



Haritz Puerto

ML/NLP PH.D. CANDIDATE

LANGUAGES

Spanish ●●●●●●

English ●●●●●●

Korean ●●●●●●

French ●●●●●●

German ●●●●●●


ML STACK


- PyTorch
- HF's Transformers, PEFT, Datasets, Evaluate
- vLLM
- Neptune.AI/Wandb
- Slurm
- Azure


RESEARCH STACK


- Grant writing
- Project management
- Seminar organization
- Public speaking
- Teaching

SOCIAL

 [haritzpuerto.github.io](https://github.com/haritzpuerto)

 [Haritz Puerto](#)

 [Haritz Puerto](#)

 [@haritzpuerto](#)

REFERENCES

Prof. Iryna Gurevych, TU Darmstadt
Prof. Seong-Joon Oh, Univeristy of Tübingen

WORK EXPERIENCE

DOCTORAL RESEARCHER

UKP Lab @ TU Darmstadt, Germany Jun 2021 - Today

- Ph.D. research, grant writing, project management, seminar organization, teaching, student supervision.

RESEARCH INTERN

Parameter Lab, Germany Jul 2024 - Feb 2025

- Internship on trustworthy AI.

DATA SCIENTIST

Coleridge Initiative, Remote Jan 2021 - May 2021

- Co-organized the Kaggle Competition [Show US the Data](#).

RESEARCH ASSISTANT

IR&NLP Lab, KAIST, Korea Sep 2018 - Feb 2021

- Research on question answering, question generation and graph neural networks.

SOFTWARE ENGINEER

Claroflex, Spain Jul 2015 - Jul 2017

- Development of a web-based enterprise resource planning (ERP) using Java EE, Python, and MySQL.

EDUCATION

Ph.D in ML & NLP

TU Darmstadt, Germany Jun 2021 - Jan 2026 (expected)

- Ph.D. topic: Reasoning in LLMs.

M.Sc. in Computer Science

KAIST, South Korea Sep 2018 - Mar 2021

- Thesis Topic: A Hierarchical SRL Graph Network for Multi-Hop Question Answering.

B.Sc. in Computer Science

University of Malaga, Spain Sep 2012 - July 2027

- Graduated with Honors.
- Exchange semester at Seoul National University.

AWARDS

- Honorable Mention Award in NYU Coleridge Initiative's Rich Context Competition (2019).
- Best Paper Award in the Korea Computer Congress (2019).
- Korean Government Scholarship Program (Korean language program and Master's Degree; 2017).
- European Institute of Innovation and Technology (EIT Digital) Master scholarship (Gracefully declined; 2017).
- Graduation with Honors from the Department of Computer Science at the University of Malaga (2017).
- Exchange Semester scholarship to study at Seoul National University (2014).
- Graduated with Honors from High School (2012).

SELECTED PUBLICATIONS

* indicates equal contribution.

- **Haritz Puerto**, Martin Gubri, Tommaso Green, Seong Joon Oh, Sangdoo Yun. **C-SEO Bench: Does Conversational SEO Work?** arxiv 2025
- Tommaso Green, Martin Gubri, **Haritz Puerto**, Seong Joon Oh, Sangdoo Yun. **Leaky Thoughts: Large Reasoning Models Are Not Private Thinkers**. arxiv 2025
- **Haritz Puerto**, Tilek Chubakov, Xiaodan Zhu, Harish Tayyar Madabushi, Iryna Gurevych. **Fine-Tuning on Diverse Reasoning Chains Drives Within-Inference CoT Refinement in LLMs**. ACL 2025
- **Haritz Puerto**, Martin Gubri, Sangdoo Yun, Seong Joon Oh. **Scaling Up Membership Inference: When and How Attacks Succeed on Large Language Models**. Findings of NAACL 2024.
- **Haritz Puerto**, Martin Tutek, Somak Aditya, Xiaodan Zhu, Iryna Gurevych. **Code Prompting Elicits Conditional Reasoning Abilities in Text+Code LLMs**. EMNLP 2024.
- **Haritz Puerto**, Tim Baumgärtner, Rachneet Sachdeva, Haishuo Fang, Hao Zhang, Sewin Tariverdian, Kexin Wang, Iryna Gurevych. **UKP-SQuARE v3: A Platform for Multi-Agent QA Research**. ACL 2023 Demo Track
- **Haritz Puerto**, Gözde Gül Şahin, Iryna Gurevych. **MetaQA: Combining Expert Agents for Multi-Skill Question Answering**. EACL, 2023.
- Junmo Kang*, Giwon Hong*, **Haritz Puerto***, Sung-Hyon Myaeng. **Regularization of Distinct Strategies for Unsupervised Question Generation**. 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).
- 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**.

Full list of publications in <https://www.semanticscholar.org/author/Haritz-Puerto-San-Roman/1389023854>.

STUDENT SUPERVISION

- Yichen Xie: Uncertainty-guided Reasoning in LLMs. 2025 M.Sc. Thesis.
- Hao Zhang: Grouding Generative LM with Knwoledge Graphs for Commonsense Reasoning. 2023. M.Sc. Thesis.
- Sewin Tariverdian: Fusing Structured with Unstructured Modalities for Multi-Hop QA. 2022. M.Sc. Thesis.
- Soulaima Khamari: NLI as a Multi-Agent System for QA. 2022. B.Sc. Thesis.

PROJECTS

Athene: Security in LLMs

Jan 2024 - Today

- Grant proposal writing and project manager.
- Research on LLMs vulnerabilities (e.g.: prompt injection, backdoors, and data leaks).
- Funded from 2024 to 2027 by the National Research Center for Applied Cybersecurity ATHENE.
- Project web: [link](#).

Microsoft: Accelerating Foundation Models Research

Jan 2024 - June 2024

- Grant manager.
- Azure credits to support our research on large language models.
- Setup of Azure OpenAI service and contact person for the whole lab.
- Publications: (Puerto et al., 2024).

DFG: Software for Question Answering Research

Jan 2022 - June 2024

- Project manager since June 2022.
- Development of a platform for question-answering research.
- Project web: square.ukp-lab.de.
- Total amount of users: 1K.
- Funded by the German Research Foundation (DFG).
- Publications: (Puerto et al., 2024), (Fang et al., 2023), (Puerto et al., 2023a), (Puerto et al., 2023b), (Sachdeva et al., 2023), (Baumgärtner et al., 2022).

IITP: Exobrain

Dec 2018 - Feb 2021

- 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.
- Project web: exobrain.kr.
- Publications: (Kang et al., 2020), (Lim et al., 2020), (Kang et al., 2019), (Puerto et al., 2019).

Rich Context Competition

Oct 2018 - Feb 2029

- 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. We solved the challenge by modeling it as a question-answering task and presented the project in the competition's workshop at New York University.
- NYU's News Post: [Link](#).

LECTURES AND TALKS

Lectures

- QA Lecture in the course NLP4Web @ TU Darmstadt (WS 22/23, WS 23/24) - Instructor: Prof. Iryna Gurevych.
- QA Lecture in the course NLP @ Koç University (SS 23) - Instructor: Prof. Gözde Gül Şahin.
- Data Analysis Software Project for Natural Language (WS 22/23) @ TU Darmstadt - Co-Instructor with Dr. Yufang Hou.
- QA Lecture in the course Deep Learning for NLP (SS 22/23) - Prof. Ivan Habernal.

Invited Talks

- Seminar of the Cyber-Physical Health and Assistive Robotics Technologies group (CHART) @ The University of Nottingham, May 2025
- UKP-SQuARE @ Huawei Search Engine Academic Workshop, November 2022.
- [ScaDS.AI Machine Learning Summer School: QA Research, July 2022](#).
- [Hacia Afuera Podcast: Making AI Answer Questions \(in Spanish\), February 2022](#).