

# Wenqi Jiang

---

## Data Engineer / Backend Engineer

My name is Wenqi Jiang. I'm a software engineer with 3 years' experience in **big data infrastructure** and **backend services**, specializing in scalability, high availability, and low latency distributed systems design and implementation. I am proficient in **Java** and **Scala** and I have studied **React** and **Flutter** to find interesting challenges. Furthermore, I enjoy powerlifting and traveling in my spare time.

### Contact:

- Mobile: KzM0IDY5MS03NTItMTU3
- Email: [jiangwenqi1995@gmail.com](mailto:jiangwenqi1995@gmail.com)
- Online Resume: [jiangwenqi.info](http://jiangwenqi.info)
- GitHub: [github.com/jiangwenqi](https://github.com/jiangwenqi)
- Blog: [dev.to/jiangwenqi](https://dev.to/jiangwenqi)

## Experience

### OEG-UPM | Ontology Engineer

Dec. 2021 - Jul. 2022

A student scholarship for the development of linking algorithms and graph linkers.

### LOOM-LD

- Analyzed ontologies from different data sources, like DBpedia, Wikidata;
- Specialized 7 linking discovery algorithms, including string-based similarity and geometry relationships;
- Implemented a Sparql-based linking application by using **Apache Jena**, and evaluated it with the OAEI-2021 datasets;
- Written the master's [thesis](#).

**Keywords:** **RDF**, **SPARQL**, **Apache Jena**, **String-based Similarity**, **DE-9IM**, **OAEI**

### TOGIC | Big Data Engineer

May 2019 - Aug. 2021 · 2 yrs 4 mos

#### Offline Data Warehouse

- Maintained the **Hadoop** cluster containing over 200 TB of data, including user and system logs from 3 different sources, which generated 300 GB per day;
- Upgraded big data components, including **Hadoop**, **Spark**, **Kafka**, etc., and monitored them on **CDH**;
- Designed and implemented a four-layer data warehouse based on the STAR model by using **Hive**, which had over 150 tables;
- Rewrote the ETL process, which saved over 80% of storage and fixed the small files' problem on **HDFS**;

**Keywords:** **CDH**, **Hadoop**, **Kafka**, **HIVE**, **Spark**

#### Data Statistics System

- Restructured the statistics system based on the new offline data warehouse;
- Rewrote the code by using **Spark DSL**, making it more readable and maintainable;
- Optimized the system's performance, reduced the process time by 25% compared to the previous version;

- Implemented more than 30 new statistical demands and fixed bugs.

**Keywords:** Sqoop, Hive, Spark, Elasticsearch, MySQL, MongoDB, Scala, Python

### Log Collection Service

- Designed and implemented the log collection pipeline using Nginx and Netty, improving performance, stability, and scalability;
- Processed all logs into the Kafka cluster, which has 5 brokers, 3 replications, and 7 days of redundancy, improving robustness.

**Keywords:** Nginx, Netty, Java, Kafka, HDFS

### WEBSTUDIO | Backend Engineer

Jul. 2018 - Apr. 2019 · 10 mos

- Implemented backend APIs of big data visualization dashboards and general reports for the BI applications;
- Developed more than 25 data operators, such as summation, date conversion, string processing etc.;
- Implemented the one-click function of syncing all user information;
- Expanded 16 data sources, including MySQL, SQL Server, MongoDB etc.;
- Maintained the big data cluster;
- Fixed bugs and developed new features for the previous projects.

**Keywords:** Dubbo, Spring Boot, MyBatis, Hibernate, Nginx, Docker, FastDFS, CDH, HBase, Impala

### KEG-THU | Software Engineering Intern

Feb. 2018 - Jul. 2018 · 6 mos

- Organized and annotated critical illness insurance documents;
- Trained models to extract the logical structure of insurance documents by using GROBID;
- Created a desktop application using above models;
- Maintained datasets in NoSQL database, like MongoDB, Neo4j;
- Completed the thesis.

**Keywords:** NLP, CRF, GROBID, Neo4j, MongoDB, Java FX

---

## Open Source

### Dapp Learning DAO

Apr. 2022 - Present

- Contributed various tutorial tasks;
- Translated the basic tasks;
- Shared the latest blockchain news and cutting-edge technics in the community.

**Keywords:** Blockchain, Web3.js, Solidity, Dapp, Ethereum, Hardhat, Alchemy, Infura, DAO

---

## Education

- **MS Data Science** · Polytechnic University of Madrid | 2021 - 2022
- **BE Computer Science & Technology** · Beijing Information Science & Technology University | 2014 - 2018