

YEN-HSIANG WANG

+886 928-666-218 [✉ heliart_sho@hotmail.com](mailto:heliart_sho@hotmail.com) [🌐 https://heliart.me/](https://heliart.me/) [🐙 github.com/kenny75557](https://github.com/kenny75557)

Summary

Currently conducting research at the [NCHU NLP Lab](#) on advancing retrieval methods and LLM applications. Previously, gained industry experience developing AI applications, bridging real-world challenges with academic insights. Skilled in both research and practical implementation of NLP solutions.

Publications

Conference/Journal Papers

(† indicates equal contribution.)

- 1 **Yen-Hsiang Wang**, Feng-Dian Su, Tzu-Yu Yeh, Yao-Chung Fan. [A Cross-Lingual Statutory Article Retrieval Dataset for Taiwan Legal Studies](#) In *ROCLING 2024*. (Best Paper Nominee)
- 2 You-Cheng Liao, Chen-Jui Yu, Chi-Yi Lin, He-Feng Yun, **Yen-Hsiang Wang**, Hsiao-Min Li, Yao-Chung Fan. [Learning-From-Mistakes Prompting for Indigenous Language Translation](#) In *ACL LoResMT 2024*.
- 3 Xingdi Yuan, Tong Wang, **Yen-Hsiang Wang**[†], Emery Fine, Rania Abdelghani, Pauline Lucas, H el ene Sauz eon, Pierre-Yves Oudeyer. [Selecting better samples from pre-trained LLMs: a case study on question generation](#) In *Findings of ACL 2023*.
- 4 Rania Abdelghani, **Yen-Hsiang Wang**, Xingdi Yuan, Tong Wang, Pauline Lucas, H el ene Sauz eon, Pierre-Yves Oudeyer. [GPT-3-driven pedagogical agents to train children’s curious question-asking skills](#) In *International Journal of Artificial Intelligence in Education*.

Tutorial

- ROCLING 2024 "Mastering Retrieval-Augmented Generation: Unlocking Advanced Retrieval Techniques," [[Website and Slide](#)]

Work Experience

National Chung Hsing University Natural Language Processing Lab

Jan. 2024 – Present

Research Assistant

Taichung, Taiwan

- Exploring how LLMs can enhance translation for low-resource or unseen languages by leveraging Chain-of-Thought reasoning and providing relevant patterns through sentence retrieval. [[LoResMT ACL'24](#)]
- Proposed a universal method for generating a synthetic legal QA dataset and validating the performance of Sparse embedding, Dense embedding, and LLM query expansion in legal domain retrieval. [[Github Link](#)]
- Researching on Distilling knowledge from dense embedding models for query expansion finetuning to enhance the performance of term-based retrieval (like BM-25) systems.
- Building a retrieval augmentation generation application for agricultural domain question-answering, focusing on retrieval and re-ranking methods to enhance accuracy. [[Shennong-TAIDE](#)].

EmotiBot

Jun. 2023 – Dec. 2023

Machine Learning Engineer

Taipei, Taiwan

- Developed AI applications for intelligent customer support, including Apple product sales and KGI Securities, leveraging Large Language Models and Retrieval-Augmented Generation (RAG) architecture.

Inria, National Institute for Research in Digital Science and Technology

Mar. 2022 – Sep. 2022

Research Engineer Intern

France

- Exploring the Educational application of large language models to automate the generation of diverse content to foster curiosity-driven questions among elementary students, reducing the need for manual intervention. [[IJAIED](#)]
- Researching on Question Generation Sampling and Evaluation of large language models. [[ACL'23](#)]

Industrial Technology Research Institute

Jul. 2021 – Aug. 2021

Software Engineer Intern

Hsinchu, Taiwan

- Training Transformer-based model for Legal Judgement Question-Answering.
- Development of Manual Annotation System for Data Labeling.

Education

National Chung Hsing University

Sep. 2020 – Sep. 2022

M.S. in Computer Science. Advisor: Yao-Chung, Fan

Taichung City, Taiwan

Fu Jen Catholic University

Sep. 2016 – Jun. 2020

B.S. Computer Science and Information Engineering.

New Taipei City, Taiwan

Selected Project

PTT Social Engine | *Scikit-learn, Elasticsearch, FastAPI, Docker*

Sep. 2024 – Dec. 2024

- **Developed a social opinion clustering application:** Applied the KNN algorithm to cluster topics within a specified time frame on PTT (a Taiwanese web forum) and employed LLMs to generate concise summaries and trend-based headlines.
- **Implemented a retrieval and analysis pipeline:** Utilized BGE-M3 as a text embedding model to retrieve similar articles based on user queries and applied In-Context Learning with LLMs to produce analysis of sentiment and stance.
- [\[Github Link\]](#)/[\[Demo Site\]](#)

PTT Search Engine | *Vue.js, Flask, Elasticsearch, Docker*

Oct. 2020 – Jan. 2021

- Built a search engine including all articles and comments crawled from popular boards on PTT (a Taiwanese web forum).
- Implemented BM25 with Elastic Search for retrieving comments from the database and processed the data using Object-Oriented Programming techniques.
- [\[Github Link\]](#) / [\[Demo Site\]](#)

Awards

Best Paper Nominee | *ACLCLP*

Nov. 2024

- Construct a Cross-Lingual Retrieval Dataset in the Legal domain and provide multiple retrieval methods.

Skills

Programming Languages

- C/C++, **Python**, JavaScript

Tools

- Framework: React, Flask, FastAPI
- Database: MySQL, **ElasticSearch**, MongoDB
- ML: PyTorch, Spark, Hadoop
- NLP: Transformers, PEFT, TRL, TGI

Language Proficiency

Chinese | *Native*

English | *TOEFL 80*

- Reading 21
- Listening 21
- Speaking 19
- Writing 19