

SANGPIL YOUM

812-830-2413 | philz0918@gmail.com | <https://github.com/philz0918>

EDUCATION

University of Florida, Gainesville, FL

Mar 2022 - present

Ph D. in Computer Science GPA : 3.83/4.00

Indiana University, Bloomington, IN

May 2021

Master of science in Data Science GPA : 3.88/4.00

SungKyunKwan University, Seoul, South Korea

Feb 2019

Bachelor in Data Science, Library and Information Science GPA : 3.92/4.50 , 4.08/4.50

University at Buffalo, Buffalo, NY

Jan 2016 - May 2016

Exchange Student GPA : 3.93/4.00

PUBLICATIONS

Peer-Reviewed Articles

- **Youm, S.**, Mather, B., Jayaweera, C., Prada, J., Dorr, B. DAHRS: Divergence-Aware Hallucination-Remediated SRL Projection. *The 29th International Conference on Natural Language & Information Systems (NLDB)*. 2024.
- Koo, S., **Youm, S.**, Shin, J. Cycle-GAN-based synthetic sonar image generation for improved underwater classification *Ocean Sensing and Monitoring XVI*. 2024.
- Jang, S. H., **Youm, S.**, & Yi, Y. J. 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. <https://doi.org/10.1177/21533687221134690>
- Jang S. H., Gerend MA, **Youm S.**, Yi Y. J. Understanding coronavirus disease 2019 (COVID-19) vaccine hesitancy: Evidence from the community-driven knowledge site Quora. *DIGITAL HEALTH*. 2022;8. <https://doi.org/10.1177/20552076221145426>
- Kim J, **Youm S.**, Shan Y, Kim J. Analysis of Fire Accident Factors on Construction Sites Using Web Crawling and Deep Learning Approach. *Sustainability*. 2021; 13(21):11694. <https://doi.org/10.3390/su132111694>

Manuscript under review

- Martinez, M., Schmer-Galunder, s., Liu, Z., **Youm, S.**, Jayaweera, C., Dorr, B. Balancing Transparency and Accuracy: A Comparative Analysis of Rule-Based and Deep Learning Models in Political Bias Classification
- Srivatsav, R., Venugopal, P., Kaliraman, D., Godhala, M., Pamarthy, V., **Youm, S.**, Dorr, B. RAG for QA: Does Linguistic Information Help?
- Khan, Y., Wu, X., Garciga, J., Sharma, R., **Youm, S.**, Ho, J., Dorr, B. DETQUS: Decomposition-Enhanced Transformers for QUery-focused Summarization
- Zhang, D., Xiao, B., Gao, C., **Youm, S.**, Dorr, B. Modeling Bilingual Sentence Processing: Evaluating RNN and Transformer Architectures for Cross-Language Structural Priming
- **Youm, S.**, Jang, S. H., Choi, J. Y., Yi, Y. J. Comparison of perceptions in the US and South Korea regarding COVID-19 vaccines on community-driven knowledge sites

WORKING PAPERS

Discourse of information seeking behavior of International Student in Reddit using network embedding

| transformers, spacy, pytorch, networkx

- Analyze information seeking behavior through Reddit post using network embedding approach
- Detect topic per user group in subreddit H1B/F1

Network Approach to Assessing the Relationship between Emotions towards parents among Asian and Asian Americans

| transformers, spacy, pytorch

- Extract emotion and depression from Reddit post using transformer models
- Analyze relation between combination of emotions and depression

Developing framework for low resource language Semantic Role Labeling

| transformers, spacy, pytorch

- Create low-resource language transferring SRL model (e.g., English-Tagalog)
- Data augmentation for fine-tuned transformer-based word alignment model

PRESENTATIONS

Academic Conference

- **Youm, S.**, Han, C.E., Jang, S.H. July 2024. Information-seeking Behavior Among International Students on Reddi: A Comparative Analysis Before and After Start of COVID-19. *10th International Conference on Computational Social Science (IC2S2)*
- **Youm, S.**, Yoo, N., Jang, S.H. 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.**, Jang, S.H., Choi, J, Yi, Y.J. July 2024. Comparison of perceptions regarding COVID-19 vaccines on community-driven knowledge sites in the US and South Korea. *WAPOR 77th and WAPOR Asia Pacific 7th Joint Annual Conference: The Soul of Public Opinion Research: Liberty, Quality and Humanity*
- Kang, P., Ko, Y. J., **Youm, S.**, Davis, S., Connaughton, D. June 2024. Consumers' value perception toward sport Non-Fungible Token (NFT): Application of Artificial Intelligence (AI) analytics on Twitter Big Data *North American Society for Sport Management (NASSM)*
- **Youm, S.**, Jang, S. H., Yi, Y.J. April 2022. Understanding COVID-19 Vaccine Hesitancy: Evidence from Quora. *In proceeding of Society of Behaviour Medicine*
- Jang, S.H., **Youm, S.**, Choi, J, Yi, Y.J. November 2021. Comparison of Perception about COVID-19 Vaccine on Social Q & A sites : Quora(U.S.) versus Knowledge-iN(South korea) *Korean Association of Health and Medical Sociology*

WORK EXPERIENCE

Instructor

Spring 2024

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

- Natural Language Processing (CAI 6307)

Graduate Teaching Assistant

Jan 2023 – Present

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

- Natural Language Processing (CAP 4641)
- Developed and operated class projects

Graduate Research Assistant

Mar 2022 – Present

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

- Developed generalizable algorithm for the Semantic Role Labeling (SRL) task in low/mid-resource languages
- Established *Natural Language Processing (NLP)* lab environment

Research Volunteer

June 2021 – Feb 2022

Luddy School of Informatics, Computing and Engineering at Indiana University

- Verified whether clustering method finds or not communities in case of perturbed network
- Construct method for finding out optimized modularity within constrained number of groups by using simulated annealing, developed in Julia and Python

Graduate Researcher

Oct 2020 – May 2021

Wisee

- Built and compared the performance between Deep learning model (BERT, LSTM) and machine learning model(Hidden Markov Model) for Name Entity Recognition.
- Processed Twitter data by applying NLP techniques with Spacy and NLTK and compared the efficiency between them

TECHNICAL SKILLS

Languages and Script : Python, C, C++, Julia, R, SQL, JavaScript, HTML, CSS, XML, Erlang

Machine Learning: Classification, Linear Regression, Logistic Regression Ridge and Lasso Regression, Support Vector machine, Hidden Markov Model, K-means Clustering, Deep Learning - CNN, LSTM, BERT, Transformers

Database: Postgres, MongoDB, MySQL

Libraries: Pandas, Scikit-learn, NLTK, spaC, NetworkX, Numpy, Tensorflow ,Pytorch, Keras, Selenium, BeautifulSoup, HTMLParser, requests, matplotlib, UMAP, genism, scipy, spacy

OS: Windows, Mac, Linux

COURSEWORK

Programming and Computer Science: Introduction to computer science, Computer Architecture Principles, Programming Language, Advanced Data structure, Algorithms, Advanced Database System, Operating System, Data Mining, Deep Learning System, Designing Database Systems

Mathematical and Statistical Foundation: Calculus 2/3, Probability and Statistics, Linear algebra, Time series analysis, Statistical Reasoning

Data Science and Application: Natural Language Processing, Computational Social Science, Network Science, Big Data Analytics, Social Data Analytics, Introduction to Data Science, Data Science and Linguistics, Health Informatics, Information Analysis and Evaluation, Semantic Web system

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