

Gilson Vieira de Souza

LinkedIn: in/gilson	Github: /gil-son	gilson.inspire@gmail.com	+55 11 94053-4609
---------------------	------------------	--------------------------	-------------------

Skills

- **Languages & Frameworks:** Python | FastAPI | LangChain | Terraform
- **Cloud & DevOps:** AWS | Azure | CI/CD | GitHub Actions | Docker | Kubernetes | SageMaker
- **MLOps & Tools:** ML Pipelines | Model Deployment | MLflow | Hugging Face (Transformers, Fine-tuning) | Ollama | LLM Applications | Generative AI
- **Best Practices:** Clean Code | SOLID | Design Patterns
- **Observability:** CloudWatch | Datadog | Grafana | Splunk
- **Methodologies:** Agile | Scrum | Kanban
- **Languages:** English – C1 Level (Fluent communication)

Experiences

IBM – Sao Paulo, Brazil

- Machine Learning Engineer / AI Engineer (04/2025 - Present)
 - Developed and supported AI-driven chatbots with NLU/NLG and Generative AI capabilities.
 - Built infrastructure and ML pipelines using AWS, Terraform, and Kubernetes for scalable and reliable deployments.
 - Integrated LLMs (OpenAI, Anthropic) into internal tools and automations.
 - Contributed to AI orchestration efforts using LangChain, FastAPI, and vector databases (Weaviate/FAISS).
 - Led model observability, logging, and CI/CD for production-grade AI services.
 - Delivered solutions used by millions of users in the banking sector.
- Mid-level Backend Developer (08/2022 - 04/2025)
 - Developed backend services using Spring Boot, Spring Batch, Apache Camel, and Apache Kafka to support scalable and resilient financial systems.
 - Built and maintained cloud infrastructure using Terraform (Infrastructure as Code), ensuring consistency and automation across environments.
 - Wrote and maintained unit and integration tests, contributing to code quality and maintainability.
- Software Developer - Junior (06/2021 – 07/2022)
 - Automated card services for over 600k users at Santander using microservices architecture.
 - Enhanced backend systems and developed efficient message queues for real-time data processing.

- Gained hands-on experience with containerized environments and cloud deployments.

UNICID – Sao Paulo, Brazil

- Data Analyst - Junior (07/2019 – 05/2020)
 - Built ETL processes to extract insights from enrollment and retention data.
 - Helped drive marketing and operational strategies to grow the student base.
 - Supported decision-making in marketing and operations using data dashboards.

Education _____

- Bachelor of Computer Science at UNICID - Sao Paulo (01/2017 – 12/2020)

Competencies _____

- **Machine Learning Engineer with 5+ years of experience**, currently focused on **Machine Learning Engineering, AI Engineering, and MLOps**. Skilled in deploying and maintaining ML workflows, integrating AWS services, and automating infrastructure with Terraform. Proficient in Generative AI, including RAG, NLP, NLU, NLG, embeddings, vector databases, and more. Experienced in combining AI models with LangChain to build advanced chatbots and automation solutions.
- Experienced in deploying AI services and ML workloads using Docker and Kubernetes in cloud environments, including Kubernetes orchestration and Amazon EKS for scalable and reliable production systems.
- In addition to backend development expertise in **Python, Java, and Go**, I bring strong knowledge of ML model deployment, observability, and scalable cloud-based pipelines. I have contributed to high-impact projects in the financial sector, serving millions of users.
- Passionate about delivering value through innovative technologies and team collaboration, my goal is to contribute to business success while continuously evolving as an engineer.