Fabien Houang

Philadelphia, PA | (+1) 215 669 5388
houang.fabien@gmail.com | linkedin.com/in/fabien-houang

EDUCATION

EPITA
MSc in Computer Science
Paris, France
SCIA - Major in Data Science and Artificial Intelligence

LINKS

LinkedIn:// **houang-fabien**
Github:// **fabienhouang**
Gitlab:// **fabien_houang**
Portfolio:// **fabienhouang.github.io**

CERTIFICATIONS

Google Cloud Certified x8 :
- Cloud Data Engineer
- Cloud Architect
- Machine Learning Engineer
- Cloud DevOps Engineer
- Cloud Database Engineer
- Cloud Security Engineer
- Cloud Developer
- Cloud Network Engineer

SKILLS

PROGRAMMING

Proficient:
Python • C/C++ • SQL
Linux shell • LaTeX
Comfortable:
Go • Javascript • Scala • Ruby
NoSQL | MongoDB • Rails • Django
Familiar:
Caml • Matlab • Lua

LIBRARIES & TOOLS

Proficient:
GCP • Terraform • BigQuery
Keras • Numpy • OpenCV
Pytorch • Google Colab • VTK
Jupyter Notebook • Matplotlib
Sacred | Omniboard • GIT
Comfortable:
Pandas • Seaborn • Scikit-Learn
OpenMP • Optuna • Neptune
Familiar:
AWS • Spark • Hadoop

LANGUAGES

English	Fluent
French	Native
Chinese(Mandarin)	Proficient

EXPERIENCE

ROCHER GROUP | Senior Data Engineer
Jan 2024 - Oct 2024 | Paris, France
Skills : Python, SQL, Google Cloud Platform, Airflow Composer, Bigquery, DataFlow
- Contributed on building a data platform to process **Consumers' retail and media data** on **Google Cloud Platform**
- Created data processing pipelines using **Cloud Composer** as orchestrator
- Created complex **SQL** scripts in **BigQuery** to process data and meet business needs
- Stored batch and streaming data from **Dataflow** in tables for reporting
- Created and maintained technical documentation on **Confluence**

L'OREAL | Data Engineer @ L'Oreal Beauty Tech Data Platform
May 2021 - Sep 2023 | Paris, France
Skills : Python, SQL, Google Cloud Platform, Terraform, Bigquery, Workflows, Flask
- Collaborated on building data platform for **Sourcing** domain-specific data, to ingest and process suppliers, spending, and contracts data on **Google Cloud Platform**
- Created data processing pipelines using **Terraform** for IaC and **Cloud Workflows/Scheduler** to orchestrate and automate processing steps
- Migrated data from different sources like SAP, Ivalua, Anaplan, Tibco to GCS/BQ
- Wrote **SQL** scripts and stored processed data in **BigQuery** for applications
- Exposed data through API implemented with **Python Flask**, running on **Cloud Run**
- Setup authentication through **Azure AD** and **Apigee** to access data from GCP

SIEMENS HEALTHINEERS | Research Engineer in Deep Learning
Mar 2019 - Mar 2020 | Princeton, New Jersey, United States
Skills : C/C++, Python, PyTorch, Optuna, OpenMP, OpenCV, Sacred | Omniboard, VTK
- Collaborated on Deep Learning project for real-time **3D cardiac chambers modelization** using 3D ultrasounds time series medical data
- Trained **Multi-Agent Deep Reinforcement Learning** for landmarks detection
- Implemented and trained 3D volumes classification models on unbalanced dataset
- Optimized production code speed and optimized models' hyperparameters
- Achieved **3x faster and 10% more accurate modelization** than the base solution

SIDE PROJECTS

VISION VENDOR | Mobile App Project
Jun 2024 – Sep 2024 | Philadelphia, Pennsylvania, United States
Skills : Flutter, Dart, GCP, Gemini AI API, Mobile App development
- Building a mobile app to **enhance online reselling experience** by integrating **Gemini AI** to streamline the process on popular platforms (eBay, FB Market)
- Submitted the app prototype at **Gemini API Developer Competition 2024**
- Built a cross-platform iOS and Android app using **Dart and Flutter Framework**
- Integrated Gemini AI API calls and prompt engineering on **Google AI Studio**

AUTONOMOUS RACING CAR | Deep Learning Project
Apr 2018 – Feb 2019 | Paris, France
Skills : Python, Keras, OpenCV, GCP, Deep Learning, Raspberry Pi, Embedded system
- Built an **autonomous RC car** connected to a **Raspberry Pi 3** with camera module
- Collected, processed and simulated data for **Deep Learning** model training on GCP
- Designed, trained **time and memory efficient** models to run on embedded system
- Led a team of 4 students and competed against more than **20 teams nationwide**