

# Yeon Seonwoo

yeon.seonwoo95@gmail.com

yeonsw.github.io

EDUCATION	<b>KAIST</b> <b>Ph.D. Student in School of Computing</b> Advisor: Alice Oh Thesis title: "Distant-Supervision for Question Answering"	Daejeon, Republic of Korea Sep 2017 - Aug 2022
	<b>KAIST</b> <b>M.S. in School of Computing</b> Advisor: Alice Oh Thesis title: "Event Prediction Using Vector Representation in Hawkes Processes"	Daejeon, Republic of Korea Mar 2016 - Aug 2017
	<b>Sungkyunkwan University</b> <b>B.S. in Computer Science &amp; Engineering</b>	Suwon, Republic of Korea Mar 2012 - Feb 2016
EXPERIENCE	<b>Applied Scientist</b> <b>Amazon (Manhattan, NY)</b>	Jul 2023 - Current
	<b>Applied Scientist Intern</b> <b>Amazon</b>	Mar 2022 - Jun 2022
	<b>NLP Research Intern</b> <b>Adobe Research</b>	Sep 2021 - Dec 2021
AWARDS	<b>Naver Ph.D. Fellowship, 2020</b>	
PUBLICATION	Juhee Son, <b>Yeon Seonwoo</b> , Alice Oh, James Thorne, Seunghyun Yoon, "Multi-hop Database Reasoning with Virtual Knowledge Graph", <i>Workshop on Knowledge Graphs and Large Language Models 2024</i> .	
	Chani Jung, Dongkwan Kim, Jiho Jin, Jiseon Kim, <b>Yeon Seonwoo</b> , Yejin Choi, Alice Oh, Hyunwoo Kim, "Perceptions to Beliefs: Exploring Precursory Inferences for Theory of Mind in Large Language Models", <i>EMNLP 2024</i> .	
	<b>Yeon Seonwoo</b> , Guoyin Wang, Sajal Choudhary, Changmin Seo, Jiwei Li, Xiang Li, Puyang Xu, Sunghyun Park, Alice Oh. "Ranking-Enhanced Unsupervised Sentence Representation Learning." <i>ACL 2023</i>	
	<b>Yeon Seonwoo</b> , Seunghyun Yoon, Franck Dernoncourt, Trung Bui, Alice Oh. "Virtual Knowledge Graph Construction for Zero-Shot Domain-Specific Document Retrieval." <i>COLING 2022</i>	
	Changyoon Lee, <b>Yeon Seonwoo</b> , and Alice Oh. "CS1QA: A Dataset for Assisting Code-based Question Answering in an Introductory Programming Course." <i>NAACL 2022</i>	
	<b>Yeon Seonwoo</b> <sup>*</sup> , Juhee Son <sup>*</sup> , Jiho Jin, Sang-Woo Lee, Ji-Hoon Kim, Jung-Woo Ha, and Alice Oh. "Two-Step Question Retrieval for Open-Domain QA." <i>ACL-Findings. 2022</i> .	
	<sup>*</sup> Contributed equally	
	<b>Yeon Seonwoo</b> , Sang-Woo Lee, Ji-Hoon Kim, Jung-Woo Ha, and Alice Oh. "Weakly Supervised Pre-Training for Multi-Hop Retriever." <i>ACL-Findings. 2021</i> . Code URL: <a href="https://github.com/yeonsw/LOUVRE">https://github.com/yeonsw/LOUVRE</a>	
	<b>Yeon Seonwoo</b> , Ji-Hoon Kim, Jung-Woo Ha, and Alice Oh. "Context-Aware An-	

swer Extraction in Question Answering.” *EMNLP (long)*. 2020.  
Code URL: <https://github.com/yeonsw/BLANC>

**Yeon Seonwoo**, Sungjoon Park, Dongkwan Kim, and Alice Oh. “Additive Compositionality of Word Vectors.” *Workshop on Noisy User-generated Text at EMNLP*. 2019.  
Code URL: <https://github.com/yeonsw/OLIVE>

Sungjoon Park, **Yeon Seonwoo**, Jiseon Kim, Jooyeon Kim and Alice Oh. “Denoising Recurrent Neural Networks for Classifying Crash-Related Events.” *IEEE Transactions on Intelligent Transportation Systems*. 2019.

**Yeon Seonwoo**, Sungjoon Park, and Alice Oh. “Hierarchical Dirichlet Gaussian Marked Hawkes Process for Narrative Reconstruction in Continuous Time Domain.” *EMNLP (long)*. 2018.

## RESEARCH PROJECTS

**Multi-Hop Question Answering (2020-2021)**  
Modeling the reasoning process in multi-hop question answering.  
[Naver Clova]

**Context-Aware Answer Extraction in Question Answering (2019-2020)**  
Alleviating the discrepancy in extractive QA models that they sometimes predict answer-spans in irrelevant context.  
[Naver Clova]

**Theoretical Analysis on Compositionality of Word Embedding Models (2018-2019)**  
Proving additive compositionality of Skip-Gram; Modeling a novel word embedding model based on our theorem.  
[Naver Clova]

**Document Clustering with Poisson Point Process (2017-2018)**  
Clustering news articles with topic information and publication time information of each article.  
[NCSoft]

## TALK

Feb. 08. 2022, Weakly-Supervised Pre-Training for Multi-Hop Retriever, KISTI  
Dec. 22. 2021, Weakly-Supervised Pre-Training for Multi-Hop Retriever, Naver Clova TechTalk  
Nov. 22. 2021, Weakly-Supervised Pre-Training for Multi-Hop Retriever, LG AI Research TechTalk  
Jul. 01. 2020, Context-Aware Answer Extraction in Question Answering, Naver Clova TechTalk  
Dec. 14. 2019, Additive Compositionality of Word Vectors, Naver Clova TechTalk  
Jan. 25. 2019, Hierarchical Dirichlet Gaussian Marked Hawkes Process for Narrative Reconstruction in Continuous Time Domain, NCSoft AI Day

## ACADEMIC SERVICE

**Reviewer**  
ACL Rolling Review (Oct 2021, Nov 2021, Jan 2022, Apr 2022, June 2022, July 2022, Sep 2022), NeurIPS2022, EMNLP 2022, ICLR 2022, NeurIPS 2021, EMNLP 2021, ACL 2021, ICLR 2021

## TEACHING EXPERIENCE

Machine Learning for Social Science, TA  
Mar 2019 - Aug 2019  
  
Artificial Intelligence & Machine Learning, TA  
Mar 2017 - Aug 2017

Data Structure, TA  
Sep 2016 - Feb 2017 (TA Excellence Award), Sep 2017 - Feb 2018, Mar 2018 - Aug 2018, Sep 2018 - Feb 2019

**SCHOLARSHIP** Korea National Science&Technology Scholarship (Mar 2012 - Feb 2016)

**COMPUTER SKILLS** **Languages:** Python, Java, C, C++,  $\text{\LaTeX}$   
**Deep Learning Libraries:** TensorFlow, PyTorch  
**Web Development:** HTML, CSS, JavaScript  
**Applications:** Vi/Vim, Git, Docker

**REFERENCES** Prof.Alice Oh, School of Computing, KAIST, [alice.oh@kaist.edu](mailto:alice.oh@kaist.edu)