From Japan

EMMANUEL THIÉRY

PYTHON DEVELOPER / DATA SCIENTIST

COORDONNÉES @

✉ emmanuel.thiery@protonmail.com

🌐 Tokyo, Japan

COMPÉTENCES ⚙

Programming Languages: Python

Data Science & ML: Machine Learning, Deep Learning, NLP, Computer Vision

AI Technologies: LLM, AI Agents, ChatGPT, Stable Diffusion

Frameworks & Libraries: Scikit-learn, TensorFlow, PyTorch, Keras, FastAI, HuggingFace

API Development: Flask, FastAPI, gRPC

Data Processing: Pandas, NumPy, Statsmodels

Visualization: Matplotlib, Seaborn, Plotly, Streamlit

Databases: MongoDB, PostgreSQL, Supabase

Cloud Platforms: AWS (Lambda, S3, EC2, API Gateway, SQS, Batch, QuickSight, Cognito), OVH VPS

DevOps & Tools: Docker, GitLab CI/CD, Serverless Framework, n8n, Make.com

Scraping & Automation: BeautifulSoup, Selenium, Scrapy, Splash

Vector Databases: FAISS, Pinecone, Qdrant

AI Frameworks: Langchain, Langraph, LlamaIndex

Orchestrator: Airflow, Prefect

Reinforcement learning : stablebaselines3, gymnasium

EXPÉRIENCES PROFESSIONNELLES 💼

MARS 2025
AVR. 2025

TRADING BOT REINFORCEMENT
Particulier

POC of a trading bot using Reinforcement learning :
- creation of trading gymnasium environnement
- dataset generation and cleaning, feature engineering
- training using PPO and RecurrentPPO with stablebaselines3
- tracking experiment with mlflow and tensorboard

DÉC. 2024
MARS 2025

AGENT CHATBOT
OMD

- chatbot to retrieve information from a huge database, generating graph

NOV. 2024
NOV. 2024

DEBUGGING RESEARCH CODE
Kuan-Yu Chiu

- debugging code on Game Theory field. several reinforcement learning tested

AOÛT 2024
OCT. 2024

MOBILE APP DEVELOPER
Mnemo Kanji

- Created mobile application "Mnemo Kanji" for Japanese language learners using mnemonic techniques.
- Developed with FlutterFlow and Supabase backend.
- Successfully deployed on Google Play Store

AVR. 2024
JUIL. 2024

RAG OPTIMIZATION SPECIALIST
Legaware

- Evaluated impact of different RAG configurations using various embedding models (OpenAI, Mistral).

• Tested multiple chunking strategies and rerankers (Jina, Cohere).

• Utilized LangSmith framework for testing and Seaborn for visualization.

• Delivered comprehensive optimization report to client

AVR. 2024
MAI 2024

E-COMMERCE AI SOLUTIONS DEVELOPER
OCITO LA GALERIE

• Developed product description generation system using GPT-4-Turbo.

• Created automated product categorization script with LLM integration.

• Implemented web scraping solutions for e-commerce product classification

FÉVR. 2024
AVR. 2024

ART MARKET ANALYSIS DEVELOPER
OMD

• Built metrics system to rank contemporary artists by popularity and revenue.

• Developed API for data exposure and web application for results visualization.

• Implemented Docker containerization and Supabase integration

FÉVR. 2023
PRÉSENT

AI CHATBOT DEVELOPER
Re-mind

• Developed chatbot using OpenAI API and Langchain framework.

• Created Streamlit application for prompt management.

• Implemented MongoDB storage for conversation analysis

OCT. 2022
PRÉSENT

DATA PIPELINE & VISUALIZATION ENGINEER
Equidata Sport

• Built automated scraping system with AWS Lambda and S3 storage.

• Implemented CRON-triggered data processing pipeline.

• Developed Streamlit visualization dashboard deployed on AWS EC2.

• Established GitLab CI/CD with Serverless Framework for AWS deployment

SEPT. 2020
DÉC. 2022

NLP, HANDWRITTEN OCR AND CLOUD SERVICES SPECIALIST

Drag'n Survey

- Migrated Flask API to FastAPI with Docker containerization.

- Fine-tuned BERT models for sentiment analysis across 16 languages.

- Tracked model performance using Neptune.ai.

- Deployed services using AWS Lambda and API Gateway.

- Optimized OCR performance with Tesseract benchmarking.

- Handwritten dataset creation for OCR fine tuning with easter2

JANV. 2022
JUIL. 2022

INFRASTRUCTURE & DEVELOPMENT ENGINEER

Omnegy

- Created Streamlit application for OCR performance visualization.

- Converted REST APIs to gRPC architecture.

- Implemented Selenium-based web scraping on AWS Lambda.

- Established GitLab CI/CD pipeline for AWS services

JUIN 2021
SEPT. 2021

DATA PROCESSING OPTIMIZATION SPECIALIST

Disney

- Refactored Python scripts for Excel file processing on AWS S3.

- Optimized database insertion performance with multithreading.

- Implemented AWS Batch services for script orchestration

OCT. 2021
DÉC. 2021

INTERFACE DEVELOPER

BCA Expertise

- Developed Streamlit interface for OCR pipeline visualization.

- Refactored monolithic code into gRPC microservices

NOV. 2020
MAI 2021

BACKEND AUTOMATION ENGINEER

Mobsuccess

- Automated Gifi.fr backend processes using Google Sheets API.

- Developed FastAPI interface for JSON and CSV management.
- Created AWS QuickSight dashboards.
- Integrated Yext API for store hours management

MAI 2023
JUIN 2023

DOCUMENT INTELLIGENCE SPECIALIST
Kiklos

- Built vector database from diverse document formats (PDF, Excel, etc.).
- Implemented ChatGPT integration for document querying.
- Created MongoDB storage for user interaction history.
- Deployed services with Docker on VPS

SEPT. 2023
SEPT. 2023

AI MODEL DEPLOYMENT
MANUSCRY

- Deployed custom Stable Diffusion models on Replicate platform.
- Created three specialized endpoints for image generation

FÉVR. 2020
MAI 2020

HEALTH DATA ANALYST
Urgotech

- Developed EEG data processing pipeline with multiprocessing.
- Created interactive Streamlit applications for data visualization and parameter adjustment

MARS 2020
AVR. 2020

CLASSIFICATION STUDY
WhiteLab Genomics

- Implemented supervised and unsupervised models for scientific article classification

OCT. 2019
DÉC. 2019

LARGE-SCALE SCIENTIFIC ARTICLES SCRAPING
WhiteLab Genomics

- Scraped and integrated millions of scientific articles into DocumentDB

ÉDUCATION

2013   **MASTER'S DEGREE IN NEUROSCIENCES**
Université Pierre and Marie Curie (UPMC)

2015   **BACHELOR'S DEGREE IN PHYSICS AND CHEMISTRY**
Université Pierre and Marie Curie (UPMC)

2011   **BACHELOR'S DEGREE IN LIFE SCIENCES**
Université Pierre and Marie Curie (UPMC)

2022   **UNIVERSITÉ (PAS DE DIPLÔME)**
University of Alberta, Alberta Machine Intelligence Institute

Reinforcement Learning Specialization
Certification ID: R8ZSFHKV5BN5
https://www.coursera.org/account/accomplishments/specialization/R8ZSFHKV5BN5

LIEN

Portfolio - Mnemo Kanji
https://play.google.com/store/apps/details?id=com.mycompany.dicojap

CENTRES D'INTÉRÊT

- Sports: Rock Climbing
- Việt Võ Đạo
- Kung Fu
- Piano

- Jogging
- Capoeira
- Arts: Theater

ACTIVITÉS EXTRASCOLAIRES

MEETING EVENT MANAGER
Climate-KIC Alumni, Paris

TREASURER AND COMMUNICATION MANAGER
PMClab (UPMC's FabLab)

VICE-PRESIDENT OF STUDENT UNION

Jussieu

STAGES

NEUROSCIENCES LABORATORY INTERNSHIP

CNRS, Paris | Paris

INTERNATIONAL RESEARCH INTERNSHIP

RIKEN, Wakou-shi, Japan | Wakou-shi