

SOHEE YANG

Ph.D. Student/Research Scientist Intern @ UCL/Google DeepMind

London, England, United Kingdom ◇ Last Update: 18 June 2024

soheeyang.cs@gmail.com ◇ github.com/soheeyang ◇ soheeyang.github.io ◇ [@soheeyang_](https://twitter.com/soheeyang_)

EDUCATION

Ph.D. in Computer Science

Mar 2023 – Present

@ **UCL NLP**, Department of Computer Science, University College London, UK
& Language Team, **Google DeepMind, UK** (50/50 Time Split)

- Advisors: **Pontus Stenetorp** and **Sebastian Riedel** @ UCL
Sebastian Riedel and **Mor Geva** @ Google DeepMind

M.S. in Artificial Intelligence

Mar 2021 – Feb 2023

@ **Language & Knowledge Lab**, Kim Jaechul Graduate School of AI, KAIST, South Korea

- Advisor: **Minjoon Seo**

B.S. in Computer Science and Engineering (*Summa Cum Laude*)

Mar 2014 – Feb 2018

@ School of CSEE, Handong Global University, South Korea

- GPA: 4.45 / 4.5 (1st rank in CSEE)

EXPERIENCES

Tel Aviv University, Israel (Working remotely from UK)

Dec 2024 – Present

*Visiting Ph.D. Student (Supervisor: **Mor Geva**)*

- Research on Reasoning Models

Language Team, Google DeepMind, UK

Mar 2023 – Present

*Part-Time Research Scientist Intern during Ph.D. (Hosts: **Sebastian Riedel** and **Nora Kassner**)*

- Research on Latent Reasoning of Large Language Models

Language & Knowledge Lab, KAIST, South Korea

Mar 2021 – Feb 2023

*Research Assistant (Supervisor: **Minjoon Seo**)*

- Research on Prompt-based Learning, Parametric Knowledge Update of LLMs, Generative Retrieval

Clova AI, NAVER Corp., South Korea

Sep 2018 – Feb 2021

Research Engineer

- Research on Dialogue System, Information Retrieval, Open-Domain Question Answering
- Text Parser of OCR @ Parsing Team [**Demo**] Oct 2019 – Feb 2021
- NLU of AiCall (Dialogue System) @ DUET TF [**Demo**] [**Japn**] [**Korean**] Sep 2018 – Oct 2019

AI Team, Kakao Corp., South Korea

Jan 2018 – Jun 2018

Research Intern

- NLU and Synthetic Data Generator for Dialogue System @ Context Part

Deep Learning Lab, Handong Global University, South Korea

Jun 2016 – Dec 2017

*Undergraduate Researcher (Supervisor: **In-Jung Kim**)*

- Deep Learning-based Dialogue State Tracker and Intent Classifier for Dialogue Systems
- Smart T-Commerce Service Using Big Data and Deep Learning [**Paper**] [**Patent**]

PUBLICATIONS

* indicates equal contribution.

Preprints

- [1] **Sohee Yang**, Sang-Woo Lee, Nora Kassner, Daniela Gottesman, Sebastian Riedel, Mor Geva. **How Well Can Reasoning Models Identify and Revert from Unhelpful Thoughts?**. *arXiv*, June 2025. [[Paper](#)] [[Tweet](#)]
- [2] Hyeonbin Hwang*, Byeongguk Jeon*, Seungone Kim, Jiyeon Kim, Hyeon Chang, **Sohee Yang**, Seungpil Won, Dohaeng Lee, Youbin Ahn, Minjoon Seo. **Let’s Predict Sentence by Sentence**. *arXiv*, May 2025. [[Paper](#)] [[Tweet](#)]
- [3] Hyeon Chang*, Jinho Park*, Hanseul Cho, **Sohee Yang**, Miyoung Ko, Hyeonbin Hwang, Seungpil Won, Dohaeng Lee, Youbin Ahn, Minjoon Seo. **The Coverage Principle: A Framework for Understanding Compositional Generalization**. *arXiv*, May 2025. [[Paper](#)] [[Tweet](#)]
- [4] Dongkeun Yoon, Seungone Kim, **Sohee Yang**, Sunkyoung Kim, Soyeon Kim, Yongil Kim, Eunbi Choi, Yireun Kim, Minjoon Seo. **Reasoning Models Better Express Their Confidence**. *arXiv*, May 2025. [[Paper](#)] [[Tweet](#)]

Peer-Reviewed Conference/Journal Papers

- [1] **Sohee Yang**, Nora Kassner, Elena Gribovskaya, Sebastian Riedel*, Mor Geva*. **Do Large Language Models Perform Latent Multi-Hop Reasoning without Exploiting Shortcuts?**. In *Findings of ACL*, 2025. [[Paper](#)] [[Code](#)] [[Dataset](#)] [[Tweet](#)]
- [2] Hyeon Chang, Jinho Park, Seonghyeon Ye, **Sohee Yang**, Youngkyung Seo, Du-Seong Chang, Minjoon Seo. **How Do Large Language Models Acquire Factual Knowledge During Pre-training?**. In *NeurIPS*, 2024. [[Paper](#)] [[Code](#)] [[Tweet](#)]
- [3] Eden Biran, Daniela Gottesman, **Sohee Yang**, Mor Geva, Amir Globerson. **Hopping Too Late: Exploring the Limitations of Large Language Models on Multi-Hop Queries**. In *EMNLP*, 2024. [[Paper](#)] [[Code](#)]
- [4] Soyoung Yoon*, Chaeun Kim*, Hyunji Lee, Joel Jang, **Sohee Yang**, Minjoon Seo. **Exploring the Practicality of Generative Retrieval on Dynamic Corpora**. In *EMNLP*, 2024. [[Paper](#)] [[Tweet](#)]
- [5] **Sohee Yang**, Elena Gribovskaya, Nora Kassner, Mor Geva*, Sebastian Riedel*. **Do Large Language Models Latently Perform Multi-Hop Reasoning?**. In *ACL*, 2024. [[Paper](#)] [[Code](#)] [[Dataset](#)] [[Poster](#)] [[Slides](#)] [[Tweet](#)]
- [6] **Sohee Yang**, Jonghyeon Kim, Joel Jang, Seonghyeon Ye, Hyunji Lee, Minjoon Seo. **Improving Probability-based Prompt Selection Through Unified Evaluation and Analysis**. *Transactions of the Association for Computational Linguistics (ACL)*, presented in *ACL*, 2024. [[Paper](#)] [[Code](#)] [[Poster](#)] [[Slides](#)] [[Tweet](#)]
- [7] Seonghyeon Ye, Hyeonbin Hwang, **Sohee Yang**, Hyeon Yu Yun, Yireun Kim, Minjoon Seo. **Investigating the Effectiveness of Task-Agnostic Prefix Prompt for Instruction Following**. In *AAAI*, 2024. [[Paper](#)] [[Code](#)] [[Tweet](#)]
- [8] Joel Jang, Dongkeun Yoon, **Sohee Yang**, Sungmin Cha, Moontae Lee, Lajanugen Logeswaran, Minjoon Seo. **Knowledge Unlearning for Mitigating Privacy Risks in Language Models**. In *ACL*, 2023. [[Paper](#)] [[Code](#)] [[Tweet](#)]
- [9] Hyunji Lee, Jaeyoung Kim, Hyeon Chang, Hanseok Oh, **Sohee Yang**, Vlad Karpukhin, Yi Lu, Minjoon Seo. **Nonparametric Decoding for Generative Retrieval**. In *Findings of ACL*, 2023. [[Paper](#)] [[Code](#)] [[Tweet](#)]

- [10] Hyunji Lee, **Sohee Yang**, Hanseok Oh, Minjoon Seo. **Generative Multi-hop Retrieval**. In *EMNLP*, 2022. [[Paper](#)] [[Code](#)] [[Tweet](#)]
- [11] Joel Jang*, Seonghyeon Ye*, Changho Lee, **Sohee Yang**, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Minjoon Seo. **TemporalWiki: A Lifelong Benchmark for Training and Evaluating Ever-Evolving Language Models**. In *EMNLP*, 2022. [[Paper](#)] [[Code](#)] [[Tweet](#)]
- [12] Joel Jang, Seonghyeon Ye, **Sohee Yang**, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Stanley Jungkyu Choi, Minjoon Seo. **Towards Continual Knowledge Learning of Language Models**. In *ICLR*, 2022. [[Paper](#)] [[Code](#)] [[Tweet](#)]
- [13] Wonseok Hwang, Jinyeong Yim, Seunghyun Park, **Sohee Yang**, Minjoon Seo. **Spatial Dependency Parsing for Semi-Structured Document Information Extraction**. In *Findings of ACL*, 2021. [[Paper](#)]
- [14] Sewon Min et al. **NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned**. *Proceedings of Machine Learning Research (PMLR)*, 2021. [[Paper](#)] [[Project](#)]
- [15] **Sohee Yang** and Minjoon Seo. **Designing a Minimal Retrieve-and-Read System for Open-Domain Question Answering**. In *NAACL*, 2021. [[Paper](#)] [[Code](#)] [[Poster](#)] [[Slides](#)] [[Video](#)]
- [16] **Sohee Yang** and Minjoon Seo. **Is Retriever Merely an Approximator of Reader?**. In *Spa-NLP Workshop at ACL 2022*, 2020. [[Paper](#)] [[Poster](#)] [[Slides](#)] [[Video](#)] [[Model](#)]
- [17] Jung-Woo Ha*, Kihyun Nam*, Jingu Kang, Sang-Woo Lee, **Sohee Yang**, Hyunhoon Jung, Eunmi Kim, Hyeji Kim, Soojin Kim, Hyun Ah Kim, Kyoungtae Doh, Chan Kyu Lee, Nako Sung, Sunghun Kim. **ClovaCall: Korean Goal-Oriented Dialog Speech Corpus for Automatic Speech Recognition of Contact Centers**. In *INTERSPEECH*, 2020. [[Paper](#)] [[Code](#)]
- [18] Sungdong Kim, **Sohee Yang**, Gyuwan Kim, Sang-Woo Lee. **Efficient Dialogue State Tracking by Selectively Overwriting Memory**. In *ACL*, 2020. [[Paper](#)] [[Code](#)]
- [19] Sang-Woo Lee, Tong Gao, **Sohee Yang**, Jaejun Yoo, Jung-Woo Ha. **Large-Scale Answerer in Questioner’s Mind for Visual Dialog Question Generation**. In *ICLR*, 2019. [[Paper](#)] [[Code](#)] [[Poster](#)] [[Slides](#)]

HONORS, AWARDS & SCHOLARSHIPS

Honors & Awards

- 1st Place, [EfficientQA Competition](#) 500MB Track Human Eval, NeurIPS 2020 Dec 2020
- Gyeongsangbuk-do Province Governor’s Commendation, HGU Feb 2018
- *Summa Cum Laude*, HGU Feb 2018
- 1st Place (Grand Prize), Senior Capstone Design Competition, HGU Dec 2017

Scholarships

- Full tuition funded by Google DeepMind at UCL Mar 2023 – Present
- 92% of tuition covered by the school at KAIST Mar 2021 – Feb 2023
- Full tuition covered by the school and Korean Government at HGU (merit-based) Mar 2014 – Feb 2018
 - National Science and Engineering Scholarship (Government of South Korea)
 - Merit Scholarship for Academic Excellence (top 1% for 4 times)
 - Entrance Scholarship (top 5%)

TALKS

- Do Large Language Models Latently Perform Multi-Hop Reasoning?** Jun 2025
- [Deep Learning: Classics and Trends](#), ML Collective (Remote)

Latent Multi-Hop Reasoning of Large Language Models	Dec 2024
· Institute for Language, Cognition and Computation (ILCC) , The University of Edinburgh, UK	
Interpreting Large Language Models: From Historical Developments to Modern Approaches	
· The University of Edinburgh, UK	Dec 2024
· CS475: ML for NLP , KAIST, South Korea (Remote)	Nov 2024
Latent Deductive Reasoning of Large Language Models	Apr 2024
· The University of Sheffield, UK	
EfficientQA & Designing a Minimal Retrieve-and-Read System for Open-Domain Question Answering	Dec 2020
· Tech Talk, Upstage AI, South Korea (Remote)	
DUET (AiCall) and Potential Needs in Dialog	May 2019
· NAVER LABS Europe × Clova AI Workshop, NAVER LABS Europe, France	
Large-Scale Answerer in Questioner’s Mind for Visual Dialog Question Generation	May 2019
· Naver Expo Lunch Talk, ICLR 2019, United States	
Deep Learning-based Dialogue System	Aug 2018
· Summer School, Deep Learning Society: Deepest, South Korea	

SERVICES

Reviewer

- EMNLP 2025 (ARR May), ICML Workshop 2025, ACL 2025 (ARR February), ARR December 2024, NeurIPS Workshops 2024, EMNLP 2024 (ARR June), ACL Workshops 2024, ICML Workshop 2024, ACL 2024 (ARR Feb), ICLR 2023, ACL Workshop 2022

Secondary Reviewer

- ICML 2019, ICLR 2019

STUDENT ASSOCIATIONS

Deep Learning Society: Deepest

Seoul National University

Jul 2018 – Feb 2021

- Director

Feb 2019 – Jun 2019

S/W Development Club: Seulgijjagi

Handong Global University

Feb 2014 – Feb 2018

- President
- President

Feb 2016 – Jun 2016

Feb 2015 – Jun 2015

TEACHING EXPERIENCE

Teaching Assistant

Handong Global University

- C Programming
- Operating Systems

Aug 2017 – Dec 2017

Feb 2017 – Jun 2017

- Computer Architecture and Organization
- Logic Design
- C Programming

Aug 2016 – Dec 2016

Feb 2016 – Jun 2016

Dec 2015 – Jan 2016

LANGUAGE PROFICIENCY

- Korean: Native Level Fluency
- English: High Level Fluency
 - IELTS Academic Overall 8.0 (taken on 31 Dec 2022)
 - TOEFL iBT 109 (taken on 02 Sep 2018)
 - GRE Writing 4.5 (taken on 19 Nov 2017)