

BASTIN Florian

bastinflorian1@gmail.com • (+33) 6 37 51 96 84 • <https://github.com/BastinFlorian>



ABOUT ME

Graduated from a mathematics and informatics bachelor, I did a research master in AI at the most prestigious French school, ENS. After an internship as a NLP Data Scientist at Lili.ai, I joined Octo Technology, part of Accenture.

I worked for several clients such as Le Monde, Casino, Club Med, Chanel, Pernod Ricard and Suez. I took different positions for these clients: Tech Lead Data, Data Scientist or Engineer, Deep Learning or LLM engineer.

During these experiences, I developed Machine and Deep Learning expertise in different fields (pose estimation, gesture recognition, churn and propensity to subscribe modeling, Gen AI RAG chatbots). I also build data architectures and robust data pipelines (see our Apache Beam 2023 [talk](#) in New York).

I improved my skills in the GCP and Azure Cloud working on it with my clients or with certifications. I discovered CI/CD and clean code & architecture good practices. I recently got awarded twice by Google at their Gen AI competition in the category Superior GenAI Business Value and ROI.

Today, I am mainly developing LLM and Gen AI use cases for my clients. I am also a Generative AI & GCP Cloud teacher two weeks per year at Dauphine Tunis University.

WORK EXPERIENCES

Suez @Octo - Current client - Gen AI Engineer

2024, March

Paris

- In charge of developing 35 Generative AI use cases using Azure Cloud Provider, OpenAI LLMs such as GPT 3.5, GPT 4, GPT 4.0, Code Interpreter, Mixtral8x7B.
- We built a private python library integrating Gen AI tools such as LangFuse, Phoenix, RAGAS and Deep Eval to speed up our LLMs Ops architecture.
- We also worked with fine-tuning embedding and models solutions such as Hugging Face, GCP and Jinja AI
- ROI estimation based on customer time saving, model quantitative and qualitative performance and use case cost.
- Keep or kill Agile decision process and development of production ready solutions on chosen use cases
- Currently building Agents to automate the development of our own use cases based on key tasks and our private library

Pernod Ricard @Octo - Gen AI Engineer

2023, November

Paris

- Development of a production ready Teams chatbot using Generative AI and RAG techniques. We used GPT 3.5 Turbo LLM and Ada 2 embeddings.
- We used Azure Cloud Service, PostgreSQL database, Azure App Service and Azure AI Studio for our architecture. As a ROI, we reduced Level 1 Human resources and IT support tickets by 45%

Chanel @Octo - Deep Learning Engineer
Paris

2023, June

- Development of a custom Face Landmarks detection Tensorflow model for automatic make up for a IOS/Android application
- We used HAGRID dataset to improve our custom labeled dataset, Google Mediapipe solutions and TF Lite and CoreML to create an edge model.

Octo - Gen AI Researcher
Paris

2023, January

- Development of Generative AI POCs for multiple clients (chatbot, large document summarization, NDA legal analyzer, ...)
- 2 Hackathons competition won with Google GCP using Mistral LLM and PALM (See this [LinkedIn Post](#) and [this one](#))
- R&D on Multimodal and Lang Chain good practises leading to [Octo Blog](#) writings

Casino @Octo - Data Engineer
Paris

2022, September

- We built a GCP architecture using DataFlow and Apache Beam to speed up events processing.
- We develop a SOTA production ready solution and were invited by Google at the Apache Beam 2023 New York conference to [present our solution](#)

Le Monde Newspaper @Octo - Tech Lead Data
Paris

2018, November

- We build 7 production ready solutions for news articles recommendations (statistical ranked, content and user based approach), user segmentation models (machine & deep learning models) and NER, NLP or graph based news detector to improve search quality.
- We used most of the GCP tools to build a scalable Data Architecture that is still used by the newspaper to improve user satisfaction.

Lili.ai - NLP Research Scientist
Paris

2018, January

- Developing unsupervised NLP techniques to detect topics on large document corpus. The goal of this use case is to detect delay in large industrial projects
- Used Spacy, NLTK and transformer model in a Python, SQL environment

CERTIFICATIONS & ACHIEVEMENTS :

Cloud Certifications

- GCP Professional Data Engineer Certification
- GCP Professional Machine Learning Engineer Certification

Hackathons

- GCP Gen AI 2024 - Ranked 1st
- GCP Gen AI 2024 - Ranked 1st
- Le Monde NLP Recommenders 2024 - Ranked 2nd
- Viva Tech - Product Recommender 2018 - Ranked 5th
- GCP Professional Machine Learning Engineer Certification

Talks and publications

- Octo Gen AI [talk](#) - 2024, June
- GCP Gen AI use case [talk](#) - 2023, December
- Google Apache Beam Conference [talk](#) - 2023, June
- Octo [writings](#)
- Generative AI and GCP Teacher as Dauphine Tunis University (2 weeks per year)

STUDIES

Ecole Normal Supérieur (ENS) - PSL, Saclay

2018

Degree: Master 2

Studies: AI Research most prestigious French degree

Paris Dauphine University - PLS, Paris

2013-2017

Degree: Licence - Master 1

Studies: Mathematics, Statistics

SKILLS

Programming Languages : Python, Deep Learning libraries (TensorFlow, Pytorch), Git, Docker, Cloud Services (GCP, Azure, AWS), SQL

Languages: French (native), English (fluent), Spanish (advanced)

Sports: Football, Tennis, Hiking, Trails