# Takyoung Kim

## Ph.D. Student at the University of Illinois, Urbana-Champaign (UIUC)

⊕ youngerous.github.io  ✉ tk30@illinois.edu  in in/youngerous  🎓 Google Scholar

## Interests

My research interest lies in the intersection of Natural Language Processing and Communication. I am particularly interested in building communicative language models (LMs) that can understand social and cognitive context, evaluating preference during interactions, and exploring how LMs compose knowledge both from parametric and external sources.

## Education

**University of Illinois at Urbana-Champaign,** *Urbana, IL, USA* ↗                    Aug 2024 - Present
Doctor of Philosophy, Computer Science
Advisor: Dilek Hakkani-Tür

**Korea University,** *Seoul, South Korea* ↗                    Sep 2020 - Aug 2022
Master of Engineering, Industrial Management Engineering
Advisor: Pilsung Kang
Thesis: *Relaxing Unrealistic Assumptions of Dataset and Evaluation Metric in Dialogue State Tracking*

**Sungkyunkwan University,** *Seoul, South Korea* ↗                    Mar 2013 - Aug 2020
Bachelor of Science in Engineering, Interdisciplinary Program of Convergent Software
Bachelor of Media & Communication *(Leave of absence for military service: Feb 2015 - Nov 2016)*

## Publications                                                         * indicates equal contribution

[1] **Revealing User Familiarity Bias in Task-Oriented Dialogue via Interactive Evaluation** 📄
Takyoung Kim, Jamin Shin, Young-Ho Kim, Sanghwan Bae, Sungdong Kim
*Workshop on NLP for Conversational AI*                                    `NLP4ConvAI, ACL 2024`

[2] **DSTEA: Improving Dialogue State Tracking via Entity Adaptive Pre-training** 📄
Yukyung Lee, Takyoung Kim, Hoonsang Yoon, Pilsung Kang, Junseong Bang, Misuk Kim
*Knowledge-Based Systems (Q1, Impact Factor: 8.8)*                                    `KBS 2024`
*Workshop on Knowledge Augmented Methods for NLP*                              `KnowledgeNLP, KDD 2023`

[3] **KoSBi: A Dataset for Mitigating Social Bias Risks Towards Safer Large Language Model Applications** 📄 </>
Hwaran Lee*, Seokhee Hong*, Joonsuk Park, Takyoung Kim, Gunhee Kim, Jung-Woo Ha
*Annual Meeting of the Association for Computational Linguistics (Industry Track)*                 `ACL 2023`

[4] **SQuARe: A Large-Scale Dataset of Sensitive Questions and Acceptable Responses Created Through Human-Machine Collaboration** 📄 </>
Hwaran Lee*, Seokhee Hong*, Joonsuk Park, Takyoung Kim, Meeyoung Cha, Yejin Choi, Byoungpil Kim, Gunhee Kim, Eun-Ju Lee, Yong Lim, Alice Oh, Sangchul Park, Jung-Woo Ha
*Annual Meeting of the Association for Computational Linguistics (Oral, Best Paper Nomination)*        `ACL 2023`

[5] **Oh My Mistake!: Toward Realistic Dialgoue State Tracking including Turnback Utterances** 📄
Takyoung Kim, Yukyung Lee, Hoonsang Yoon, Pilsung Kang, Junseong Bang, Misuk Kim
*Towards Semi-Supervised and Reinforced Task-Oriented Dialog Systems*              `SereTOD, EMNLP 2022`

[6] **Mismatch between Multi-turn Dialogue and its Evaluation Metric in Dialogue State Tracking** 📄
Takyoung Kim, Hoonsang Yoon, Yukyung Lee, Pilsung Kang, Misuk Kim
*Annual Meeting of the Association for Computational Linguistics (Short)*                 `ACL 2022`

[7] **Building an Integrated Framework of Korean Text Summarization and Text-to-Speech** 📄 </>
Takyoung Kim, Jina Kim, Hyeongwon Kang, Subin Kim, Pilsung Kang
*Journal of Korean Institute of Industrial Engineers*                                    `JKIIE 2022`

**Preprint & Under Review**

[8] **Learning to Explore and Select for Coverage-Conditioned Retrieval-Augmented Generation** 📄 </>
Takyoung Kim, Kyungjae Lee, Young Rok Jang, Ji Yong Cho, Gangwoo Kim, Minseok Cho, Moontae Lee
*Preprint*

## Research Experience

**Korea University | Data Science & Business Analytics Lab,** *Seoul, South Korea* ⬀  Apr 2024 - Jul 2024
*Research Associate* (Advisor: Pilsung Kang)
Worked on retrieval-augmented language models specialized in Korean digital media

**LG AI Research | Advanced Machine Learning Lab,** *Seoul, South Korea* ⬀  Apr 2023 - Mar 2024
*Research Intern* (Mentors: Kyungjae Lee, Moontae Lee)
Worked on improving long-form responses of retrieval-augmented language models via planning [8]

**NAVER AI Lab | Language Research Team,** *Seongnam, South Korea* ⬀  Jul 2022 - Jan 2023
*Research Intern* (Mentor: Sungdong Kim)
Worked on revealing innate bias in task-oriented dialogue research via interactive study [1]
Worked on constructing safety benchmark for sensitive questions [4] and social biases [3]

**Korea University | Data Science & Business Analytics Lab,** *Seoul, South Korea* ⬀  Sep 2020 - Jun 2022
*Research Assistant* (Advisor: Pilsung Kang)
Worked on analyzing realistic biases and limitations of dialogue state tracking [6, 5, 2]
Developed algorithms and pipelines in various domains, such as news [7], car, and ocean waves

**Korea University | Data Science & Business Analytics Lab,** *Seoul, South Korea* ⬀  Jan 2020 - Aug 2020
*Undergraduate Research Intern* (Advisor: Pilsung Kang)
Designed spatial/temporal architecture reflecting optical flow for video deepfake detection

## Presentations and Talks

**Selective Masking for Task-Adaptive Pre-Training**
*Korean Institute of Industrial Engineers Conference (Spring)*  Jun 2021

**A Life of Turning into Software**
*Software Edu Fest, Naver Connect Foundation*  Jul 2020

## Honors and Awards

**Best Student Feedback Award | Research Ideation Contest**  Korea University  Feb 2021, Dec 2021
**2nd Place (2,500 USD) | MyData Acting Hackathon**  Ministry of Science and ICT  Oct 2019
**Merit-based Scholarship | Computer Science Coursework**  Sungkyunkwan University  Spring 2018, Fall 2018

## Teaching Experience

**T02459101 - Natural Language Processing**  Hankuk University of Foreign Studies  Feb 2024 - Jun 2024
*Teaching Assistant (Instructor: Yukyung Lee)*

**Industrial Machine Learning**  LG Electronics  Jan 2022 - Apr 2022
*Head Teaching Assistant (Instructor: Pilsung Kang)*

**Research Mentoring**  Hyundai Mobis  Jul 2021 - Sep 2021
*Research Project Assistant (Instructor: Pilsung Kang)*

**Industrial Machine Learning and Deep Learning**  LG Innotek, SK Hynix  Jul 2021 - Aug 2021
*Teaching Assistant (Instructor: Pilsung Kang)*

## Academic Service

**Conference Reviewer**  NeurIPS Datasets & Benchmarks (2022, 2023, 2024), EMNLP (2023, 2024), EMNLP Industry Track (2023, 2024), EACL (2024)

**Journal Reviewer**  Knowledge-Based Systems (2023, 2024)

**Others**  ACL Rolling Review (2023, 2024)