



Valeria Postiglione

📍 Bologna ✉ valeria.postiglione.98@gmail.com ☎ +39-3336551894

in <https://www.linkedin.com/in/valeria-postiglione-a55413202> 🌐 ValeriaPosty

Profile

Passionate Data Scientist, with a specialization in Natural Language Processing (NLP) and experience creating GenAI applications. My academic training in Bioinformatics and my professional experiences have allowed me to acquire advanced skills in data analysis, programming, and development of machine learning models. I approach every activity, be it academic or professional, with profound enthusiasm, as I am motivated by a strong curiosity to continuously learn and grow.

Education

University of Bologna- Alma mater studiorum

2021 – 2023

Master's degree in Bioinformatics

- Grade : 104
- **Coursework:** Algorithms and Data structures, Python Programming, Biomedical Databases, Big Data Processing Infrastructures, Applied Machine Learning Basic and Advanced

University of Naples Federico II

2017 – 2020

Bachelor degree in Molecular Biology

- Grade : 110 cum laude
- **Coursework:** Chemistry, Mathematics, Physics, Biochemistry, Molecular Biology, Genetics

Experience

Data Scientist

Roma, RO

Accenture

November 2023 – current

During my experience at Accenture I held the role of Data Scientist, focusing on the development of applications based on Generative Artificial Intelligence (GenAI) and the design of advanced chatbots. I worked intensively with Natural Language Processing (NLP) models to improve the interaction between users and systems. Furthermore, in the role of Platform Data Engineer, I optimized data pipelines, ensuring the integrity and security of information. The main activities and responsibilities I carried out include:

- Developed and optimized Retrieval-Augmented Generation (RAG) pipelines, integrating search and generation models to provide contextually accurate responses and improve knowledge retrieval efficiency.
- Optimized data pipelines for efficient handling of large datasets, ensuring data integrity and compliance with industry standards.
- Performed data cleaning, transformation, and feature engineering on large-scale structured and unstructured datasets.
- Cloud Infrastructure management with Azure for scalable data processing, model deployment, and data pipeline management.
- Utilized Azure DevOps for continuous integration and deployment (CI/CD) pipeline.

HPC optimization workflow Intern

Bologna, BO

Istituto Nazionale di Fisica Nucleare- INFN

Febr. 2023 – July 2023

At the INFN in Bologna, I worked as a researcher in Bioinformatics, focusing on the automation and optimization

of bioinformatics workflows in high performance computing (HPC) and Cloud environments. I have developed and implemented solutions to improve the efficiency and scalability of bioinformatics processes. The main activities and responsibilities I carried out include:

- Designed and presented the project "Dynamic K8S for HPC Applications on Cloud" at the INFN Computing Workshop in Loano, Italy, showcasing innovative solutions to dynamically manage containerized bioinformatics applications in cloud-based HPC environments.
- Automated routine tasks (e.g., data pre-processing, job scheduling) with custom Bash and Python scripts, improving workflow efficiency.
- Utilized workflow orchestration tools (e.g., SLURM) to manage complex bioinformatics jobs on HPC clusters, ensuring efficient parallel processing and resource allocation.

Scientific editor
Microbiologia Italia

Napoli, NA
2020

During my experience as a science communicator at Microbiologia Italia, I had the opportunity to combine my passion for science with effective communication. The main activities and responsibilities I carried out include:

- Developed informative content on microbiology topics, highlighting recent advancements and scientific breakthroughs, while ensuring clarity and accuracy for readers with varying levels of expertise.
- Utilized SEO best practices to optimize articles for online visibility, ensuring that scientific content reached a wider audience through search engines and social media platforms.

Some Projects

Web App implementation based on GenAI for energy company

- The goal was to improve the document review process. Key functionalities implemented include: document summarization; chat functionality for querying aggregated data; dashboard creation for visualizing key insights; interactive chat feature to explore both documentation and domain-specific knowledge.
- Tools used: Cognitive Services, Form Recognizer, Cognitive Search, Cosmos DB, OpenAI Embedding, OpenAI Generative LLM, Langchain, Streamlit, Azure functions, FastAPI

AI Document Generation Project

- This project involved designing and developing a web application powered by GenAI to automate the creation of tender documents. Using LLM agents, the system was able to generate initial drafts of procurement contracts based on pre-defined templates and clauses. It also included mechanisms for evaluating and reviewing the generated documents against compliance criteria, in accordance with internal regulations and industry standards.

GenAI-based Chatbot in Economy and Finance field

- The chatbot aimed to help users navigate and analyze financial documents. Key functionalities developed include: document summarization; version comparison between two documents; document comparison for highlighting differences; RAGAS pipeline for evaluating and improving response accuracy.
- Tools used: Cognitive Services, Form Recognizer, Cognitive Search, Cosmos DB, OpenAI Generative LLM, Langchain, Azure, pytest

E-commerce Back-End application development

- CRM and Shopify Integration Solution for e-commerce
- The solution aimed to streamline user data management, loyalty program subscriptions, and discount code handling.
- Tools Used: Shopify API, CRM API, SQL-Alchemy, FastAPI

Technologies

Programming languages: Python, SQL, R

Main Libraries and Tools:

- **Web Development and APIs:** Flask, FastAPI
- **Generative AI, NLP, and Machine Learning:** LangChain, TensorFlow, Keras
- **Version Control and Collaboration:** Git
- **Scripting and System Management:** Bash, Linux
- **Data Processing and Analysis:** Pandas, NumPy

Cloud Computing: Azure AZ-900, Azure AI-900, AWS Cloud Practitioner

Database Architecture and Management: SQL and NoSQL (such as CosmosDB and MongoDB)

Agile framework: Asana, DevOps

CERTIFICATES AND PROFESSIONALIZING COURSES

Microsoft Certified Azure AZ-900

Microsoft Certified Azure AI-900

AWS Cloud Practitioner

Neo4j Certified Professional

Agile TDD Best Practices for APIs: Test-driven Development Strategy

IBM AI Developer Specialization

Flask-RESTful: Integrating the API Server with a MySQL Database

Professional certificate "Google Data Analytics" (Google)

Machine Learning for all (London University)

Python for Data Science, AI Development (IBM)

Building Generative AI-Powered Applications with Python (IBM)

NLP Case Studies: Developing an AI Chatbot

ChatGPT Prompt Engineering Examples and Use Cases

Python for Developers

Programming Foundations with JavaScript, HTML and CSS (Duke University)

Capstone: Retrieving, Processing, and Visualizing Data with Python (University of Michigan)

Responsible AI Delivery Bootcamp Project Management

R Programming (Johns Hopkins University)

Data Analysis with R Programming (Google)

Share Data Through the Art of Visualization (Google)

Core Skills Series: Relationship Development

Ielts C1 Level (University of Cambridge)

Soft skills

Ability to work in a team or independently;

Strong orientation towards problem solving;

Flexibility to change;

Strong time management skills;

Great attention to details;

Exceptional communication skills.