Thilini Wijesiriwardene

Ph.D., College of Engineering and Computing, University of South Carolina

Homepage | thilini@sc.edu | thilini-w | Thiliniiw | Google Scholar

Pittsburgh, Pennsylvania - 15229, United States

Summary

I hold a Ph.D. in Computer Science and Engineering and bring 5+ years of specialized expertise in Natural Language Processing and Large Language Models. My research advances analogical reasoning and concept abstraction, resulting in peer-reviewed publications across neurosymbolic AI, language understanding, and biomedical prediction. I lead multidisciplinary AI initiatives that deliver scalable solutions in healthcare and education, combining technical mastery of Python, PyTorch, and Huggingface with strong capabilities in mentoring junior researchers and communicating insights across diverse audiences.

Work Experience

- **Artificial Intelligence Institute, University of South Carolina [🌐]** — *08/2019 - 12/2025*
  Graduate Research Assistant (R.1) — South Carolina, USA
- **National Library of Medicine [🌐]** — *05/2021 - 01/2022*
  Research Intern (R.2) with Dr. Olivier Bodenreider — Maryland, USA
- **Wright State University [🌐]** — *01/2018 - 08/2019*
  Graduate Research Assistant (R.3) — Ohio, USA
- **Wright State University [🌐]** — *09/2017 - 12/2017*
  Volunteer Researcher (R.3) — Ohio, USA
- **University of Colombo School of Computing [🌐]** — *03/2015 - 06/2017*
  Assistant Lecturer — Colombo, Sri Lanka
- **University of Colombo School of Computing [🌐]** — *10/2014 - 03/2015*
  Graduate Instructor — Colombo, Sri Lanka
- **attune Consulting [🌐]** — *02/2014 - 10/2014*
  Associate SAP Consultant — Colombo, Sri Lanka

Education

- **University of South Carolina** — *08/2019 - Fall 2025*
  Doctor of Philosophy, Computer Science — South Carolina, USA
  - **Advisor:** Dr. Amit Sheth.
  - **Relevant Coursework:** Artificial Intelligence, Reinforcement Learning, Natural Language Processing, Big Data Analytics.
  - **Grade Point Average:** 4.0
- **Wright State University** — *01/2018 - 08/2019*
  MS in Computer Science (Transferred to Ph.D.) — Ohio, USA
  - **Advisor:** Dr. Amit Sheth.
  - **Relevant Coursework:** Algorithm Design and Analysis, Deep Learning, Information Retrieval.
  - **Grade Point Average:** 3.89
- **University of Colombo School of Computing** — *08/2009 - 02/2014*
  Bachelor of Science in Information and Communication Technology — Colombo, Sri Lanka
  - **Relevant Coursework:** Data Mining, , Research Methods.
  - **Grade Point Average:** 3.61

Research Projects

R=Research Project

[R.1] Analogy and Abstraction in the Era of Large Language Models [🌐] | *Publications: C.1, C.2, C.3, J.1, W.2*

- Spearheaded a research initiative on Large Language Models, directing a team of junior researchers to investigate analogical reasoning and abstraction with external knowledge, driving innovative insights and project success.
- Orchestrated the development and publication of multiple high-impact datasets and methods (See C.3 and C.1), leading collaborative efforts of several junior and senior researchers to advance research in the field.

[R.2] Vocabulary Alignment in UMLS Metathesaurus | *Publications: C.5, W.1, PP.1*

- Led the development and training of a BERT-based language model for synonymy prediction in the UMLS Metathesaurus, utilizing Python and PyTorch to achieve over 90% accuracy, enabling precise biomedical vocabulary alignment for healthcare applications.
- Directed the design and evaluation of scalable machine learning models for vocabulary alignment in the UMLS Metathesaurus, collaborating with cross-functional teams to build robust data pipelines and feature engineering workflows, advancing interoperability in medical informatics.

[R.3] Context-Aware Harassment Detection on Social Media [🌐] | *Publication: C.6*

- Led a research initiative to develop NLP models for contextual analysis of online interactions, leveraging Python and TensorFlow to identify and mitigate toxic content in social media, enhancing decision-support capabilities for user-centric AI applications.
- Orchestrated the creation of a novel dataset of tweets paired with contextual metadata, collaborating with cross-functional teams to design scalable data pipelines and feature engineering workflows, advancing NLP research with potential applications in sentiment analysis for healthcare communication.

Publications

C=Conference, J=Journal, W=Workshops, *=Shared authorship

[C.1] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Sreeram Reddy Vennam, Vinija Jain, Aman Chadha, Amitava Das, Ponnurangam Kumaraguru and Amit Sheth. **"KnowledgePrompts: Exploring the Abilities of Large Language Models to Solve Proportional Analogies via Knowledge-Enhanced Prompting."** In Proceedings of the 31st International Conference on Computational Linguistics (COLING) 2025.

[C.2] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Aishwarya Naresh Reganti, Vinija Jain, Aman Chadha, Amit Sheth and Amitava Das. **"On the relationship between sentence analogy identification and sentence structure encoding in large language models."** In Findings of the Association for Computational Linguistics: *EACL* 2024.

[J.1] Thilini Wijesiriwardene, Amit Sheth, Valerie L. Shalin and Amitava Das. **"Why Do We Need Neurosymbolic AI to Model Pragmatic Analogies?."** In IEEE Intelligent Systems 2023.

[C.3] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Bimal Gajera, Shreeyash Gowaikar, Chandan Gupta, Aman Chadha, Aishwarya Naresh Reganti, Amit Sheth, and Amitava Das. **"ANALOGICAL - A Novel Benchmark for Long Text Analogy Evaluation in Large Language Models."** In Findings of the Association for Computational Linguistics: ACL 2023.

[C.4] Riccardo Tommasini, Filip Ilievski, and Thilini Wijesiriwardene. **"IMKG: The internet meme knowledge graph."** In The Semantic Web: 20th International Conference, ESWC 2023.

[W.1] Goonmeet Bajaj, Vinh Nguyen, Thilini Wijesiriwardene, Hong Yung Yip, Vishesh Javangula, Amit Sheth, Srinivasan Parthasarathy, and Olivier Bodenreider. **"Evaluating Biomedical Word Embeddings for Vocabulary Alignment at Scale in the UMLS Metathesaurus Using Siamese Networks."** In Proceedings of the Third Workshop on Insights from Negative Results in NLP 2022.

[C.5] Vinh Nguyen, Hong Yung Yip, Goonmeet Bajaj, Thilini Wijesiriwardene, Vishesh Javangula, Srinivasan Parthasarathy, Amit Sheth, and Olivier Bodenreider. **"Context-enriched learning models for aligning biomedical vocabularies at scale in the UMLS Metathesaurus."** In Proceedings of the ACM Web Conference 2022 (WWW '22). Association for Computing Machinery.

[W.2] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Valerie L Shalin, Amit P Sheth. **"Towards efficient scoring of student-generated long-form analogies in STEM."** Workshop on Analogies: from Theory to Applications at ICCBR 2022.

[C.6] Thilini Wijesiriwardene*, Hale Inan*, Ugur Kursuncu, Manas Gaur, Valerie L Shalin, Krishnaprasad Thirunarayan, Amit Sheth, and I Budak Arpinar. **"Alone: A dataset for toxic behavior among adolescents on twitter."** In Social Informatics: 12th International Conference, SocInfo 2020.

Other Articles

PP=Preprint

[PP.1] Thilini Wijesiriwardene,Vinh Nguyen, Goonmeet Bajaj, Hong Yung Yip, Vishesh Javangula, Yuqing Mao, Kin Wah Fung, Srinivasan Parthasarathy, Amit P Sheth, and Olivier Bodenreider. **"Ubert: A novel language model for synonymy prediction at scale in the umls metathesaurus."** arXiv preprint arXiv:2204.12716 (2022).

Skills

- **Programming Languages:** Python, C, Java, R, SQL, Prolog.
- **Frameworks and Libraries:** PyTorch, Scikit-learn, Pandas, Numpy, Matplotlib, Huggingface, Langchain.
- **Software & Tools:** GitHub, Anaconda, MS Office, MATLAB, Protege, Neo4j.

Travel Grants & Other Awards

- Student Volunteer: The 31st International Conference on Computational Linguistics 2025, Abu Dhabi, UAE
- Student Volunteer: The 62nd Annual Meeting of the Association for Computational Linguistics 2023, Toronto, Canada [🌐]
- Student Scholarship: The 37th AAAI Conference on Artificial Intelligence 2023, Wasgington DC, USA [🌐]
- CRA-WP Grad Cohort Workshop for Women 2022, New Orleans, Louisiana, USA [🌐]
- Grace Hopper Celebration 2020, Virtual [🌐]
- Travel Award: WSTNET Web Science Summer School 2018, Hannover, Germany [🌐]
- Travel Award: Tri-State Celebration of Women in Computing (TRIWiC) 2018, Cincinnati, OH, USA [🌐]

Academic Service

R=Reviewer, AC=Area Chair, PC=Program Committee Member

[R] ACL Rolling Review 2024 (April, June, August, October, December),
Empirical Methods in Natural Language Processing (EMNLP) 2023,
ACM SIGCHI Conference on Computer-Supported Cooperative Work & Social Computing (CSCW) 2022,
ACM Transactions on Computing for Healthcare 2024.

[AC] ACL Rolling Review (2025 February).

[PCM] Extended Semantic Web Conference (ESWC) - Resource Track 2025.