

PERSONAL INFORMATION

Haritz Puerto



 Daejeon, South Korea (willing to relocate to Germany)

 +82 010 9494 1576

 haritzpuerto94@gmail.com

 <https://haritzpuerto.github.io>

semanticscholar.org/author/Haritz-Puerto/1389023854

 [linkedin.com/in/haritzpuerto/](https://www.linkedin.com/in/haritzpuerto/)

Nationality Spanish

WORK EXPERIENCE

Jan. 2021 -

Data Scientist

Coleridge Initiative

Working on the Rich Context project developing a platform for dataset discovery from research publications. This project is the continuation of the Rich Context Competition I won in 2019 (see Projects and Awards Sections). <https://coleridgeinitiative.org/data-products/>

Aug. 2018 - Feb. 2021

Research Assistant

IR&NLP Lab@KAIST

Daejeon, South Korea

Research on Machine Learning and Natural Language Processing. Specifically, Question Answering, Question Generation, Unsupervised Learning, Transformers, and Graph Neural Networks. Lab seminar organization and server administration. <http://ir.kaist.ac.kr/about/>

Jul. 2015 – Jul. 2017

Software Engineer

CLAROFLEX

Malaga, Spain

Development of an web-based enterprise resource planning (ERP) using Java EE, Python, and MySQL. <http://www.claroflex.com/en/productos/claroflex-erp-4/>

EDUCATION

Aug. 2018 – Feb. 2021

MSc in Computer Science

GPA: 92.77/100

KAIST, South Korea

Korean Government Scholarship Program. <http://niied.go.kr/user/nd35203.do>

Aug. 2017 – Aug. 2018

Korean Language Program

Dongseo University, South Korea

Korean Government Scholarship Program

Aug. 2014 – Dec. 2014

Exchange Semester

Seoul National University, South Korea

Sep. 2012 – Jul. 2017

BSc in Computer Science and Engineering

GPA: 9.02/10

University of Malaga, Spain

Graduated with Honors

LANGUAGES

Mother tongue Spanish

Other languages	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	C1	C1	C1	C1	C1
Korean	B1	B1	B1	B1	B1
Test of Proficiency in Korean (TOPIK) Level 4					
French	A2	A2	A2	A2	A2

Levels: A1 and A2: Basic user – B1 and B2: Independent user – C1 and C2: Proficient user
[Common European Framework of Reference for Languages](#)

PROJECTS

MSc Thesis

A Hierarchical SRL Graph Network for Multi-Hop Question Answering (PDF submitted)

Dec. 2018 - Dec. 2020 **Exobrain**

Modeling multi-hop question answering using graph neural networks. Project funded by the Institute for Information & Communications Technology Planning & Evaluation (IITP) of the Korean Government. <http://exobrain.kr/pages/en/>

Oct. 2018 - Feb. 2019 **Rich Context Competition**

Won the honorable mention award in New York University's Coleridge Initiative Rich Context competition. The goal is to develop a machine learning model that identifies the datasets used in scientific publications. I solved the challenge modeling it as a Question Answering task and presented the project in the competition's workshop at New York University. <https://coleridgeinitiative.org/richcontext/richcontextcompetition/>

PUBLICATIONS

* indicates equal contribution.

- First author publication based on my MSc thesis to be submitted to EMNLP 2021.
- Junmo Kang*, Giwon Hong*, Haritz Puerto*, Sung-Hyon Myaeng. **Regularization of Distinct Strategies for Unsupervised Question Generation**. In Findings of EMNLP, 2020.
- Doyeon Lim*, Haritz Puerto*, Sung-Hyon Myaeng. **Analysis of the Semantic Answer Types to Understand the Limitations of MRQA Models**. Journal of KIISE (2020).
- Book chapter **Finding Datasets in Publications: The KAIST Approach** in Rich Search and Discovery for Research Datasets. SAGE 2020.
- Junmo Kang*, Haritz Puerto*, and Sung-Hyon Myaeng. **Let Me Know What to Ask: Interrogative-Word-Aware Question Generation**. In Proceedings of the 2nd Workshop on Machine Reading for Question Answering, EMNLP 2019.
- Haritz Puerto*, Doyeon Lim*, and Sung-Hyon Myaeng. **Analysis of Answer Type Application Ability of State-of-the-Art Reading Comprehension Models for Question Answering Task**. In Proceedings of the Korea Computer Congress (2019). Best Paper Award.

AWARDS

- Honorable Mention Award in NYU's Coleridge Initiative Rich Context Competition.
- Best Paper Award in the Korea Computer Congress 2019.
- Korean Government Scholarship Program (Korean language program and Master's Degree).
- European Institute of Innovation and Technology (EIT Digital) Master scholarship (Gracefully declined).
- Graduation with Honors from the Department of Computer Science at the University of Malaga.
- Honors in 14 courses at the University of Malaga (*Matricula de honor* in Spanish).
- Exchange Semester scholarship by the University of Malaga (to study at Seoul National University).
- Graduated with Honors from High School.