

# SANGPIL YOUM

812-830-2413 | [youms@ufl.edu](mailto:youms@ufl.edu) | <https://github.com/philz0918> | <https://huggingface.co/yeomtong>

## EDUCATION

<b>University of Florida, Gainesville, FL, USA</b> <i>Ph D. in Computer Science</i>	May 2022 - Expected: August 2026
<b>Indiana University, Bloomington, IN, USA</b> <i>Master of Science in Data Science</i>	Advisor: Bonnie J. Dorr May 2021
<b>SungKyunKwan University, Seoul, South Korea</b> <i>Bachelor in Library and Information Science, Data Science</i>	Feb 2019

## PUBLICATIONS

- Youm, S., Chaeeun Han, Hojeong Yoo, Sou Hyun Jang, Bonnie Dorr. “I Still Need Your Help”: Online Information Seeking Behavior Among International Students on Reddit, *Public Library of Science One (PLOS One)*, to appear, 2026.
- Youm, S., Leah Jones, Bonnie Dorr. Revisiting Semantic Role Labeling: Modernizing for Structure, Speed, and Multilinguality, submitted to LREC
- Khan, Y., Xinlei Wu, Sangpil Youm, Justin Ho, Aryaan Mehboob Shaikh, Jairo Garciga, Rohan Sharma, Bonnie Dorr. DETQUS: Decomposition-Enhanced Transformers for QUery-focused Summarization, *Proceedings of the 2025 Annual Conference of The Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL)*, New Mexico, pp. 2720–2731, 2025. *Paper Link*
- Youm, S., Brodie Mather, Chathuri Jayaweera, Juliana Prada, Bonnie Dorr. DAHRS: Divergence-Aware Hallucination-Remediated SRL Projection. *The 29th International Conference on Natural Language & Information Systems (NLDB)*, Turin, Italy, pp. 423–438, 2024. *Paper Link*
- Jayaweera, C., Sangpil Youm, Bonnie Dorr. AMREx: AMR for Explainable Fact Verification, *Proceedings of the Seventh Fact Extraction and VERification Workshop (FEVER)*, Conference on Empirical Methods in Natural Language Processing (EMNLP), Miami, pp. 234–244, 2024. *Paper Link*
- Zhang, D., Bushi Xiao, Chao Gao, Sangpil Youm, Bonnie Dorr. Modeling Bilingual Sentence Processing: Evaluating RNN and Transformer Architectures for Cross-Language Structural Priming, *Proceedings of the Fourth Workshop on Multilingual Representation Learning (MRL 2024)*, Conference on Empirical Methods in Natural Language Processing (EMNLP), Miami, pp. 127–136, 2024. *Paper Link*
- Martinez, M., Sonja Schmer-Galunder, Zoey Liu, Sangpil Youm, Chathuri Jayaweera, Bonnie Dorr. Balancing Transparency and Accuracy: A Comparative Analysis of Rule-Based and Deep Learning Models in Political Bias Classification, *Proceedings of the Second Workshop on Social Influence in Conversations (SICon 2024)*, Conference on Empirical Methods in Natural Language Processing (EMNLP), Miami, pp. 102–115, 2024. *Paper Link*

## WORK EXPERIENCE

<b>Instructor</b> <i>Department of Computer and Information Science and Engineering at University of Florida</i>	Spring 2024
• Designed and taught Natural Language Processing (CAI 6307), covering syntax, semantics, and large language models (LLMs), with emphasis on real-world applications	
• Led semester-long team-based projects, including: Tabular-based question answering models, Political bias classification systems, Bilingual sentence processing, with student work resulting in publications at NLP conferences	
<b>Graduate Teaching Assistant</b> <i>Department of Computer and Information Science and Engineering at University of Florida</i>	Jan 2023 – Present
• Supported instruction for Natural Language Processing (CAP 4641), spanning classical NLP methods through modern LLMs	
• Managed GPU-based developing environments using the University of Florida’s high-performance computing infrastructure	
<b>Graduate Research Assistant</b> <i>Department of Computer and Information Science and Engineering at University of Florida</i>	Mar 2022 – Present
• Developed a modernized Semantic Role Labeling (SRL) model improving efficiency and robustness	
• Approximately $\approx X10$ speed up in modernized SRL model and 13% accuracy improvement in cross-lingual SRL compared to prior baselines	
<b>Graduate Researcher</b> <i>Wissee</i>	Oct 2020 – May 2021
• Developed an LSTM-CRF model for Named Entity Recognition (NER) on real-world data	
• Achieved 69.3% F1 score under noisy and domain-specific conditions	