

# Donghyeon Kim

Ph.D. in Computer Science and Engineering  
Senior Research Engineer at AIRS, Hyundai Motor Company

[donghyeonk@me.com](mailto:donghyeonk@me.com)  
<https://donghyeonk.github.io>

## INTERESTS

---

Language Modeling, Question Answering, Information Retrieval

## EDUCATION

---

**Korea University, Seoul** March 2011 - Feb 2020  
Ph.D. in Computer Science and Engineering (On leave: startups, Oct 2014 - Mar 2018)  
Advisor: Prof. Jaewoo Kang

**Korea University, Seoul** Mar 2003 - Feb 2011  
B.S. in Computer Science Education (On leave: military service, Aug 2006 - Jul 2008)  
Advisor: Prof. Seongbin Park

## WORK EXPERIENCE

---

**Hyundai Motor Company** Dec 2019 - Present  
Senior Research Engineer Seoul, Korea

- Hyperscale AI: Transformer-based language models
- Question Answering: Passage retrieval & reranking, reading comprehension
- Machine Translation: H-Translator

**Konolabs, Inc.** Apr 2016 - Mar 2018 (Dispatch from Korea Univ. Apr 2016 - Dec 2016)  
Data Scientist & Software Engineer Seoul, Korea

- Event start time prediction by learning online calendar event titles, week contexts and user preferences
- Scheduling related sentence classification
- Email bot trigger, architecture design & implementation

**Opinion8, Inc.** Oct 2014 - Apr 2016 (Dispatch from Korea Univ.)  
CTO & Co-founder Seoul, Korea

- Opinion, location, and time based entity(restaurant, attraction, hotel) search engine implementation
- Sentiment analysis using dependency trees and part-of-speech tags

## SELECTED PUBLICATIONS

---

**BERN2: an advanced neural biomedical named entity recognition and normalization tool**  
Mujeen Sung, Minbyul Jeong, Yonghwa Choi, **Donghyeon Kim**, Jinhyuk Lee and Jaewoo Kang  
Bioinformatics, Aug 2022, <http://bern2.korea.ac.kr>

**Relevance Similarity Scorer and Entity Guided Reranking for Knowledge Grounded Dialog System**  
Hyunkyung Bae, Minwoo Lee, Ahhyeon Kim, Hwanhee Lee, Cheongjae Lee, Cheoneum Park, **Donghyeon Kim**, Kyomin Jung  
DSTC9 Workshop at AACL, Feb 2021

**Transferability of Natural Language Inference to Biomedical Question Answering**  
Minbyul Jeong, Mujeen Sung, Gangwoo Kim, **Donghyeon Kim**, Wonjin Yoon, Jaehyo Yoo and Jaewoo Kang  
BioASQ Lab Workshop at CLEF, Sep 2020, cited by 19+

### **Building a PubMed knowledge graph**

Jian Xu, Sunkyu Kim, Min Song, Minbyul Jeong, **Donghyeon Kim**, Jaewoo Kang, Justin F. Rousseau, Xin Li, Weijia Xu, Vetle I. Torvik, Yi Bu, Chongyan Chen, Islam Akef Ebeid, Daifeng Li and Ying Ding  
Scientific Data, Jun 2020, cited by 65+

### **Pre-trained Language Models for Biomedical Question Answering**

Wonjin Yoon, Jinhyuk Lee, **Donghyeon Kim**, Minbyul Jeong and Jaewoo Kang  
BioASQ Workshop at ECML PKDD, Sep 2019, cited by 55+

### **BioBERT: a pre-trained biomedical language representation model for biomedical text mining**

Jinhyuk Lee<sup>†</sup>, Wonjin Yoon<sup>†</sup>, Sungdong Kim, **Donghyeon Kim**, Sunkyu Kim, Chan Ho So and Jaewoo Kang  
Bioinformatics, Sep 2019, cited by 2515+

### **A Neural Named Entity Recognition and Multi-Type Normalization Tool for Biomedical Text Mining**

**Donghyeon Kim**, Jinhyuk Lee, Chan Ho So, Hwisang Jeon, Minbyul Jeong, Yonghwa Choi, Wonjin Yoon, Mujeen Sung and Jaewoo Kang  
IEEE Access, Jun 2019, cited by 78+, <https://bern.korea.ac.kr>

### **Learning User Preferences and Understanding Calendar Contexts for Event Scheduling**

**Donghyeon Kim**<sup>†</sup>, Jinhyuk Lee<sup>†</sup>, Donghee Choi, Jaehoon Choi and Jaewoo Kang  
CIKM (full, oral), Oct 2018, cited by 5+

### **BEST: Next-Generation Biomedical Entity Search Tool for Knowledge Discovery from Biomedical Literature**

Sunwon Lee<sup>†</sup>, **Donghyeon Kim**<sup>†</sup>, Kyubum Lee, Jaehoon Choi, Seongsoon Kim, Minji Jeon, Sangrak Lim, Donghee Choi, Sunkyu Kim, Aik-Choon Tan and Jaewoo Kang  
PLoS One, Oct 2016, cited by 87+, <http://best.korea.ac.kr>

### **Smith Search: Opinion-Based Restaurant Search Engine**

Jaehoon Choi, **Donghyeon Kim**, Donghee Choi, Sangrak Lim, Seongsoon Kim, Youngjae Choi and Jaewoo Kang  
WWW (demo), May 2015, cited by 2+

### **Consento: A Consensus Search Engine for Answering Subjective Queries**

Jaehoon Choi, **Donghyeon Kim**, Seongsoon Kim, Junkyu Lee, Sangrak Im, Sunwon Lee and Jaewoo Kang  
WWW (poster), Apr 2012, cited by 2+

## **DOMESTIC PUBLICATIONS**

---

### **Efficient codebook expressions of word representations in language modeling**

Byungjoo Kim, Yookyung Koh, Chan Ho So, **Donghyeon Kim** and Jaewoo Kang  
Korea Computer Congress, Jun 2018

### **Utilizing Social Q&A Data for Decision Making**

Jukyoung Lee, Jaehoon Choi, **Donghyeon Kim**, Donghee Choi, Minhwan Yoo and Jaewoo Kang  
The 42nd Annual Meeting and Winter Conference of KIISE, Dec 2015

## **AWARDS**

---

### **Winner, biomedical semantic question answering (Task 8b, Phase B, Exact/ideal answers), BioASQ Challenge**

Minbyul Jeong, Mujeen Sung, Gangwoo Kim, **Donghyeon Kim**, Wonjin Yoon, Jaehyo Yoo and Jaewoo Kang  
Sep 2020

### **Winner, biomedical semantic question answering (Task 7b, Phase B, Exact answers), BioASQ Challenge**

Wonjin Yoon, Jinhyuk Lee, **Donghyeon Kim**, Minbyul Jeong and Jaewoo Kang  
Sep 2019

### **The grand prize, Start-up competition, Korea University**

Jaehoon Choi, **Donghyeon Kim** and Donghee Choi  
Jan 2015