

SANGPIL YOUM

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

EDUCATION

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

WORK EXPERIENCE

Instructor <i>Department of Computer and Information Science and Engineering at University of Florida</i>	Spring 2024
<ul style="list-style-type: none">Designed and taught Natural Language Processing (CAI 6307), covering syntax, semantics, and large language models (LLMs), with emphasis on real-world applicationsLed 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	
Graduate Teaching Assistant <i>Department of Computer and Information Science and Engineering at University of Florida</i>	Jan 2023 – Present
<ul style="list-style-type: none">Supported instruction for Natural Language Processing (CAP 4641), spanning classical NLP methods through modern LLMsDesigned and maintained course projects integrating practical NLP pipelinesManaged GPU-based developing environments using the University of Florida's high-performance computing infrastructure	
Graduate Research Assistant <i>Department of Computer and Information Science and Engineering at University of Florida</i>	Mar 2022 – Present
<ul style="list-style-type: none">Developed a modernized Semantic Role Labeling (SRL) model improving efficiency and robustnessImplemented a projection-based cross-lingual SRL frameworkApproximately $\approx X10$ speed up in modernized SRL model and 13% accuracy improvement in cross-lingual SRL compared to prior baselines	
Graduate Researcher <i>Wissee</i>	Oct 2020 – May 2021
<ul style="list-style-type: none">Developed an LSTM-CRF model for Named Entity Recognition (NER) on real-world dataAchieved 69.3% F1 score under noisy and domain-specific conditions	

PUBLICATIONS

Peer-Reviewed Articles
<ul style="list-style-type: none">Youm, S., Cheeun Han, Hojeong Yoo, Sou Hyun Jang, Bonnie Dorr. “I Still Need Your Help”: Online Information Seeking Behavior Among International Students on Reddit, <i>PLOS One</i>, 2026. <i>To appear</i>Coggins, W., Jasmine McKenzie, Sangpil Youm, Pradham Mummaleti, Juan Gilbert, Eric Ragan, and Bonnie Dorr. 2025. That Ain’t Right: Assessing LLM Performance on QA in African American and West African English Dialects. In <i>Proceedings of the 9th Widening NLP Workshop</i>, pages 123–129, Suzhou, China. Association for Computational Linguistics. <i>Paper Link</i>Khan, Y., Xinlei Wu, Sangpil Youm, Justin Ho, Aryaan Mehboob Shaikh, Jairo Garciga, Rohan Sharma, and Bonnie Dorr. DETQUS: Decomposition-Enhanced Transformers for QUery-focused Summarization, in <i>Proceedings of the 2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL)</i>, 2025. <i>Paper Link</i>Kang, P., Ko, Y., Youm, S., Chun, Y., Davis, S., Connaughton, D. Application of Machine and Deep Learning Techniques on Sport Non-Fungible Token Tweets: Exploration of Perceived Values and Risks, <i>International Journal of Sports Marketing and Sponsorship</i>, 2025. <i>Paper Link</i>Jayaweera, C., Sangpil Youm, Bonnie Dorr. AMREx: AMR for Explainable Fact Verification in <i>Proceedings of the Seventh Fact Extraction and VERification Workshop (FEVER) on Conference on Empirical Methods in Natural Language Processing (EMNLP)</i>, 2024. <i>Paper Link</i>

- Zhang, D., Bushi Xiao, Chao Gao, **Sangpil Youm**, and Bonnie Dorr. Modeling Bilingual Sentence Processing: Evaluating RNN and Transformer Architectures for Cross-Language Structural Priming in *Proceedings of the Fourth Workshop on Multilingual Representation Learning (MRL 2024) on Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024. *Paper Link*
- Martinez, M., Sonja Schmer-Galunder, Zoey Liu, **Sangpil Youm**, Chathuri Jayaweera, and Bonnie Dorr. Balancing Transparency and Accuracy: A Comparative Analysis of Rule-Based and Deep Learning Models in Political Bias Classification in *Proceedings of the Second Workshop on Social Influence in Conversations (SICon 2024) on Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024. *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)*. 2024. *Paper Link*
- Koo, S., **Sangpil Youm**, Jane Shin. Cycle-GAN-based synthetic sonar image generation for improved underwater classification *Ocean Sensing and Monitoring XVI*. 2024. *Paper Link*
- Jang, S. H., **Sangpil Youm**, Yong Jeong Yi. Anti-Asian Discourse in Quora: Comparison of Before and During the COVID-19 Pandemic with Machine- and Deep-Learning Approaches *Race and Justice*. 2023; 13(1): 55–79. *Paper Link*
- Jang S. H., Mary A. Gerent, **Sangpil Youm**, Yong Jeong Yi. Understanding coronavirus disease 2019 (COVID-19) vaccine hesitancy: Evidence from the community-driven knowledge site Quora. *DIGITAL HEALTH*. 2022. *Paper Link*
- Kim J, **Sangpil Youm**, Yongwei Shan, Jonghoon Kim. Analysis of Fire Accident Factors on Construction Sites Using Web Crawling and Deep Learning Approach. *Sustainability*. 2021; 13(21):11694. *Paper Link*

Manuscript under review

- **Youm, S.**, Leah Jones, Bonnie Dorr. Revisiting Semantic Role Labeling: Modernizing for Structure, Speed, and Multilinguality, submitted to LREC
- **Youm, S.**, Sou Hyun Jang, Haewoon Kwak, Jaeyoung Choi, Yong Jeong Yi. Cross-Cultural Vaccine Perceptions: Comparing COVID-19 Vaccine Public Discourse in the U.S. and South Korea using Natural Language Processing, *Journal of Medical Internet Research (JMIR)*, Revise & Resubmit.

WORKING PROJECTS

- Role-aware Extraction of Forms of Linguistic Expressions for Communication Tracking** *Dissertation*
- Develop and Evaluate Semantic Role Labeling (SRL)-aware BERT-based and prompt-based Transformer models for detecting internal states (emotion, sentiment) and external linguistic cues (politeness, formality).
 - Generalize linguistic-cue detection to multiple languages through SRL projection techniques.
 - Use Case 1—Evaluating Team Performance in High-Stress Environments: Analyze communication breakdowns and their relationship to behavioral proxies of team performance.
 - Use Case 2—Monitoring Adolescent Mental Health: Identify linguistic signals associated with self-reported mental well-being.

Mixed Emotions on Reddit Predict Depressive Symptoms in Asian American Families

- Adopt BERT-based GoEmotion detection model onto Reddit data
- Conduct statistical Analysis for the relation between mixed emotions and depressive symptoms

PROFESSIONAL SERVICE

Peer Review

- Computers in Human Behavior
- International Conference on Computational Social Science (IC2S2)
- Language Resources and Evaluation

PRESENTATIONS

Academic Conference

- **Youm, S.**, Brodie Mather, Chathuri Jayaweera, Juliana Prada, Bonnie Dorr. July 2024. DAHRS: Divergence-Aware Hallucination-Remediated SRL Projection. *The 29th International Conference on Natural Language & Information Systems (NLDB)*.
- **Youm, S.**, Chaeeun Han, Sou Hyun Jang. July 2024. Information-seeking Behavior Among International Students on Reddit: A Comparative Analysis Before and After Start of COVID-19. *10th International Conference on Computational Social Science (IC2S2)*
- **Youm, S.**, Nari Yoo, Sou Hyun Jang. July 2024. Beyond Positive or Negative: How Mixed Emotions on Reddit Predict Depressive Symptoms in Asian American Families. *10th International Conference on Computational Social Science (IC2S2)*
- **Youm, S.**, Sou Hyun Jang, Yong Jeong Yi. April 2022. Understanding COVID-19 Vaccine Hesitancy: Evidence from Quora. *In proceeding of Society of Behaviour Medicine*

LEADERSHIP, AWARDS AND OTHERS

Excellent Project Award in Python Programming Bootcamp	August 2017
<i>SungKyunKwan University(SKKU), Seoul, South Korea</i>	
Outstanding Performance Scholarship Awards	April, September 2017
<i>SungKyunKwan University(SKKU), Seoul, South Korea</i>	
President of the student council	December 2016 – December 2017
<i>Department of Data Science and Library and Information Science</i>	
Army Sergeant Discharge in Full Time	December 2011 - September 2013
<i>Republic of Korea Army, Pocheon, South Korea</i>	

REFERENCES

Available upon request