

# Qi Wang

Data Scientist, NLP Engineer

avenue de Lingenfeld  
77200 Torcy, France  
☎ +33 (0)6 47 36 96 20  
✉ wq2021@outlook.com  
📁 wq2021.github.io

*Master's graduate in Data Science and NLP. Possessing a spirit of organization and analytical thinking through my professional and academic career.*

## Professional Experiences

- Feb. - Aug. 2022 **Data Science Intern**, Astek, Boulogne-Billancourt, France.
- Build a personalized recommender system for documents using novel Deep Learning methods (transformer models).
  - Design an evaluation protocole in terms of relevance and ranking of generated results.
  - Prepare monthly presentations for team meeting, synthesize all works in several research documents,.
  - Tools: Python(PyTorch, Tensorflow, Pandas, Selenium, BeautifulSoup), GitHub, Markdown, L<sup>A</sup>T<sub>E</sub>X, etc.
- Apr. - Oct. 2021 **Text Mining Intern**, BNP Paribas, Paris, France.
- Develop a new functionality of an internal platform : automatic detection of key events from daily economic and financial news in English and French.
  - Implementation and comparaison of 3 types of approaches (statistical methods, lexical properties and neural networks).
  - Build an evaluation model for quantitative and qualitative studies.
  - Tools: Python(Pandas, SpaCy, lxml), Excel, L<sup>A</sup>T<sub>E</sub>X, etc.

## Education and Degrees

- 2021-2022 **MSc. in Machine Learning for Data Science**, Université Paris Cité (ex. Université Paris-Descartes), With high honours, Major courses : Supervised learning, Unsupervised learning, Deep learning, Text mining, NLP, Big data, Recommender systems, Time series, Data Visualization, Graph reconstruction, etc.
- 2019-2021 **MA in Natural Language Processing (Multilingual Engineering)**, Inalco, With high honours, Major courses : Algorithmic, Text mining, Machine learning for NLP, Data Visualization, Object-Oriented Programming, Database management, Knowledge engineering, Advanced linguistics, etc.

## Project Works

- Building a scoring model using supervised algorithms applied to banking and relational data (Python Pandas, Scikit-Learn, Numpy, SciPy, Matplotlib, etc).
- Scientific papers metadata analysis using unsupervised learning methods (Pandas, Scikit-Learn, Numpy, SciPy, Seaborn, NLTK, etc).

## Certifications

- IBM Data Science Professional Certificate
- DeepLearning.AI & Stanford Online : Machine Learning Specialization
- DeepLearning.AI & AWS : Generative AI with Large Language Models
- Python3 : from the basics to advanced language concepts

## Achievements

- Jun. 2023 Won the first place in 5th HackaTAL(Hackathon in NLP), by providing a effective solution for French citation detection task using machine learning methods and linguistic patterns.

## Computer Skills and Tools

- Advanced Python(Pandas, Tensorflow, PyTorch, Scikit-learn, Transformers, SpaCy, NLTK, Matplotlib, Seaborn, etc), MySQL, HTML5, CSS3, VS Code, XML, MicroSoft Office, iWork
- Intermediate R(Shiny), Shell, Java, PHP, JavaScript, Git, L<sup>A</sup>T<sub>E</sub>X, MySQL, Dataiku, Power BI, Docker

## Languages

- Native **Mandarin**
- Professional **French (DALF C2), English (B2)**