Geetanjali Bihani

bihani.geet@gmail.com | Website | Github | Google Scholar

Summary

Ph.D. with 7+ years of experience developing machine learning and deep learning frameworks, including 5+ years specializing in Natural Language Processing (NLP). Published in top AI/NLP venues (COLING, AAAI, NAACL), with research interests in enhancing language model reasoning by addressing challenges in resolving linguistic ambiguity and enhancing generalization. Focused on developing reliable, high-impact AI solutions.

Education

2020 – 2025 **Ph.D. in Computer and Information Technology**
Purdue University, West Lafayette, USA
Dissertation: *On the Reasoning Capabilities of Language Models for Detecting Grooming and Coercive Discourse*

2018 – 2020 **M.S. in Computer and Information Technology**
Purdue University, West Lafayette, USA
Thesis: *Longitudinal Comparison of Word Associations in Shallow Word Embeddings*

2012 – 2016 **B.Tech. in Electronics and Communications Engineering**
Delhi Technological University, Delhi, India

Industry Experience

Summer 2023 **Research Intern – Microsoft Research** (*Redmond, WA*)
Mentors: Sujay Jauhar & Milenko Drinic
Team: Knowledge Technologies and Intelligent Experiences (KTX)
Project: Automatic Induction of Interpretable Document Templates using Large Language Models
- Developed a real-time document clustering framework leveraging LLM embeddings and responses for dynamic assignments, improving automatic clustering and operational efficiency.
- Implemented error-handling and retry mechanisms, integrating LLM reasoning to enhance clustering consistency and fault tolerance.

Summer 2022 **Research Intern – Microsoft Research** (*Redmond, WA* [*Hybrid*])
Mentors: Sujay Jauhar, Bahareh Sarrafzadeh & Milenko Drinic
Team: Knowledge Technologies and Intelligent Experiences (KTX)
Project: Automatic Template Discovery using Language Model Embeddings
- Developed a structure-aware document template understanding model incorporating attention mechanisms, positional encoding, and language model (LM) embeddings for structural annotations.
- Achieved an 18% improvement in identifying document template subcategories, surpassing previous text-only and text+structure neural baselines.

Summer 2019 **Entrepreneurial Lead – NSF I-Corps** (*Chicago, IL*)
- Developed a time series ML-based decision prediction system for solar energy, optimizing cost models and revenue streams.
- Conducted customer discovery with 100+ stakeholders in the NSF I-Corps program to validate and refine predictive models, ensuring alignment with market needs.

2016 – 2018 **Data Scientist – Transorg Analytics** (*Delhi, India*)
- Developed and deployed machine learning models, improving demand forecasting and targeted marketing by 20%, optimizing resource allocation and sales performance.

Technical Skills

Programming: Python, SQL, R, Git, Linux
Libraries: PyTorch, TensorFlow, DeepSpeed, LangChain, HuggingFace, Scikit-learn, NLTK, Pandas, NumPy, Keras
NLP Frameworks: LLMs, In-context learning, Supervised fine-tuning, RAG, Transfer Learning, Model Evaluation
Visualization: Matplotlib, Seaborn, Plotly, Procreate, W&B
Natural Languages: English (fluent), Hindi (native)

Selected Publications

- Damin Zhang, Yi Zhang, **Geetanjali Bihani**, and Julia Rayz. Hire Me or Not? Examining Language Model's Behavior with Occupation Attributes. - In *Proceedings of the 31st International Conference on Computational Linguistics (COLING 2025), pages 7891–7911. Association for Computational Linguistics*.

- **Geetanjali Bihani** and Julia Rayz. The Reliability Paradox: Exploring How Shortcut Learning Undermines Language Model Calibration. - In *Proceedings of the 58th Hawaii International Conference on System Sciences (HICSS 2025), pages 856–865*.

- **Geetanjali Bihani** and Julia Taylor Rayz. A Fuzzy Evaluation of Sentence Encoders on Grooming Risk Classification. In *Proceedings of the NAFIPS International Conference on Fuzzy Systems, Soft Computing, and Explainable AI (NAFIPS 2024)*.**Received Outstanding Student Paper Award.**

- **Geetanjali Bihani**, Julia Taylor Rayz. Calibration Error Estimation Using Fuzzy Binning. In *Cohen, K., Ernest, N., Bede, B., Kreinovich, V. (eds) Fuzzy Information Processing 2023. NAFIPS 2023. Lecture Notes in Networks and Systems, vol 751. Springer, Cham.* **Honorable Mention for Best Student Paper.**

- **Geetanjali Bihani**. Interpretable Privacy Preservation of Text Representations Using Vector Steganography. In *Proceedings of the AAAI Conference on Artificial Intelligence, 36(11), 12872-12873.*

- **Geetanjali Bihani**, Julia Taylor Rayz. Low Anisotropy Sense Retrofitting (LASeR) : Towards Isotropic and Sense Enriched Representations. In *Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, pages 81–95. Association for Computational Linguistics.*

- **Geetanjali Bihani**, Julia Taylor Rayz. Model Choices Influence Attributive Word Associations: A Semi-supervised Analysis of Static Word Embeddings. *Proceedings of the 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*.

Selected Talks

Spring 2025	Bridging the Gap: Advancing AI to Detect Covert Harms, *HICSS-58*
Fall 2023	Automatic Induction of Interpretable Document Templates, *Microsoft Research*
Fall 2022	Automatic Template Discovery, *Microsoft Research*
Spring 2022	Language, Representations and Leakage (Invited Talk), *RAISE Lab, Syracuse University*
Spring 2022	Interpretable Privacy Preservation of Text Representations Using Vector Steganography, *AAAI'22*

Honors and Awards

2025	**A.H. Ismail Interdisciplinary Travel Support Grant.** *Purdue University*. Amount: $1000
2025	**Graduate Student Travel Grant.** *PPI Dean's Office*. Amount: $2000
2024	**Travel Grant Award.** *CIT Graduate Student Association.* Amount: $500
2024	**Third Place.** *Holistic Safety and Security Research Symposium, Purdue Polytechnic Institute.*
2023	**Holistic Safety and Security Research Travel Grant.** *Purdue Polytechnic Institute.* Amount: $250
2022	**CIT Research Travel Grant Award.** *Purdue Polytechnic Institute.* Amount: $400
2022	**Best Poster (runner-up).** *Holistic Safety and Security Research Symposium, Purdue Polytechnic Institue.*
2021	**Ross-Lynn Graduate Student Fellowship.** *Purdue Research Foundation.*
2022	**Dean's Graduate Student Travel Grant.** *Purdue Polytechnic Institute.* Amount: $250
2022	**Purdue Graduate Student Government Travel Grant.** *Purdue University.* Amount: $500
2022	**Holistic Safety and Security Research Travel Grant.** *Purdue Polytechnic Institute.* Amount: $250
2021	**Second Place.** *Holistic Safety and Security Research Symposium, Purdue Polytechnic Institute.*
2019	**National Science Foundation (NSF) I-Corps Teams National Award.**

Service

2023 – 2025	**Research & Engagement Officer**, *Purdue CIT Graduate Student Association.*
2024	**Organizer**, *Special Session on Natural Language Uncertainty in the Era of LLMs (NAFIPS).*
2023	**Program Committee Member**, *European Interdisciplinary Cybersecurity Conference.*
2019 – 2020	**Graduate Student Advisor**, *Purdue CIT Student Council.*
2015 – 2016	**Head, Public Relations**, *IEEE Student Branch (Delhi Technological University, India)*
2015 – 2016	**Secretary**, *IEEE Women in Engineering (Delhi Technological University, India)*

References

*Available upon request