

# **Yen-Hsiang, Wang**

---

**Yen-Hsiang (Kenny), Wang**

+886-928-666-218

E-mail: [heliart\\_sho@hotmail.com](mailto:heliart_sho@hotmail.com)

Linkedin: [linkedin.com/in/heliart](https://linkedin.com/in/heliart)

GitHub: [github.com/kenny75557](https://github.com/kenny75557)

---

## **About me**

Master's graduate with previous internship experience at Inria and 2+ years of research expertise in Natural Language Processing and machine learning, specializing in text generation, summarization, and data analysis. Experienced in implementing Natural Language Generation systems, including novel model application and full-stack development.

Seeking a research developer role in a company where knowledge and expertise in the field can be effectively utilized.

---

## **Education**

**National Chung Hsing University / Master of Computer Science**

Sep. 2020 -- Sep. 2022, Taichung City, Taiwan

**Fu Jen Catholic University/ Bachelor of Computer Science and Information Engineering**

Sep. 2016 -- Jun. 2020, New Taipei City, Taiwan

---

## **Experience**

**Inria/ Research Engineer Intern**

Mar. 2022 -- Sep. 2022, Bordeaux, France (Remote)

1. Learning curiosity-driven Question Generation and evaluation with Large Language Models (GPT-3), towards educational technology applications.
2. Collaborating with Microsoft Research Montreal, published two papers. (one of these two papers accepted by ACL 2023 Findings)

**Industrial Technology Research Institute/ Summer Intern**

July 2021 -- August 2021, Hsinchu, Taiwan

1. Leverage Question-Answering language models(based on Transformer) for judgment-related query with over 90% accuracy.
2. Developed a Legal document labeling system for information extraction training.

---

## Selected Projects

### Legal Document QA System

Jul. 2021 -- Oct. 2021

Question Answering bot based on the transformer-based model for querying entity owners in the legal documents.

### Legal Document Labeling Tool

Jul. 2021 -- Oct. 2021 [Github](#)

Legal document labeling system for information extraction training.

### PTT Search Engine | Vue.js, Flask, ElasticSearch, Docker

Oct. 2020 -- Jan. 2021 [Github](#)

In charge of Full-Stack. Cooperating with other teams (Web Crawler, Server) and adopting the Scrum workflow.

Maintaining a search engine and database using ElasticSearch, handling hundreds of millions of data.

### Taipei QA Bot

Sep. 2020 [Github](#)

Ask a question and the bot will tell you which government department you should go to.

---

---

## Publications

### Selecting Better Samples from Pre-trained LLMs: A case Study on Question Generation

Sep. 2022, In ACL 2023 Findings. [arXiv](#)

### GPT-3 Driven pedagogical agents for training children's question-asking skills

Sep. 2022 [arXiv](#)

---

---

## Technical Skills

### Programming Language:

- C, C++, Java, Python, JavaScript

### Toolchains:

- Frameworks: PyTorch Lightning, Vue.js, Flask, Django
- Libraries: PyTorch, Transformers, Pandas, SpaCy, NLTK
- Database: MySQL, MongoDB, ElasticSearch
- DevOps: Docker, AWS, Google Cloud