Geetanjali Bihani

✉ gbihani@purdue.edu | 🌐 https://www.gbihani.com | in linkedin.com/in/gbihani/

EDUCATION

Purdue University, West Lafayette, USA 2020 – 2025 *(Expected)*
Ph.D. in Computer and Information Technology (GPA - 3.82)
Research Interests: Reliability in LLMs, Lexical Semantics, Text Representations

Purdue University, West Lafayette, USA 2018 – 2020
M.S. in Computer and Information Technology (GPA - 3.69)
Thesis: Longitudinal Comparison of Word Associations in Shallow Word Embeddings [link]

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

FELLOWSHIPS & ASSISTANTSHIPS

2022 – Current **Research Assistantship** funded through U.S. Department of Justice Award (15PJDP-22-GK-03107-MECP). **Title:** *LIve-streaming Offender Network-based Chat Analysis Triage Tool (LION-CATT).* **Contribution:** Develop methods to flag online predatory intents using Large Language Models (LLMs). **Advisor:** Julia Taylor Rayz.

2021 – 2022 **Ross-Lynn Graduate Student Fellowship** *(previously Purdue Research Foundation (PRF) Fellowship).* **Title:** *Steganographic Approaches using Natural Language* **Contribution:** Wrote the proposal in its entirety; conducted research on natural language steganography using language model (LM) embeddings. **Mentor:** Julia Taylor Rayz.

2020 – 2021 **Research Assistantship** funded through NSF Smart & Connected Communities Grant (1737591). **Title:** *Sociotechnical Systems to Enable Smart and Connected Energy-Aware Residential Communities* **Contribution:** Researched multi-intent classification to address intent ambiguity in natural language utterances. **Advisor:** Julia Taylor Rayz.

2019 **Research Assistantship** funded through NSF I-Corps Team National Award (1932343). **Title:** *Decision Support Tool to Assess Distributed Electricity Needs* **Contribution:** Participated in the NSF I-Corps program, conducted customer discovery with over a hundred stakeholders in three months to define cost models and revenue streams. **Advisor:** Lisa Bosman.

2018 – 2019 **Research Assistantship** funded through Purdue Polytechnic Institute Charrette Research Award. **Title:** *Decision Support Tool to Assess Distributed Electricity Needs* **Contribution:** Developed real-time data workflow for solar energy in three U.S. locations, analyzed grid impacts through time series modeling, and performed sensitivity analysis on solar and net metering data. **Advisor:** Lisa Bosman.

TRAVEL GRANTS

Holistic Safety and Security Research Travel Grant. Amount: $250	*2023*
CIT Research Travel Grant Award, Purdue CIT. Amount: $400	*2022*
Dean's Graduate Student Travel Grant. Amount: $250	*2022*
Purdue Graduate Student Government Travel Grant. Amount: $500	*2022*
Holistic Safety and Security Research Travel Grant. Amount: $250	*2022*

HONORS AND AWARDS

Outstanding Student Paper Award. *NAFIPS 2024.* **Amount:** $350	*2024*
Third Place *PPI Holistic Safety and Security Research Impact Area Poster Presentation.* **Amount:** $200	*2024*
Honorable Mention for Student Paper *NAFIPS 2023.* **Amount:** $150	*2023*
Runner-up *PPI Holistic Safety and Security Research Impact Area Poster Presentation.* **Amount:** $250	*2022*
Ross-Lynn Graduate Student Fellowship, Purdue Research Foundation.	*2021*
Second Place, *PPI Holistic Safety and Security Student Presentation Competition*	*2021*
National Science Foundation (NSF) I-Corps Teams National Award	*2019*

Publications

- **May. 2024 -** Damin Zhang, Yi Zhang, **Geetanjali Bihani**, and Julia Rayz."Hire Me or Not? Examining Language Model's Behavior with Occupation Attributes." - Under Review

- **May. 2024 - Geetanjali Bihani**, Tatiana Ringenberg, and Julia Taylor Rayz."Evaluating Language Models on Grooming Risk Estimation Using Fuzzy Theory." Manuscript accepted to Annual Conference of the North American Fuzzy Information Processing Society, NAFIPS 2024

- **May. 2024 - Geetanjali Bihani** and Julia Taylor Rayz."A Fuzzy Evaluation of Sentence Encoders on Grooming Risk Classification." Manuscript accepted to Annual Conference of the North American Fuzzy Information Processing Society, NAFIPS 2024

- **Jan. 2024 - Geetanjali Bihani**, Julia Taylor Rayz. Learning Shortcuts: On the Misleading Promise of NLU in Language Models. *Proceedings of the 57th Hawaii International Conference on System Sciences (HICSS-57)* [preprint]

- **May. 2023 - Geetanjali Bihani**, Julia Taylor Rayz. Calibration Error Estimation Using Fuzzy Binning. *Fuzzy Information Processing 2023 (NAFIPS 2023), Lecture Notes in Networks and Systems, vol 751. Springer, Cham.* [paper]

- **Apr. 2022 - Geetanjali Bihani**, Julia Taylor Rayz. On Information Hiding in Natural Language Systems. *The International FLAIRS Conference Proceedings, 35.* [paper]

- **Aug. 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. *Building and Environment, 109252.* [paper]

- **Jun. 2022 - Geetanjali Bihani**. Interpretable Privacy Preservation of Text Representations Using Vector Steganography. *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence.* [paper]

- **Jun. 2021 - Geetanjali Bihani**, Julia Taylor Rayz. Low Anisotropy Sense Retrofitting (LASeR) : Towards Isotropic and Sense Enriched Representations. *Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, NAACL 2021.* [paper]

- **Jul. 2021 - Geetanjali Bihani**, Julia Taylor Rayz. Fuzzy Classification of Multi-intent Utterances. *Proceedings of the Annual Conference of the North American Fuzzy Information Processing Society.* [paper]

- **Dec. 2020 - 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.* [paper]

Posters

Spring 2024	*Detecting Online Grooming Risk Using Transformers* 2024 Purdue Polytechnic RIA 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.* 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 2021	

INDUSTRY EXPERIENCE

Microsoft Research – _Research Intern_ Summer 2023
Project: Automatic Induction of Interpretable Document Templates using LLMs

Microsoft Research – _Research Intern_ Summer 2022
Project: : Automatic Template Discovery using Language Model embeddings

Transorg Analytics – _Data Scientist_ 2016 – 2018
Implemented data mining and machine learning solutions for demand forecasting, and targeted marketing

India Smart Grid Forum – _Research Intern_ Summer 2015
Assisted in research and data collection for the Machine-to-Machine Communication Roadmap (Power Sector)

TALKS & PANELS

Fall 2023	**Talk:** _Automatic Induction of Interpretable Document Templates_ Microsoft Research
Spring 2023	**Panelist:** _How do next-generation researchers see Mode 2 science?_ International Workshop of the Society of Design and Process Science (SDPS)
Fall 2022	**Talk:** _Automatic Template Discovery_ Microsoft Research
Spring 2022	**Invited Talk:** _Language, Representations and Leakage_ RAISE Lab, Syracuse University
Spring 2022	**Talk:** _Interpretable Privacy Preservation of Text Representations Using Vector Steganography_ AAAI 2022 Doctoral Consortium CERIAS Security Symposium 2022

TEACHING

Fall 2023, **Spring 2023,** **Fall 2022,** **Spring 2022**	**Guest Instructor - Natural Language Technologies (CNIT 519)** Course Instructor: Dr. Julia Taylor Rayz Lectures covering Transformer based language modeling, contextual word representations and word-sense relations
Spring 2023	**Guest Instructor - Research Methodology and Design (CNIT 322)** Course Instructor: Dr. Tatiana Renae Ringenberg Lectures covering research methods in Natural Language Processing

MENTORSHIP

2023	Krishnakanth Alagiri (Graduate). **Topic:** Instruction tuning LLMs to decipher grooming strategies in online chats.

REVIEWING

The 11th International Conference on Learning Representations (ICLR)	_2023_
The 61st Annual Meeting of the Association for Computational Linguistics (ACL)	_2023_
Spring Undergraduate Research Conference, Purdue University	_2022_
International Conference on Cyberworlds (CW)	_2021_

SERVICE

Officer (Research and Engagement): Purdue CIT Graduate Student Association	_2023-25_
Organizer: Special Session on Natural Language Uncertainty in the Era of LLMs (NAFIPS 2024)	_2024_
Program Committee Member: European Interdisciplinary Cybersecurity Conference (EICC)	_2023_
Student Organizer: Conference of the North American Fuzzy Information Processing Society (NAFIPS 2021)	_2021_
Graduate Student Advisor: Purdue CIT Student Council	_2019-20_
Head, Public Relations: IEEE Delhi Technological University Student Branch (IEEE-DTU)	_2015-16_
Secretary: Women in Engineering, Delhi Technological University Student Branch (WIE-DTU)	_2015-16_

Professional Affiliations

- Association of Computational Linguistics (ACL)

- Center for Education and Research in Information Assurance and Security (CERIAS)

- Institute of Electrical and Electronic Engineers (IEEE)

- North American Fuzzy Information Processing Society (NAFIPS)

- Association for the Advancement of Artificial Intelligence (AAAI)

- Association for Computing Machinery (ACM)

Technical Skills

Programming: Python, SQL (Postgres), R, LaTeX, Git
Libraries: pandas, NumPy, Matplotlib, pytorch, beautifulsoup4, tensorflow, gensim
Visualization: Tableau, Prezi, Photoshop, Procreate
Statistics: Statistics Probability Theory, GLMs, LMEMs, Bayesian Models
Natural Languages: English, Hindi

References

*Available upon request