

Youssef Ben Moussa

AI Engineer | LLM Developer | MLOps Enthusiast
Milan, Italy

✉ yussyjob@gmail.com

🌐 youssef-ben.com

📷 yousse-f

🌐 youssef-ben-moussa

Profile

AI Engineer passionate about building intelligent systems using modern LLM and MLOps stacks. From data pre-processing to scalable deployment, I specialize in developing context-aware, efficient, and production-ready AI solutions. Currently interning at Publicis Sapient on enterprise AI tools, while developing personal and collaborative projects in parallel.

Experience

AI Engineering Intern *Publicis Sapient, Milan* *Jun 2025 – Present*

- Supporting the development of an internal AI tool to analyze social media profiles for the business strategy team.
- Implementing vision-based AI techniques to extract profile descriptions and detect post intent across platforms.
- Contributing to core components using Python, FastAPI, and OpenCV in a collaborative team setting. **Founder &**

AI Developer *Studio Innovativo (Freelance)* *Apr 2024 – Present*

- Designed and built intelligent web platforms for SMEs using modern AI tools.
- Developed AutoGenius (workflow automation) and DocuZen (document assistant).
- Stack: Python, FastAPI, LangChain, Supabase, Vercel.

Co-founder & AI Engineer *GeniApp (Startup)* *Jun 2025 – Present*

- Building Insurly, an AI platform for the insurance sector powered by agent orchestration and advanced RAG pipelines.
- Enables insurance consultants to query company databases and documents using natural language.
- Handles charts, PDFs, images, and internal data to provide instant, intelligent responses via chat.
- Also developing Docus, a second product focused on document automation and intelligent retrieval.
- Stack: Python, FastAPI, mem0, OpenAI API, LangChain, Vision AI, RAG, ChromaDB, Docker.

Software Developer Intern (C#) *Zena Office S.r.l., Piacenza* *May 2024 – Sep 2024*

- Maintained and optimized business software applications in C#.
- Assisted in debugging, refactoring, and improving internal toolsets.
- Discontinued to pursue full-time AI specialization.

Quality Control Specialist *Scigno Group, Piacenza* *Feb 2022 – May 2024*

- Managed quality inspections and audits using BDWIN ERP.
- Enhanced traceability, documentation, and process optimization.

CNC Programmer *Motridal America Inc., Piacenza* *Jun 2018 – Jan 2021*

- Developed and optimized CNC machining cycles.
- First hands-on experience with automation, logic systems, and process efficiency.

Selected Projects

MindsetOS – Multi-Agent AI Coaching Platform

Personalized AI assistant for fitness, nutrition and mindset. Agents interact through orchestrator with memory and intent-based tasking. *Stack: Python, FastAPI, mem0, OpenAI API, Docker*

Vision Strategy Agent – Social Media Visual Analysis

Tool that analyzes Instagram content (image, video, text) to extract strategic insights. *Stack: OpenCV, BLIP, CLIP,*

LangChain, Selenium

ZenPocket AI – Productivity App with Built-in Assistant
Web app for intelligent note-taking, summarization and context-aware Q&A. *Stack: Next.js, Tailwind CSS, Node.js, Framer Motion*

Education

Education

AI Developer – Advanced Specialization (IFTS - EQF Level 5) *Galdus Formazione, Milan* Nov 2024 – Sep 2025 (ongoing)

Specialized program in AI/ML, NLP, Computer Vision, MLOps and cloud deployment.

Internship: Publicis Sapient

Tools: Python, TensorFlow, PyTorch, scikit-learn, LangChain, Hugging Face, FastAPI, Docker, MLflow, GitHub Actions, Google Vertex AI

Course in Sustainability and Environmental Impact *Galdus Formazione* 2025
Focused on sustainable technologies, environmental impact, and social responsibility in digital innovation.

Computer Science Technical Diploma *"A. Volta", Lodi* 2024
5-year curriculum in programming, databases, software engineering, and information systems.

Technical Skills

Languages: Python, SQL, Bash, C#, Java
AI/ML: scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM, Transformers, LangChain, RAG (FAISS, ChromaDB), OpenAI API, CrewAI
Vision AI: OpenCV, YOLO, CLIP, BLIP
Data Engineering: Pandas, NumPy, Feature Engineering, Preprocessing
MLOps: Docker, MLflow, DVC, GitHub Actions, CI/CD, Vertex AI, KServe
Web Dev: FastAPI, Node.js, REST APIs, Streamlit, HTML/CSS, Tailwind, Next.js
Cloud/DB: Supabase, Firebase, PostgreSQL, Redis, MongoDB, Pinecone, Railway, GCP, AWS
Tools: VS Code, Jupyter, Google Colab, PyCharm

Languages

Italian: Native
Arabic: Native
English: B2
French: B2