

# Jinfeng HU

📍 Paris, France    ✉ jinfenghu99@outlook.com    ☎ +33 07 61 77 39 66    in linkedin.com/in/jinfeng-hu  
🔗 github.com/JinfengHU99

## Profil

---

Actuellement freelance NLP/AI Engineer, spécialisée dans le développement de systèmes intelligents multilingues (anglais, français, chinois), l'analyse de données structurées et non-structurées, et la modélisation prédictive. Expérimentée dans la création de chatbots, le fine-tuning de modèles LLM et les systèmes de questions-réponses RAG. Capable de transformer des données complexes en insights exploitables et solutions automatisées pour soutenir la prise de décision stratégique.

## Expériences

---

### AI Engineer

Paris

*Syncro (Mission freelance)*

07/2025 – 10/2025

- Développement d'un système de veille automatisée sur l'IA, permettant la collecte en temps réel des dernières actualités et publications dans le domaine de l'intelligence artificielle.
- Analyse et synthèse des tendances émergentes de l'IA, fournissant des insights stratégiques pour anticiper les évolutions technologiques et guider la prise de décision.

### Data Analyst

Paris

*Kawaiipanda (Mission freelance)*

07/2025 – 10/2025

- Construction d'une base de données structurée en JSON à partir des données clients couvrant la période 2016–2024, facilitant l'exploitation et l'analyse des indicateurs clés sur plusieurs années.
- Analyse des performances commerciales des stands sur plusieurs salons, identification des facteurs influençant le chiffre d'affaires, évaluation de la viabilité économique de chaque salon et modélisation de prévisions de revenus pour orienter les décisions stratégiques à long terme.

### NLP Engineer

Paris

*e-Peak People HR (Mission freelance)*

03/2025 – 05/2025

- Conception d'un système NER multilingue (anglais, français, chinois, hindi, italien, arabe) permettant d'identifier avec précision les expressions professionnelles et d'accélérer significativement la revue manuelle des profils.
- Développement d'un système de questions-réponses basé sur RAG et le fine-tuning de modèles Mistral, exploitant les données internes pour générer des évaluations cohérentes avec les compétences des employés et améliorer la précision des recommandations de développement RH.

### Stagiaire en Automatisation NLP

Paris

*Cyberfreelance*

02/2024 – 07/2024

- Extraction et classification de plus de 500 annonces JSON via Web Scraping et détection de similarité, améliorant la structuration et l'accès aux données.
- Création d'un chatbot de recherche de profils utilisant LangChain et GPT-4, permettant une identification rapide des candidats pertinents
- Déploiement d'une interface web en HTML et Flask, facilitant l'interaction utilisateur avec le chatbot et la consultation des données.

### Stagiaire en Recherche NLP

Blois

*LIFAT (Laboratoire d'Informatique Fondamentale et Appliquée de Tours)*

03/2023 – 07/2023

- Étude des différences de cohérence entre résumés humains et générés automatiquement (corpus SummEval), fournissant des insights sur la qualité et la structure des résumés.
- Annotation de chaînes de corréférence et construction de métriques de similarité sémantique (F1-score), permettant l'évaluation quantitative de la cohérence textuelle.

## Formation

---

**Université d'Orléans**

*Master de Sciences du Langage - Linguistique Outillée et Traitement Automatique des Langues*

09/2022 - 09/2024

**Université Normale de Chine du Sud**

*Licence de Lettres – Langue et Littérature Française*

09/2018 - 06/2022

**Université de Tours**

*Programme d'échange – Langues Étrangères Appliquées (Français-Anglais-Chinois)*

01/2022 - 05/2022

## Compétences

---

**Programmation :** Python, JavaScript, R, SQL, Bash

**NLP :** Transformers, BERT, SpaCy, NLTK, NER, Classification, Scikit-learn, TensorFlow, PyTorch

**LLM :** Prompting, RAG, Fine-tuning, Hugging Face, LangChain

**Data & Analyse :** Pandas, NumPy, Matplotlib, Seaborn, Analyse de données structurées et non-structurées, Nettoyage et transformation de données, Extraction d'informations

**Machine Learning :** Modélisation prédictive, Entraînement de modèles, Évaluation, Optimisation

**Web & Déploiement :** Flask, FastAPI, HTML, CSS, Docker, Git, GitHub, AWS, Jupyter Notebook

**Web Scraping :** BeautifulSoup, Selenium, Scrapy

**Langages de balisage :** XML, XPath, TEI

**Analyse de données linguistiques :** Iramuteq, Lexico 3, TXM, Praat, ELAN

**Linguistiques :** Français (niveau C1), Anglais (niveau C1), Chinois (langue maternelle)

## Centre d'intérêt

---

- Peinture à l'aquarelle et en numérique
- Ski et snowboard
- Photographie argentique et appareils numériques rétro