

Malek BEN YOUSSEF

✉ malekbenyoussef00@gmail.com [in](#) LinkedIn [G](#)itHub [☎](#) +33744774402

📍 Paris, France [🔗](#) mbenyoussef.tech



EDUCATION

Data Science MSc Degree,
EDC Paris Business School [🔗](#)
October 2025 – Present
Paris, France

Computer Science Engineer,
National Institute of Applied
Sciences and Technology
(INSAT) [🔗](#)
2022 – 2025 | Tunis, Tunisia

Preparatory Cycle, National
Institute of Applied Sciences and
Technology (INSAT) [🔗](#)
2020 – 2022 | Tunis, Tunisia

Relevant Courses: Databases,
Algorithms and Data Structures,
Big Data, Data modeling,
Optimization Methods

CERTIFICATES

Professional Data Engineer [🔗](#)
Google Cloud

**Professional Machine Learning
Engineer** [🔗](#)
Google Cloud

Core Designer Certificate [🔗](#)
Dataiku

Deep Learning Specialization [🔗](#)
DeepLearning.AI

LANGUAGES

English: Proficient

French: Proficient

German: Basic

SKILLS

Programming Languages
Python, SQL, C, C++

**Machine Learning & AI
Frameworks**
Google ADK, LangGraph,
LangChain, LlamaIndex, SpaCy,
NLTK, Hugging Face, RAG, Prompt
Engineering, Context Caching

Cloud & MLOps
GCP, AWS, Vertex AI, Dataflow,
Pub/Sub, Cloud Composer,
BigQuery, Apache Airflow, Docker

PROFESSIONAL EXPERIENCE

AI Engineer, Ceva Logistics (CMA CGM) [🔗](#)
October 2025 – Present | Paris, France

- Developed a dual-agent system that reduced data labeling costs by 92% (from \$500k to \$41k) for 100k+ enterprise records.
- Accelerated data processing speed by 5x through Google Cloud Run Jobs.
- Implemented Context Caching strategies for Gemini 3 Pro, reducing input token usage and API costs by 90%.
- Integrated Vertex AI Search grounding to improve LLM accuracy, reducing hallucinations and accelerating factual retrieval by 40%
- **Skills:** Google Cloud, Vertex AI, Python, Fast API, Prompt Engineering, REST APIs

Data Engineer, Crewmeister [🔗](#)

February 2025 – September 2025 | Munich, Germany

- Automated Airflow ETL/ELT pipelines to ingest and transform data into BigQuery.
- Created cost efficient data workflows using Airflow automation, reducing BigQuery costs by over 57%.
- Implemented Medallion architecture using dbt for modular models
- **Skills:** BigQuery, dbt, SQL, Apache Airflow, Looker Studio, Python, Docker, Kafka, Git

Machine Learning Intern, Premier Cloud [🔗](#)

July 2024 – September 2024 | Victoria, British Columbia, Canada · Remote

- Deployed GCP Agent Builder chatbot on Cloud Run, automating reporting.
- Built Pub/Sub Dataflow streaming pipelines, reducing latency by 65%.
- Optimized partitioned and clustered BigQuery tables, reducing query costs by 48% across 27M+ rows.
- **Skills:** GCP, BigQuery, SQL, Vertex AI, Docker, Streamlit, Git, Pub/Sub, Cloud Functions, Flask

Machine Learning Intern, Silver Brain AI AG [🔗](#)

November 2023 – January 2024 | Geneva, Switzerland · Remote

- Developed a chatbot using Llama 2 LLM (7B parameters) to provide detailed service information to clients.
- Enhanced retrieval performance by implementing RAG, improving response relevance by 30%.
- Integrated and deployed the chatbot application, managing a CI/CD pipeline.
- **Skills:** Python, LangChain, LlamaIndex, Prompt Engineering, Git

PROJECTS

Query Assistant

- Built an NL2SQL system converting natural language into SQL queries.
- Deployed the application using Google Cloud Run and managed a CI/CD pipeline with Google Cloud Build, accelerating the workload by 55%.
- Optimized queries and ensured 95% of its schema compliance with Dataplex.
- **Skills:** GCP, BigQuery, SQL, Vertex AI, Docker, Git, Prompt Engineering, streamlit

Sales Stream

- Built a real-time sales analytics pipeline with Google Pub/Sub, and Apache Beam, processing over 5,000 transactions per day with average latency under 5 seconds
- Delivered business insights using Looker Studio and automated data aggregation in BigQuery.
- Improved data quality and transformation consistency using dbt with the Medallion architecture.
- **Skills:** Google Pub/Sub, Apache Beam, dbt, BigQuery, Looker Studio, FastAPI, Python

Soccer Chatbot

- Built an information retrieval system about soccer based on 1000+ PDF file pages.
- Used Llama 2 LLM and RAG techniques achieving 95% accuracy in information retrieval.
- **Skills:** Python, LangChain, Chainlit, FAISS, OpenAI Embeddings, Llama2