
Hideaki Joko

NLP and IR researcher based in the Netherlands

ABOUT

Hideaki Joko is a research staff and PhD candidate at the Institute for Computer and Information Science, Radboud University, Netherlands. He received his MSc from the University of Tokyo for his research on natural language processing (NLP), and BEng in applied mathematics from Waseda University. He was a research engineer at Mitsubishi Electric Information and Technology R&D Center, Japan, from 2016 to 2020.

His research interests include information retrieval (IR) and NLP, and he has 6+ years of research and 4+ years of hands-on experience in these fields.

He has published 10+ papers in conferences and peer-reviewed journals, and filed several patents on IR and NLP algorithms and software. He received five awards including the incentive research award for his NLP research from the Japanese Society of Artificial Intelligence (JSAIL) Workshop on Access and Visual Mining, and the Mitsubishi Electric R&D Center president award for making the design process efficient by developing an IR algorithm and software.

SKILLS

- **Programming Languages:** Python, C#, Java, SPARQL, etc.
- **Software Libraries:** PyTorch, Keras, Scikit-learn, Pandas, MySQL, Numpy, Elasticsearch, etc.
- **Operating Systems:** Linux, Mac OS, Windows
- **Languages:** English and Japanese

EXPERIENCE

Radboud University, Netherlands - *Research Staff*

September 2020 - PRESENT

- Project: conversational search
- Created the conversational search dataset using crowdsourcing, and published it at the SIGIR 2021 conference.

Mitsubishi Electric, Japan - *Research Engineer*

April 2016 - August 2020

- Researched and developed deep-learning-based document search algorithm and software with small computational complexity to solve the word sense disambiguation problem
- Researched and developed an algorithm of deep-learning-based question answering system
- Developed a text-based automatic error diagnosis algorithm and software
- Led an R&D project in collaboration with a manufacturing division, and developed a C#-based information retrieval software to make a design process efficient

EDUCATION

Radboud University, Netherlands - *PhD, Data Science*

September 2020 - August 2024

Fully-funded PhD program, focusing on conversational search and dialogue system.

University of Tokyo, Japan - *MSc, Computing System, Multidisciplinary Science, GPA 3.74/4.00*

March 2016

Master's thesis: Automatic Synonym Acquisition Using a Context-Restricted Skip-gram Model

Waseda University, Japan - *BEng. Applied Mathematics*

March 2014

Bachelor's thesis: Accent Detection based on Optimality Theory

PROFESSIONAL DEVELOPMENT

Other than research, his interest is in solving business challenges using his expertise in NLP and IR. He has 200+ hours of data analysis, marketing, and project management education from several institutes including Wharton Online and Mitsubishi Electric.

CERTIFICATIONS

Data Analysis

- A Crash Course in Causality: Inferring Causal Effects from Observational Data by University of Pennsylvania on Coursera
- Bayesian Statistics: From Concept to Data Analysis by University of California on Coursera

Marketing

- Entrepreneurship 1: Developing the Opportunity by Wharton Online
- Financial Markets by Yale University on Coursera
- Introduction to Marketing by Wharton Online

AWARDS

Received five awards including:

- Incentive Research Award, The Japanese Society for Artificial Intelligence Workshop on Access and Visual Mining, 2019.
 - Our research paper "*Intention Understanding with Small Training Data Sets by Utilizing Multi-Task Transfer Learning*" was selected as one of the best two out of more than 20 papers.
- Mitsubishi Electric R&D Center president award for making the design process efficient by developing an IR algorithm and software, 2019.

FIRST AUTHOR PUBLICATIONS

Refereed Paper for International Conference

- Conversational Entity Linking: Problem Definition and Datasets, Hideaki Joko, et. al., The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2021.
- Learning Word Embeddings Using Spatial Information, Hideaki Joko, et. al., IEEE International Conference on System, Man, and Cybernetics, 2019.
- Intention Understanding in Small Training Data Sets by Using Transfer Learning, Hideaki Joko, et. al., Proceedings of The 11th International Conference on Mobile Computing and Ubiquitous Networking, 2018.

Refereed Journal Paper (In Japanese)

- Accelerating Contextualized Representation based Document Retrieval Using Approximate Nearest Neighbor Search, Hideaki Joko, et. al., IEICE Transactions on Information and Systems, 2020.
- Automatic Synonym Acquisition Using a Context-Restricted Skip-gram Model, Hideaki Joko, et. al., Journal of Natural Language Processing, 2017.

Domestic Conference (In Japanese)

- Learning Word Embeddings Using Spatial Information, Hideaki Joko, et. al., Proceedings of the Japanese Society for Artificial Intelligence Workshop on Access and Visual Mining, 2019.
- Intention Understanding with Small Training Data Sets by Utilizing Multi-Task Transfer Learning, Hideaki Joko, et. al., Proceedings of the Japanese Society for Artificial Intelligence Workshop on Access and Visual Mining, 2018. Incentive Research Award.
- Intention Understanding in Small Training Data Sets by Using Transfer Learning, Hideaki Joko, et. al., Proceedings of the 2018 IEICE General Conference, 2018.
- Automatic Synonym Acquisition Using a Context-Restricted Skip-gram Model, Hideaki Joko, et. al., Proceedings of the Annual Meeting of The Association for Natural Language Processing, 2016.
- Evaluation of Word Vectors by Synonym Identification, Hideaki Joko, et. al., Proceedings of the Japanese Society for Artificial Intelligence Workshop on Access and Visual Mining, 2015.

PATENTS

- INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING PROGRAM, Hideaki Joko, WO/2020/095373, Patent description, Pending Patent.
- LANGUAGE PROCESSING DEVICE, LANGUAGE PROCESSING SYSTEM AND LANGUAGE PROCESSING METHOD, Hideaki Joko, WO/2019/106758, Pending Patent. (A related application also pending).

TEST SCORES

Japanese Higher Civil Service Examination - Field: Engineering

Passed the exam and interview for a technocrat, one of the most difficult exams in Japan with a pass rate of about 6% in total. This proves his high ability in mathematics and engineering.