

# Mathieu LAÏ-KING

📍 Paris, France ✉ laiking.mathieu@gmail.com ☎ +33 6 89 26 86 14 📺 in/mathieulaiking 🌐 mathieulaiking.github.io/

---

## SUMMARY

PhD from Université Paris Saclay in Natural Language Processing (defense in April 2025) looking for a role as a research scientist, research engineer or data scientist.

---

## EXPERIENCE

### Data Science/NLP R&D Intern

SEGULA Technologies

February 2021 – July 2021, Lyon, Auvergne-Rhône-Alpes, France

- MS last year internship in Research and Innovation team. Using machine learning and NLP to find recurrent problems in customer service free text messages data about vehicle malfunctions.
- Implementation of NLP pipeline (preprocessing, tokenization & embedding, clustering, topic modelling) to extract similar groups of problems and identify their main topic.
- Added information about main problems in the database (unused before this project).

### Machine learning R&D Intern

Fraunhofer Institute for Production Systems and Design Technology IPK

July 2019 – June 2020, Berlin, Germany

- Gap year internship in an applied sciences research institute. Working in different teams for 2 projects.
- First project : using AI for product requirements. Computer vision + Machine learning to extract different parts of a scanned JPG requirement file then OCR to extract text and classification to distinguish requirements in the documents. NLP pipeline to compare requirements (text sequences). Work done for an industry partner of the institute, as an R&D project.
- Second project : Web development using Java, JSF and MySQL + deployment on Glassfish server. Application used internally in the institute to manage projects.

---

## EDUCATION

### Doctor of Philosophy – PhD, Computer Science

Université Paris-Saclay • Paris, France • January 2025

- Research on automatic assessment of quality in biomedical scientific articles using Natural Language Processing, and data selection for pretraining biomedical models
- Participation in conferences (TALN, ACL 2024), and 3 publications (n2c2 Shared Task, BioNLP@ACL2024, Revue TAL 65.2)

### Master of Science – MS, Computer Science

CPE Lyon • Lyon, France • July 2021 • 3.732

- Major Courses in Electronics, Mathematics, Computer Science, Machine Learning
- Member of Student Union ("Bureau des Eleves"), organization of events (ski trip, integration week end at a resort, parties) for ~300-400 students for one year

---

## CERTIFICATIONS

### Natural Language Processing Specialization

Coursera • 2021

### CAE (C1 – Grade B)

Cambridge University Press & Assessment English • 2020

---

## SKILLS

General Skills : Statistics, Machine Learning, Deep Learning, Natural Language Processing, Model Evaluation, Large Language Models, Research Project Planning,

Computer Science : Python, Java, Linux, Git

NLP / Deep Learning / Machine Learning : Pytorch, Transformers, vLLM, W&B, accelerate, deepspeed, spacy, stanza, nltk, Sentence Transformers, datasets, scikit-learn