

Eugene Jang

jang.eu@northeastern.edu
<http://genesith.github.io>

EDUCATION	Ph.D, Computer Science Northeastern University Advisor: Terra Blevins	2025–Present
	M.S., Computer Science Korea Advanced Institute of Science and Technology (KAIST) Advisor: Jong C. Park	2019–2021
	B.S., Computer Science (Minor: Physics) Korea Advanced Institute of Science and Technology (KAIST)	2014–2019
EXPERIENCE	Senior Researcher (AI Team), S2W Inc. • Detection of various types of cybercrime [2, 5-7, 10] • Automation of cyber threat intelligence [3-4, 6] • Vulnerabilities of LLMs [1]	2021–2025
PUBLICATIONS	[1] Eugene Jang, Kimin Lee, Jin-Woo Chung, Keuntae Park, Seungwon Shin. <i>Improbable Bigrams Expose Vulnerabilities of Incomplete Tokens in Byte-Level Tokenizers</i>. EMNLP 2025 [2] Minkyoo Song, Eugene Jang, Jaehan Kim, Seungwon Shin. <i>Covering Cracks in Content Moderation: Delexicalized Distant Supervision for Illicit Drug Jargon Detection</i>. KDD 2025 [3] Jian Cui, Hanna Kim, Eugene Jang, Dayeon Yim, Kicheol Kim, Yongjae Lee, Jin-Woo Chung, Seungwon Shin, Xiaojing Liao. <i>Tweezers: A Framework for Security Event Detection via Event Attribution-centric Tweet Embedding</i>. NDSS 2025 [4] Eugene Jang, Jian Cui, Dayeon Yim, Youngjin Jin, Jin-Woo Chung, Seungwon Shin, Yongjae Lee. <i>Ignore Me But Don't Replace Me: Utilizing Non-Linguistic Elements for Pretraining on the Cybersecurity Domain</i>. Findings of NAACL 2024 [5] Hanna Kim, Jian Cui, Eugene Jang, Chanhee Lee, Yongjae Lee, Jin-Woo Chung, Seungwon Shin. <i>DRAINCLoG: Detecting Rogue Accounts with Illegally-obtained NFTs using Classifiers Learned on Graphs</i>. NDSS 2024 [6] Youngjin Jin, Eugene Jang, Jian Cui, Jin-Woo Chung, Yongjae Lee, Seungwon Shin. <i>DarkBERT: A Language Model for the Dark Side of the Internet</i>. ACL 2023 [7] Youngjin Jin, Eugene Jang, Yongjae Lee, Seungwon Shin, Jin-Woo Chung. <i>Shedding new light on the language of the dark web</i>. NAACL 2022 [8] ChaeHun Park, Eugene Jang, Wonsuk Yang, Jong C Park. <i>Generating negative samples by manipulating golden responses for unsupervised learning of a response evaluation model</i>. NAACL 2021 [9] Huije Lee, Wonsuk Yang, Chaehun Park, Hoyun Song, Eugene Jang, Jong C Park. <i>Optimizing Domain Specificity of Transformer-based Language Models for Extractive Summarization of Financial News Articles in Korean</i>. PACLIC 2021	
PUBLICATIONS (DOMESTIC)	[10] Dayeon Yim, Eugene Jang, Jin-Woo Chung. <i>Topic classification for cybercrime-related telegram messages using BERT</i>. KSC 2023 [11] Eugene Jang, Wonsuk Yang, Jong C Park. <i>Target-Agnostic Detection of Stances Toward Entities in News Articles</i>. HCIK 2021	

VULN. DISCLOSURE	[12] Daejin Lee, Seunghoe Kim, Donguk Kim, Eugene Jang . <i>CVE-2022–24548: Microsoft Defender Denial of Service Vulnerability</i> (Severity: Important). S2W Medium
THESIS	[13] Eugene Jang . <i>Detecting quoted claims and claim speakers in news articles using transformer-based language models</i> . Master's Thesis, 2021
PATENTS (DOMESTIC)	[14] (Approved) Eugene Jang , Jian Cui, Dayeon Yim, Youngjin Jin, Jin-Woo Chung, Seungwon Shin, Yongjae Lee. <i>A Method for Training a Language Model for Cybersecurity and A System for its Implementation</i> .
SERVICES	Program Committee / Reviewer, • COLING (2020, 2022, 2025) • LREC-COLING 2024 • EMNLP 2023 • ACL 2023 • ACL Rolling Review (2023, 2024, 2025) • Great Reviews: Oct 2023, Dec 2023, Feb 2024, Oct 2024 • COLM (2024, 2025)
TEACHING EXPERIENCE	Data Structure (CS 206) Fall 2020 Teaching Assistant, KAIST Natural Language Processing with Python (CS 372) Spring 2020 Teaching Assistant, KAIST Introduction to Programming (CS 101) Fall 2018, Spring 2019 Teaching Assistant, KAIST
SKILLS	Languages Korean (<i>native</i>), English (<i>native-level, preferred</i>), Mandarin Chinese (<i>intermediate</i>) Programming Python, C/C++, MATLAB, Java, Perl, Prolog.