

# Hongjin Qian

PhD. student at  
Gaoling School of Artificial Intelligence  
Renmin University of China

ian@ruc.edu.cn  
chien.io  
+86 13051011770

## Education

### Renmin University of China

*PhD. Student in Artificial Intelligence (Supervised by Prof. Zhicheng Dou and Prof. Ji-Rong Wen)*

Beijing, China  
Sept. 2020 – Present

### University of Sydney

*Master of Information Technology, Major in Data Management and Analytics*

Sydney, Australia  
Jul. 2017 – Jan. 2019

### Nankai University

*Bachelor of Science, Major in Electronic information Science and Technology*

Tianjin, China  
Sept. 2013 – Jun. 2017

## Experience

### PhD. Researcher

*Gaoling School of Artificial Intelligence, Renmin University of China*

Sept. 2020 – Present  
Beijing, China

- Personalized Chatbot
- Conversational Information seeking
- Document Ranking

### NLP Researcher

*Beijing Academy of Artificial Intelligence*

Jun. 2020 – Mar. 2021  
Beijing, China

- QA System dedicated in Governance Domain
- Dense Vector Search
- Fine-Grained Named Entity Recognition

### NLP Engineer

*Elens.AI*

Jan. 2019 – Sept. 2020  
Beijing, China

- Core NLP Toolkit for Chinese, English and Uyghur
- NLP Applications in Multiple Domains (Financial, Security, Media etc.)
- Large-Scale Pretrained Language Model and Text Generation

## Publications

\* indicates equal contribution.

### Pchatbot: A Large-Scale Dataset for Personalized Chatbot. SIGIR 2021

Hongjin Qian, Xiaohe Li, Hanxun Zhong, Yu Guo, Yueyuan Ma, Yutao Zhu, Zhanliang Liu, Zhicheng Dou and Ji-Rong Wen

### Speaker or Listener? The Role of a Dialog Agent. Findings of the Association for Computational Linguistics: EMNLP 2020

Yafei Liu\*, Hongjin Qian\*, Hengpeng Xu, Jinmao Wei

## Patents

### A Latin-alphabet-based Uyghur language processing method and system (CN111428509A)

Hongjin Qian, Zhen Huang, Zhicheng Dou, Zhanliang Liu

### A method for long text retrieval in open-domain question answering tasks (CN111881264A)

Hongjin Qian, Zhanliang Liu, Zhicheng Dou, Jiajun Liu

### Optical character recognition method incorporating pretrained language model (CN111738251A)

Hongjin Qian, Zhanliang Liu, Zhicheng Dou, Jiajun Liu

### A method for Uyghur language Named Entity Recognition (CN111814433A)

Hongjin Qian, Zhanliang Liu, Zhicheng Dou, Jiajun Liu

### A semantic parsing method based on rules and learning (CN112347793A)

Hongjin Qian, Xiaotong Li, Zhanliang Liu, Yushu Yang, Zhicheng Dou, Gang Cao, Ji-Rong Wen

### A method for Pinyin-based Chinese language representation (CN110162789A)

Zhicheng Dou, Hongjin Qian, Zhen Huang

### A natural language processing method and system based on index data (CN111488423A)

Zhanliang Liu, Hongjin Qian, Zhicheng Dou, Jiajun Liu

**A natural language semantic representation method based on function declaration (CN111967247A)**

Zhanliang Liu, **Hongjin Qian**, Zhicheng Dou

**Multi-source NL2SQL system based on semantic rules and multi-dimensional model (CN112559550A)**

Zhi Li, **Hongjin Qian**, Zhanliang Liu

**A method for automatically constructing FAQ knowledge base (CN112784022A)**

Sixu Guo, **Hongjin Qian**, Yushu Yang, Zhanliang Liu, Zhicheng Dou, Gang Cao, Ji-Rong Wen

**A method for automatically generating FAQ knowledge base based on complex data format (CN112800177A)**

Sixu Guo, **Hongjin Qian**, Yushu Yang, Zhanliang Liu, Zhicheng Dou, Gang Cao, Ji-Rong Wen

**A document-ranking method for text retrieval (CN111930928A)**

Yu Zhang, **Hongjin Qian**, Zhanliang Liu, Zhicheng Dou

**A method for extracting text from image document (CN112036406A)**

Yuanyuan Huang, **Hongjin Qian**, Zhanliang Liu, Zhicheng Dou

**A semantic retrieval method (CN112035730A)**

Yang Zhou, **Hongjin Qian**, Zhanliang Liu, Zhicheng Dou

**A method for quickly realizing NL2SQL based on vectorized semantic rules (CN112001188A)**

Chaofeng Xiao, Zhi Li, **Hongjin Qian**, Zhanliang Liu, Zhicheng Dou

**A method for testing interface performance (CN111881060A)**

Huan Zhang, Zhi Li, **Hongjin Qian**, Zhanliang Liu

**A method for automatically constructing FAQ knowledge base based on tabular data (CN112800032A)**

Sixu Guo, Yushu Yang, **Hongjin Qian**, Zhanliang Liu, Zhicheng Dou, Gang Cao, Ji-Rong Wen

---

#### Technical Skills, Language Skills

**Programming Languages:** Python, C, Go

**Deep Learning Framework:** Pytorch, Tensorflow, PaddlePaddle

**Writing:**  $\text{\LaTeX}$ , Office

**Languages:** Chinese (native), English (fluent)

Last updated: May, 2021