Bruno Guzzo
Cloud Engineer — AI & ML Specialist

Work Experience

AI Specialist (Mid-Senior Consultant) — January 2025 – Present
Go Reply, Turin

- Coordinated a multi-company effort to gather customer requirements for a **research-oriented healthcare platform**, enabling centralized and rapid clinical or **genomic data retrieval**. Lead project key phases form defining a complete **functional analysis** document, design a **cloud-native architecture** and implementation throught sprint planning.

- Directed a cross-functional team of developers in delivering a **B2C cloud-native mobile application** enabling home-automation, after-sales support, digital documentation access via **Gen-AI RAG chat-bot** and seamless multi-cloud **CRM integration**. Guided the adoption of a distributed serverless **micro-services** architecture coordinated using **Google Pub/Sub** to enforce **fault isolation** an scalability.

- Designed and implemented a scalable **legal treaties search engine** and **RAG solution** using **Vertex AI** and **BigQuery** with a highly efficient and concurrent data ingestion pipeline, enabling insight generation and optimization of customer workflow. Successfully improved data processing speed by **94%** after adopting a serverless cloud-native master/worker architecture relying on **Cloud Task** and **Cloud Run** for concurrent workloads distribution.

- Advised clients on **AI and agent-based system** integration, facilitating the adoption of Google Cloud enterprise ecosystem via **custom-tailored RFP solutions** to improve critical business processes. Increased client adoption rate of Vertex AI by **30%**.

- Managed **strategic client engagement** within the global development industry, coordinating **multi-project cloud advisory** and **resource allocation** that mobilized **20%** of the BU workforce and drove **30%** of unit profitability.

- Mentored junior engineers, **promoted agile best practices** through knowledge-sharing events, and **coordinated sprint reviews** and retrospectives to accelerate team performance and continuous improvement. Successfully decreased code-review time by **60%**.

Cloud Engineer (Consultant) — June 2022 – December 2024
Go Reply, Turin

- Contributed to the development of a digital platform for the **agriculture and food sector**, enhancing **data collection and analytics** capabilities to support evidence-based decision-making on sustainable farming and antimicrobial resistance. Lead the platform **migration to Google Cloud** relying on IAC best-practices as full **Terraform foundation** and **GCP Fabric** modules to ensure custom service configuration and resources tracking. Successfully reduced infrastructural fixed-costs by **30%**.

- Designed efficient multi-stage **Git-Hub** and **Git-Lab CI/CD pipeline** for quickest application deployment to production using **Cloud Build** for **Docker** image building and a fine-grained configured **Cloud Run** instance (or Firebase hosting) for computing. Bug-free releases are accomplished leveraging a comprehensive test-cases suite targeting both back-end, using **Junit** or **PyTest**, and front-end, using **Selenium** to mimic end-user actions.

- Participated in the implementation and deployment of a **digital learning solution** in the **education industry**, delivering scalable and cloud-based features aligned with client innovation goals. Data security and **GDPR compliance** are guaranteed by the adoption of rigid **IAM policy** and region locks through the definition of project **IAC**, using **Terraform**.

- Modernized a large-scale **telecommunications cybersecurity platform**, improving scalability and enhancing user experience through **architectural refactoring**, cloud integration, and automation to strengthen **service reliability** across mobile and fixed networks.

Personal & Academic Projects

MSc Thesis: Anomaly Detection with Deep Autoencoders

- Investigated a semi-supervised anomaly detection pipeline using a novel Autoencoder (AE-SAD) architecture, specifically addressing the challenge of severe class imbalance during training. Demonstrated that applying advanced under-sampling techniques can significantly boost detection accuracy (AUC) in complex many-vs-many scenarios while reducing training time by up to two orders of magnitude.

Transformer-GNN for Link Prediction

- Built a scalable data pipeline to create a large-scale graph dataset from Wikipedia using web crawling and NLP. Trained and evaluated multiple Graph Attention Network (GAT) models on the custom dataset, analyzing graph metrics and model performance.

Education

MSc, AI & ML Engineering — Oct 2022 – July 2025
UNICAL, Rende (CS), Italy

BSc, Computer Engineering — Oct 2018 – June 2022
UNICAL, Rende (CS), Italy

- 13 Jan 2000
- +393398305458
- brunoguzzo18@gmail.com
- github.com/bGuzzo
- linkedin.com/in/guzzobruno

Technical Skills

Google Cloud (*Advanced*)
- **Compute**: Cloud Run, App Engine, Cloud Functions, Compute Engine
- **AI/ML**: Gemini, Vertex AI, Agent Builder
- **Data**: BigQuery, Cloud SQL, Firestore, and Firebase Realtime Database, Cloud Storage
- **Infra**: VPC, Load Balancer, IAM, Cloud Armor, Cloud DNS, Cloud CDN, Cloud Task, Cloud Pub/Sub

AI & Machine Learning (*Academic*)
- **Frameworks**: PyTorch, PyTorch Geometric, Langchain, Hugging Face Transformers, PEFT (Parameter-Efficient Fine-Tuning), TRL (Transformer Reinforcement Learning), Unsloth
- **Libraries**: Pandas, NumPy, Scikit-learn, SciPy, NLTK, Sentence-Transformers, FAISS (Vector Database), BitsAndBytes (for Quantization)
- **Concepts**: LLM, RAG, NLP, Fine-Tuning (QLoRA), Vector Databases, Embeddings, Anomaly Detection, Time-Series Analysis, Reinforcement Learning, GNNs, CNNs, Autoencoder, Transformer, GPT, BERT, Hyperparameter Tuning, Model Evaluation

DevOps (*Advanced*)
- **IaC**: Terraform, Fabric
- **CI/CD**: Cloud Build, GitLab CI, GitHub Actions, Artifact Registry
- **Containers**: Docker
- **Others**: Linux, Git, A/B Testing

Software Development (*Advanced*)
- **Languages**: Python, Java, C, JavaScript, Typescript, Dart, SQL
- **Frameworks**: Spring Boot, Hibernate, JPA, Flask, Django, SQLAlchemy Angular, React, Flutter
- **Testing**: Selenium, Junit, PyTest
- **Databases**: PostgreSQL, MySQL, NoSQL
- **Others**: Agile, Scrum, Jira, Sprint Planning, Cloud-native

Languages
- **Italian**: Native
- **English**: B2, Cambridge certified