation. *Findings of EMNLP 2024*. Poster.

Workshop papers and posters

1. H. Dang*, **B. Nguyen***, N. Ziems, M. Jiang. Embedding Mental Health Discourse for Community Recommendation. *CODI Workshop @ ACL 2023*.
2. **B. Nguyen**. Queering NLP: A Non-Heteronormative Approach to Quantifying and Investigating Sentiment Bias against LGBTQ+ Identities in Word Embeddings. Poster, *oSTEM 2021*.

TALKS AND PRESENTATIONS

- ReplicatorBench: An End-to-End Benchmark for AI Agents in Research Replication. Talk (with D. Soós), [FORRT AI in Metascience Online Conference](#), Sep 2026
- ReplicatorBench: Benchmarking LLM Agents for Replicability in Social and Behavioral Sciences. Oral presentation, KDD 2026 Workshop on SciSoc Agents & LLMs, 2026
- QG-SMS: Enhancing Test Item Analysis via Student Modeling and Simulation. Talk, oSTEM Annual Conference, 2025
- Classroom Simulations for Large Language Model Applications. Invited talk for K–12 teachers, Research Experiences for Teachers (RET) program, University of Notre Dame, June 2024. Presented the research and advised teachers on using AI to support their students

- Enhancing Curriculum Design through Large Language Model (LLM) Simulations of Student’s Learning Effectiveness. Trustworthy AI Lab for Education (TALE) Summit, Lucy Family Institute for Data & Society, University of Notre Dame, Nov 2023
- Paper presentations at lab meetings (Best Presentation Awards): “Answering Correctly or Realistically? Character Knowledge in LLM Role-playing” (Fall 2025); “Self-Supervised Bootstrapping of Action-Predictive Embodied Reasoning” (Spring 2026)

HONORS AND AWARDS

- Best Presentation Award (top 2 of 10), DM2 Lab, Notre Dame, Fall 2025, Spring 2026
- Named one of the Best Reviewers, KDD 2026 AI for Sciences Track, 2026
- Researcher Access Program, OpenAI, 2025
- 3rd of 20 team presentations, course project presentations (“Predicting Movie Genres Using Character Conversations”), Data Science course, Notre Dame, Fall 2022
- Critical Digital Engagement Award for Senior Thesis (top 3 of 175), College of Wooster, 2022
- Phi Beta Kappa, College of Wooster, 2022
- Dewald Recognition for students with a cumulative GPA of 3.95 or higher, College of Wooster, 2021
- oSTEM Hackathon Winner, Out in STEM Conference, 2021

Grants and travel awards

- oSTEM Collegiate Grant, to present at the oSTEM Annual Conference, 2025
- Graduate Student Government Conference Presentation Grant, University of Notre Dame, Jul 2025, Jul 2026
- Notebaert Professional Development Fund, University of Notre Dame, Sep 2025

MENTORING

- **Ethan Sotka**, undergraduate researcher, University of Notre Dame: AI agents that replicate graphs and other visual results from social-science studies, Fall 2025
- **Olivia Zino**, undergraduate researcher, University of Notre Dame: educational question generation (co-mentored with another PhD student); co-author on an AAAI 2026 paper, Fall 2024 – Fall 2025
- **Tingting Du**, visiting undergraduate, Notre Dame International Summer Undergraduate Research Experience (iSURE): student modeling for test-item analysis; second author on QG-SMS (ACL 2025), Summer 2024
- Peer mentor for students in introductory statistics and data science, STEM Zone, College of Wooster, 2020 – 2022

SERVICE AND LEADERSHIP

- **Lab manager**, Data Mining Lab (15 members), University of Notre Dame, Fall 2024 – Fall 2025
- **Organizer**, Midwest Speech and Language Days (MSLD) 2025
- **Reviewer**: ARR (Jan, Mar, May 2026; Feb, Oct 2025), KDD AI4Sciences, ACM CSUR 2026, IEEE TKDE 2023, KnowledgeNLP workshops (ACL 2024, KDD 2023)
- **Creative and Performance Director**, Chasing Dance Team, College of Wooster, 2019 – 2022: directed 5+ campus performances and led team choreography

SKILLS

Languages: Python (proficient), R, C++, HTML/CSS.

ML / LLMs: PyTorch, Hugging Face; fine-tuning and alignment of language models (e.g., SFT, chain-of-thought fine-tuning, DPO, RL with AI feedback, PEFT, token classification); LLM agents and tool use; evaluation design and benchmarking.

Infrastructure and data: Docker, sandboxed experiment environments, GPU training on HPC clusters (Notre Dame CRC); Tableau.