



Yahya Momtaz

📍 Italy 🔗 <https://yayamomt.tech> ✉ yahyamomtaz@gmail.com

🌐 [linkedin.com/in/yahya-momtaz-601b34108](https://www.linkedin.com/in/yahya-momtaz-601b34108) 🐙 github.com/yahyamomtaz

Professional Experience

Researcher (Data, AI & Machine Learning), 05/2024 – present | Naples, Italy
University of Naples Federico II

- Designing and maintaining end-to-end ETL pipelines using Python, Airflow, and SQL to ingest, transform, and deliver data across systems.
- Fine-tuned a post-OCR language model with PyTorch to improve text extraction from early printed books.
- Applied deep learning and image processing techniques to restore historical manuscripts, reducing bleed-through and improving readability.
- Developed a full-stack web application for bleed-through removal—built the backend with Flask, the frontend with React, containerised with Docker, and deployed on Google Cloud Platform (GCP).
Demo: <https://bleed-through-cleaner-app-469035949643.europe-west1.run.app/> 🔗
- Launched the project website (developed the frontend with React): <https://www.magic.unina.it/> 🔗

Data Scientist (internship), Spici srl 01/2022 – 08/2022 | Naples, Italy

- Developed a data-driven web scraper to collect and structure data from LinkedIn, automating dataset generation for marketing analysis, ensuring proper data management and quality, and integrating with GCP data warehouse solutions to create efficient data pipelines.
- Implemented a Flask API for the Neo4j *database*, automating common tasks and reducing graph database implementation time by 40% for the team.

Network Engineer, Saba INV 06/2018 – 11/2020 | Tehran, Iran

- Maintained and monitored the company's internal network infrastructure, ensuring reliable connectivity and system uptime.
- Configured and managed servers, routers, and network devices to support day-to-day business operations.
- Installed and implemented hardware solutions for employees, improving workstation performance and network access.

Education

Master in Data Science, University of Naples Federico II 10/2021 – 03/2024 | Naples, Italy

Apple Developer Academy 09/2022 – 06/2023 | Naples, Italy

Bachelor in Software Engineering, Karaj Azad University 09/2011 – 09/2016 | Karaj, Iran


Languages

English (C1) • Italian (B1)

Projects


AI Image API

- Developed and deployed a real-time API for object detection and image enhancement using *YOLO and PyTorch*, including *API integrations for seamless data exchange*.
- Built the backend with *Flask*, containerized with *Docker*, and deployed on *GCP*.
- Created a responsive *JavaScript frontend* for live interaction.

Demo: <https://aiimagelive.web.app> 

Multi Agent TradeGraph

- Contributed to a multi-agent AI financial advisor using *LangGraph*, *OpenAI GPT-4*, and *Firecrawl* for real-time stock analysis and portfolio optimization.
- Resolved critical configuration and validation issues, enabling end-to-end analysis pipeline for multi-stock portfolio recommendations

Repository: <https://github.com/ApexAILabs/TradeGraph> 

Retrieval Augmented Generation (RAG) Chatbot:

- Developed a *document-based QA chatbot* for Italian legal texts, combining *LLMs + IR techniques* for accurate, context-aware answers.
- Implemented backend in *Python*, web interface with *Streamlit*, and integrated text-mining algorithms.

Repository: <https://github.com/yahyamomtaz/ChatBot> 

Prompt Analyzer

- Developed a *Streamlit app* for interactive dataset exploration with *LLM-powered insights*.
- Focused on industrialization of the Prompt Analyzer by creating *CI/CD pipelines* using *GitHub Actions* and implementing *MLOps best practices*.
- Integrated *OpenAI API* for prompt-driven analysis.
- *AI Skills: Prompt Engineering, Context Engineering*

Repository: <https://pandasai-dataset-analysis.streamlit.app/> 

Certificates

<https://www.linkedin.com/in/yahya-momtaz-601b34108/details/certifications> 

Skills

Programming

Python

Data Engineering & Databases

SQL, Neo4j, Data Pipelines, Data Modeling,
Apache Airflow

Cloud & DevOps

Docker, Git, GitHub Actions (CI/CD), Google Cloud
Platform (GCP), AWS

Machine Learning & AI

PyTorch, TensorFlow, Scikit-Learn, Hugging Face
Transformers, NLP (LLMs, RAG, Text Mining),
Agentic AI, LangChain

Data Analytics & Visualization

Pandas, NumPy, Matplotlib, Plotly, Tableau

Frontend

React, JavaScript

Publications

Modular Pipeline for Text Recognition in Early Printed Books Using Kraken and ByT5,
Yahya Momtaz et al. *Electronics* 2025, 14(15), 3083 [↗](#)

Minimizing Bleed-Through Effect in Medieval Manuscripts with Machine Learning and Robust Statistics,
Adriano Ettari, Massimo Brescia, Stefania Conte, Yahya Momtaz and Guido Russo. [↗](#)
J. Imaging 2025, 11(5), 136;

Preserving Clarity: The MAGIC project approach to ancient manuscripts,
Yahya Momtaz et al. Presented at AIUCD Conference, Verona, Italy [↗](#)

Artificial Intelligence in the Digital Humanities,
Journal of Artificial Intelligence Research and Innovation [↗](#)
Guido Russo, Yahya Momtaz, Sohail Maqsood

13th International Conference on Radiation, Natural Sciences, Medicine, Engineering, Technology and Ecology (RAD 2025 Conference) [↗](#)
Guido Russo, Stefania Conte, Lorenza Laccetti, Yahya Momtaz