

# ZIJIAN WANG

## Master NLP & Data Science & NLP Engineering

### EDUCATION

#### Paris Sorbonne University, Master Language and Computer Science

2018-2020

Programming, NLP, Machine Learning, Deep Learning

#### Paris Sorbonne University, Master Applied French and Literature

2016-2018

### EXPERIENCE

#### Laboratoire STIH, Sorbonne University / Named Entity Recognition

05/2020 - present

- Use Different Tools NLP to detect Named Entities in German and French.
- Retrain the Models, evaluate the data, optimise parameters.
- <https://projet.liris.cnrs.fr/aphns-magis/AtelierAssisesMAGIS2020.html>

#### Shanghai Data Exchange Center / Data Scientist and Developer

07/2019 - 08/2019

- Database Management with Spring and CURL Database MySQL
- Software Test
- Github: <https://github.com/WANGZijian1994/fuxing-tools>

### IMPORTANT PROJECTS

#### Phonetic and Semantic Comparaison (Thesis for Master 1)

- Creation of algorithms to find minimum pairs (phonology).
- Discovery of French wordnet to find hyponymy and hypernymy
- Creation of services to search minimum pairs by Spring.
- Youtube: <https://www.youtube.com/watch?v=fQY9d4JoKxE>

#### Solr Research

- Clean data of Ligue 1 French football for 5 years using pandas
- Annotation and indexation of corpus
- Creation of services to search by Solr Interface and by Spring

#### Amazon Sentiment analysis

- Data Scrapy and Data Cleaning Data, using nltk and pandas
- Vectorization and Word-Embeddings to pre-train data, using sklearn
- Sentiments Classification using Machine Learning Models
- 10 "hot topics" using unsupervised learning.

#### Speech Recognition

- Collection of phoetic data by Praat and Cleaning Data by Python
- Training an HMM model to indentify isolated word

#### Algorithmes and Data Structure

- Participate in offering answers for Leetcode exercices
- Github: <https://github.com/AlgoStudyGroup/Leetcode>

### CERTIFICATIONS

- Coursera SQL for Data Science
- Coursera C++
- GRE scores: Verbal 152/170 Quantity 170/170 Writing 4/6

Paris, France  
+33 07 69 22 16 16  
[wangzijian1994@hotmail.com](mailto:wangzijian1994@hotmail.com)  
<https://www.linkedin.com/in/wang-zijian-260442165/>  
<https://github.com/WANGZijian1994>

### Skills

- Python
- Pandas
- NumPy
- Nltk
- Spacy
- Data Visualization
- Data Cleaning
- Tensorflow
- PyTorch
- selenium
- Bash
- Git
- MySQL
- neo4j cypher
- Solr
- Java
- Maven and Spring
- StanfordNLP
- C++
- JavaScript
- HTML CSS

### Languages

- French DALF C1
- English IELTS 6.5
- Chinese Mother Tongue
- Norwegian, German, Spanish, and Japanese -Elementary