Geetanjali Bihani

gbihani@purdue.edu | 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 reputed AI/NLP venues (COLING, AAAI, NAACL), with an interest in studying and enhancing reasoning and inference in language models, towards developing architectures that advance generalization capabilities. Seeking to develop reliable and 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)

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**, Julia Taylor Rayz. Learning Shortcuts: On the Misleading Promise of NLU in Language Models. In *Artificial Intelligence for Design and Process Science, pages 147-158. Cham: Springer Nature Switzerland, 2025.*
- **Geetanjali Bihani**, Tatiana Ringenberg, and Julia Taylor Rayz. Evaluating Language Models on Grooming Risk Estimation Using Fuzzy Theory. In *Proceedings of the NAFIPS International Conference on Fuzzy Systems, Soft Computing, and Explainable AI* (*NAFIPS 2024*).
- **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**, Julia Taylor Rayz. On Information Hiding in Natural Language Systems. In *The International FLAIRS Conference Proceedings, vol. 35. 2022.*
- Huijeong Kim, Sangwoo Ham, Marlen Promann, Hemanth Devarapalli, **Geetanjali Bihani**, Tatiana Ringenberg, Vanessa Kwarteng, Ilias Bilionis, James E.Braun, Julia Taylor Rayz, Leigh Raymond, Torsten Reimer, Panagiota Karava. MySmartE – An eco-feedback and gaming platform to promote energy conserving thermostat-adjustment behaviors in multi-unit residential buildings. In *Building and Environment 221* (*2022*)*: 109252.*
- **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. Fuzzy Classification of Multi-intent Utterances. In *Explainable AI and Other Applications of Fuzzy Techniques: Proceedings of the 2021 Annual Conference of the North American Fuzzy Information Processing Society, NAFIPS 2021, pp. 37-51. Springer International Publishing, 2022.*
- **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*.

Posters

Spring 2024	**Detecting Online Grooming Risk Using Transformers** 2024 Purdue Polytechnic Research Impact Area Student Poster Session
Spring 2023	**Text Augmentation: Improving Classification Accuracy at the Expense of Calibration?** CERIAS Security Symposium
Spring 2022	**Interpretable Privacy Preservation of Text Representations Using Vector Steganography.** The Thirty-Fifth AAAI Conference on Artificial Intelligence
Spring 2022	**Permutation-based Privacy in Text Vectors.** CERIAS Security Symposium
Spring 2021	**Low Anisotropy Sense Retrofitting (LASeR): Towards Isotropic and Sense Enriched Representations.** The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, NAACL

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.**
2015	**2nd Position.** *Renesas Electronics Design Challenge, DTU.*
2010	**Silver certificate (Top 10%).** *In United Kingdom Senior Mathematics Challenge (UKMT).*

Fellowships & Assistantships

Aug 2022 – Present
Research Assistantship funded by the U.S. Department of Justice (Award No. 15PJDP-22-GK-03107-MECP).
Project: LIve-streaming Offender Network-based Chat Analysis Triage Tool (LION-CATT)
Contribution: Authored part of the proposal; Developed bi-encoder models for real-time grooming risk prediction in chat conversations; Evaluated model performance across victim, decoy, and law enforcement data, revealing generalization gaps in LM-based risk detection.
Advisor: Julia Taylor Rayz.

Aug 2021 – May 2022
Ross-Lynn Graduate Student Fellowship (*formerly Purdue Research Foundation Fellowship*)
Project: Steganographic Approaches Using Natural Language
Contribution: Authored the successful proposal and conducted research on encoding hidden messages within natural language using language model embeddings.
Advisor: Julia Taylor Rayz.

May 2020 – Jul 2021
Research Assistantship funded by NSF Smart & Connected Communities Grant (No. 1737591).
Project: Sociotechnical Systems to Enable Smart and Connected Energy-Aware Residential Communities
Contribution: Developed classification models to address intent ambiguity in user-generated natural language queries.
Advisor: Julia Taylor Rayz.

May 2019 – Aug 2019
Research Assistantship funded by NSF I-Corps Team Grant (No. 1932343).
Project: Decision Support Tool to Assess Distributed Electricity Needs
Contribution: Participated in the national I-Corps program; conducted over 100 stakeholder interviews to define business models and value propositions.
Advisor: Lisa Bosman.

Aug 2018 – Apr 2019
Research Assistantship funded by Purdue Polytechnic Institute Charrette Research Award.
Project: Decision Support Tool to Assess Distributed Electricity Needs
Contribution: Created real-time solar data workflows for three U.S. regions; performed time series analysis and evaluated grid impacts under variable conditions.
Advisor: Lisa Bosman.

Teaching

Spring'22 – Fall'24 **Guest Instructor - Natural Language Technologies (CNIT 519)**
Course Professor: Dr. Julia Taylor Rayz
Lectures covering Transformer based language modeling, contextual word representations and word-sense relations.

Spring'23 **Guest Instructor - Research Methodology and Design (CNIT 322)**
Course Professor: Dr. Tatiana Renae Ringenberg
Lectures covering research methods in Natural Language Processing.

Reviewing

2023 The 11th International Conference on Learning Representations (ICLR)
2023 The 61st Annual Meeting of the Association for Computational Linguistics (ACL)
2022 Spring Undergraduate Research Conference, Purdue University
2021 International Conference on Cyberworlds (CW)

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