

SANGPIL YOUM

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

EDUCATION

University of Florida, Gainesville, FL, USA

Ph D. in Computer Science

May 2022 - Expected: August 2026

Advisor: Bonnie J. Dorr

Indiana University, Bloomington, IN, USA

Master of Science in Data Science

May 2021

SungKyunKwan University, Seoul, South Korea

Bachelor in Data Science, Library and Information Science

Feb 2019

WORK EXPERIENCE

Instructor

Spring 2024

Department of Computer and Information Science and Engineering at University of Florida

- 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

Graduate Teaching Assistant

Jan 2023 – Present

Department of Computer and Information Science and Engineering at University of Florida

- Supported instruction for Natural Language Processing (CAP 4641), spanning classical NLP methods through modern LLMs
- Designed and maintained course projects integrating practical NLP pipelines
- Managed GPU-based developing environments using the University of Florida's high-performance computing infrastructure

Graduate Research Assistant

Mar 2022 – Present

Department of Computer and Information Science and Engineering at University of Florida

- Developed a modernized Semantic Role Labeling (SRL) model improving efficiency and robustness
- Implemented a projection-based cross-lingual SRL framework
- Approximately $\approx X10$ speed up in modernized SRL model and 13% accuracy improvement in cross-lingual SRL compared to prior baselines

Graduate Researcher

Oct 2020 – May 2021

Wissee

- 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

PUBLICATIONS

Peer-Reviewed Articles

- **Youm, S.**, Chaeun 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.**, Sou Hyun Jang, Haewoon Kwak, Jaeyoung Choi, Yong Jeong Yi. Cross-Cultural Differences in Public Discourse on COVID-19 Vaccination in the U.S. and South Korea: Cross-Sectional Analysis Using Natural Language Processing, *Journal of Medical Internet Research (JMIR) Formative Research*, to appear, 2026.
- Han, C., **Sangpil Youm**, Hojeong Yoo, Sou Hyun Jang. "Revisiting Conceptualizing Stigma": Unpacking Minority College Students' Stigmatization on Reddit, *Cyberpsychology, Behavior, and Social Networking*, to appear, 2026
- Coggins, W., Jasmine McKenzie, **Sangpil Youm**, Pradham Mummaleti, Juan Gilbert, Eric Ragan, Bonnie Dorr. That Ain't Right: Assessing LLM Performance on QA in African American and West African English Dialects, *Proceedings of the 9th Widening NLP Workshop, Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Suzhou, China, pp. 123–129, 2025. [Paper Link](#)
- 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](#)

- Kang, P., Yongjae Ko, **Sangpil Youm**, Yoonki Chun, Sean Davis, Daniel Connaughton. Application of Machine and Deep Learning Techniques on Sport Non-Fungible Token Tweets: Exploration of Perceived Values and Risks, *International Journal of Sports Marketing and Sponsorship*, 26:4, pp. 862–883, 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](#)
- Koo, S., **Sangpil Youm**, Jane Shin. Cycle-GAN-based synthetic sonar image generation for improved underwater classification, *Proceedings of the Ocean Sensing and Monitoring XVI*, 13061, pp. 69–83, 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*, 13:1, pp. 55–79, 2023. [Paper Link](#)
- Jang S. H., Mary A. Gerend, **Sangpil Youm**, Yong Jeong Yi. Understanding coronavirus disease 2019 (COVID-19) vaccine hesitancy: Evidence from the community-driven knowledge site Quora, *DIGITAL HEALTH*, 8, 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*, 13:21, 2021. [Paper Link](#)

Manuscript under review

- **Youm, S.**, Leah Jones, Bonnie Dorr. Revisiting Semantic Role Labeling: Modernizing for Structure, Speed, and Multilinguality, submitted to LREC

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.**, Chaeun 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

SungKyunKwan University(SKKU), Seoul, South Korea

Outstanding Performance Scholarship Awards

April, September 2017

SungKyunKwan University(SKKU), Seoul, South Korea

President of the student council

December 2016 – December 2017

Department of Data Science and Library and Information Science

Army Sergeant Discharge in Full Time

December 2011 - September 2013

Republic of Korea Army, Pocheon, South Korea

REFERENCES

Available upon request