Marc Giannuzzi, Data Engineer

(+33)695646454, marc.giannuzzi@gmail.com

LINKS

GitHub
Linkedin

EXPERIENCE

Apr. 2025 — *present*

LCL [SFEIR] — Data Architect

Villejuif, FR

Context: LCL (Le Crédit Lyonnais) is a major French bank, a subsidiary of the Crédit Agricole group, which offers banking services to individuals, professionals, and companies.

Optimization of the internal organization of the SNI entity and implementation of analytical dashboards for its CODIR and other squad leaders. The challenge was to consolidate data from different sources, orchestrate and automate their processing through a scalable and maintainable architecture.

- Audit of the existing IS: analysis of the stacks in place (databases, scripts, tools), comparison of PostgreSQL, MongoDB, Teradata
- Hybrid technical architecture (on-premise first, comparison and proposals for Cloud GCP equivalents)
- Kubernetes orchestration of dockerized Python scripts.
- Storage via PostgreSQL (analytical mode) and S3
- Design of data schemas (tables, columns, indexes).
- Selection of Python libraries (pandas, SQLAlchemy)
- Integration of Jira APIs and other software packages into the pipelines
- Dataviz: dynamic Power BI dashboards for the CODIR.
- Technical supervision: collaboration with 2 data engineers and 1 PO, technical review and implementation support Best practices, guides centralized in Confluence (architecture, code)

Skills: Kubernetes · GCP · GitLab CI/CD · Python · AWS S3 · SQL · MongoDB · PostgreSQL · Teradata · PowerBI · Elastic Stack (ELK)

Mar. 2023 — present

SFEIR — Data Consultant

Neuilly Sur Seine, FR

Context: SFEIR is a French company, the leading Google Cloud Platform GCP partner (Services 2023, Training 2024, EMEA 2025), specializing in software development and technological support for companies. It helps its clients create web, cloud, mobile or data applications using modern technologies.

- **Recruitment** of technical profiles Data Engineers - Data Scientists - ML Ops - Devops - Fullstack

Skills: GCP (Google Cloud Platform)

Jan. 2023 — Jan. 2025   **Engie [Visian] — Data Consultant**   Paris, FR

Context: Engie is the leading supplier of natural gas and green electricity, as well as the third largest group in the world in the energy sector.
Data Engineer part of the agile team (5 people) developing the Mon Bonus Engie / Ecodefi+, ICE (Impact Carbone Elec) and Ecogestes Dynamiques products evolving in the SAFe framework

- Ecodefi+ (Mon Bonus Engie) is a free service offering individuals to be rewarded for their energy sobriety during challenges (**300k customers**)
- ICE (Impact Carbone Elec) allows customers to **analyze carbon consumption peaks in the future and adapt their consumption** accordingly (**1M customers**).
- Dynamic Ecogestures helps to explain the impacts of electrical objects in homes (**1M customers**)

Missions

- Implementation and development of products
- Large-scale data management (databases, flow orchestration...)
- Elaboration of technical and functional KPIs
- Monitoring and Alerting of solutions
- Creation of REST APIs (100k calls/day), latency < 1sec
- Participation in the architecture of AWS micro-services and digital platforms
- Optimizations (memory, execution time, refactoring...)
- Database flow optimizations (structure, indexes...)
- Automated reporting via SNS
- Gitlab CICD
- Writing of roadmaps as part of the migration of a central technology used for several years within the SAFe train
- Participation in technical-functional architectures
- Scaling of Deep Learning algorithms (DeepAR, Temporal Fusion Transformer...)

Skills: c3ai · JavaScript · Jenkins · AWS Lambda · Amazon Kinesis · Databricks · Elastic Stack (ELK) · Python (programming language) · AWS S3 · Terraform · Amazon CloudWatch · Splunk · MLflow · Amazon DynamoDB · TIDB · Apache Airflow · Amazon Simple Queue Service (SQS) · Amazon Simple Notification Service (SNS)

Jan. 2023 — Jan. 2025   **Visian — Data Consultant**   La Défense, FR

Context: Visian is a **consulting** company specializing in projects using large-scale data. My video presentation: Youtube

- Recruitment of technical profiles (Data Engineers, Data Scientists, Devops, Fullstack)
- **Direction of the Pysquad project**, developing digital projects internally
- Coordination of the generic implementation of a generative **Artificial Intelligence (AI) project** for different inference tasks (textual, vision, multimodal etc...) with the use of ML/DL Azure and Azure AI Studio services
- **Creation of a data platform** to help sales to retrieve data on missions, prospects, customers, consultants etc

Skills: Docker · Azure AI Studio · Django · Streamlit · PowerBI · FastAPI · Elastic Stack (ELK) · AWS RDS · Redis · AWS Code Artifacts

Aug. 2021 — Oct. 2022	**Aqua Digital Rising — Lead Data Scientist**	Paris, FR

Context: Aqua Digital Rising is a **fintech startup** that aims to revolutionize sports betting by analyzing all player and team statistics.

Data Management

- Management of storage (creation, structuring, cleaning) of several databases (MongoDB, SQL server, HDF5) from data recovered by the suppliers Sport Radar and Opta
- Structuring of data flows in Python through the different components of the architecture and optimization of the reading speed thanks to the SQL Server indexes (several hundred Go)
- Creation of a highly configurable REST API allowing the front-end team to retrieve database aggregations and other static information
- Temporary backups of API query results thanks to Redis
- Live update of statistical data with the use of RabbitMq
- Analysis and data visualization
- Machine learning algorithms (clustering) and visualizations for the analysis of player and team statistics

Other missions

- Configuration of various Azure services (Azure Functions, Azure Web Services, Azure Artifacts, Azure Pipelines, Azure Keyvault) in line with the company's Cloud strategy
- Quantitative analysis of cryptocurrencies with InfluxDb to feed the investment strategy
- Coordination of 2 engineers (a data scientist and a front-end developer)

Skills: Microsoft SQL Server · Scikit-learn · Redis · JIRA · RabbitMQ · Databases · InfluxDB · Bash · MongoDB · Microsoft Azure

Feb. 2021 — Jul. 2021	**Orange Business Services — Data Scientist**	Paris, FR

Context: Orange Business Services is a **consulting** company in digital transformation in communication projects.

- Analysis and diagnosis of VoIP network anomalies, particularly with machine learning (clustering)
- Creation of a dynamic VoIP network graph with Kibana then Neo4J for easier visualization, integrated into a web page
- Analysis of Devops metrics (CPU usage, storage...) of a cluster of Elastic servers
- Automation of the deletion of Elastic logs with log4j

Note: VoIP = Voice over Internet Protocol is a network that allows telephone calls via the Internet

Skills: HTML · Scikit-learn · Elastic Stack (ELK) · Cascading Style Sheets (CSS) · Python (programming language) · Neo4j · JavaScript · Databases · Bash

Aug. 2020 — Mar. 2021	**Digiberries — Software Engineer**	Paris, FR

Context: Digiberries is a **web referencing agency**, improving the SEO strategies of clients.

- Development of a script allowing the automation of the creation of excel reports from Google Ads data
- Creation of an SEO analysis tool and automated PDF report creation with Lighthouse and NodeJS packages
- Automation of sending emails

Skills: Lighthouse · Python (programming language) · Google Ads · Node.js

Apr. 2020 — Aug. 2020	**Eudonet CRM — Data Scientist**	Courbevoie, FR

Context: Eudonet **develops CRM software** in a B2B context.

- Implementation of the Elastic suite (ElasticSearch, Kibana, Logstash + Filebeat) for the analysis of Windows IIS log files
- Creation of a monitoring tool with dynamic dashboards in Kibana
- Creation of a Python algorithm for detecting a drop in CRM usage by customers. Use of the FBProphet library
- Creation of a dynamic web platform (Javascript, Node.js) allowing a visualization of the results of the previous algorithm
- PowerBI visualizations for the internal management of employee projects
- Creation of powershell scripts for the management of IIS applications
- SQL queries for the creation of reports on the use of customers and the modification of parameters
- Weekly agile sprints

Skills: Microsoft SQL Server · HTML · Windows PowerShell · Elastic Stack (ELK) · Cascading Style Sheets (CSS) · Python (programming language) · Microsoft Power BI · JavaScript · Databases · Bash

Jul. 2019 — Sep. 2019	**Veesion — Software Engineer**	Paris

Context: <u>Veesion</u> is a French scale-up wanting to **automatically detect thefts** in supermarkets.
- Creation of an internal platform allowing to visualize and label (in order to improve the company's theft detection algorithm) live videos with Django
- Classification of data allowing the learning of the computer vision model

Skills: HTML · Cascading Style Sheets (CSS) · Python (programming language) · JavaScript · Amazon S3 · Django

---

PROJECTS

2024	**Artificial Intelligence**

Automated creation of multimodal digital content

2024	**Finance**

Algorithms and visualizations to aid financial analysis (Python)

2021	**Sorting Algorithm**

Sorting algorithm visuali

2020	**NLP (Natural Language Processing)**

Conversational agent displaying the lyrics of a song (Chatbot Messenger)

2020	**Computer vision**

Computer vision algorithm transforming finger gestures into equations via a webcam (Tensorflow, OpenCV)

2019	**Deep Learning**

Health scoring algorithm for cooking recipes (Tensorflow)

EDUCATION

Sep. 2018 — Jun. 2021	**ESILV (École Supérieure d'Ingénieurs Léonard De Vinci)**	Courbevoie, FR

- Master's Degree in Computer Science, Big Data and Connected Objects
- Specialization in Data Science

Sep. 2019 — Jan. 2020	**UCR (University Of California Riverside)**	Riverside, California USA

- <u>CS 235</u> Data Mining Techniques
- <u>CS 171</u> Introduction to Machine Learning and Data Mining
- CS 009 Python programming

Sep. 2016 — Dec. 2018	**CPGE Fénelon Sainte-Marie**	Paris, FR

- MPSI
- MP

LANGUAGES	**English**: Fluent **French**: Native