



Marc Giannuzzi, Data Engineer

(+33) 6 95 64 64 54, marc.giannuzzi@gmail.com

LIENS

[GitHub](#)
[Linkedin](#)

EXPERIENCE

Sept. 2025	Waiyer — Auto entrepreneur	Paris, FR
Aujourd'hui	<p><i>Contexte :</i> <u>Waiyer</u> est une entreprise du digital spécialisée en data et en architecture cloud. Waiyer permet d'optimiser la gestion des données grâce à des solutions sur mesure, en utilisant les dernières technologies. Son but est de concevoir et de mettre en œuvre des solutions efficaces et pérennes pour aider les entreprises à tirer le meilleur parti de leurs données.</p> <ul style="list-style-type: none">• Mise en place de l'architecture des données• Orchestration de pipelines de données• Création de dashboards analytiques• Business Intelligence• Software engineering : développement de scripts, APIs, outils d'automatisation• Vision fonctionnelle : Définition des besoins, des KPIs, des dashboards adaptés• Veille technologique : suivi des évolutions dans le domaine de la data, identification des opportunités d'innovation pour les clients <p><i>Compétences :</i> Kubernetes · Github Actions · Python · AWS S3 · SQL · MongoDB · PostgreSQL · PowerBI · Flutter · Go · Typescript · gRPC · Redis</p>	
Avr. 2025	SFEIR — Data Architect pour LCL	Villejuif, FR
Sept. 2025	<p><i>Contexte :</i> <u>LCL</u> (Le Crédit Lyonnais) est une banque française majeure, filiale du groupe Crédit Agricole, qui propose des services bancaires aux particuliers, professionnels et entreprises.</p> <p>Optimisation de l'organisation interne de l'entité SNI et mise en place des tableaux de bord analytiques pour son CODIR. L'enjeu était de consolider les données issues de différentes sources, orchestrer et automatiser leur traitement via une architecture scalable et maintenable.</p> <ul style="list-style-type: none">• Audit du SI existant : analyse des stacks en place (bases, scripts, outils), comparaison PostgreSQL, MongoDB, Teradata• Architecture technique hybride (on-premise first, comparaison et propositions d'équivalents Cloud GCP)• Orchestration Kubernetes de scripts Python dockerisés.• Stockage via PostgreSQL (mode analytique) et S3• Design des schémas de données (tables, colonnes, index).• Sélection des librairies Python (pandas, SQLAlchemy)• Intégration des API Jira et autres progiciels dans les pipelines• Dataviz : dashboards Power BI dynamiques pour le CODIR.• Encadrement technique : collaboration avec 2 data engineers et 1 PO, revue technique et support à l'implémentation• Bonnes pratiques, guides centralisés dans Confluence (architecture, code) <p><i>Compétences :</i> Kubernetes · GCP · GitLab CI/CD · Python · AWS S3 · SQL · MongoDB · PostgreSQL · Teradata · PowerBI · Elastic Stack (ELK)</p>	
Mar. 2025	SFEIR — Data Consultant	Neuilly Sur Seine, FR
Sept. 2025	<p><i>Contexte :</i> <u>SFEIR</u> est une entreprise française, premier partenaire Google Cloud Platform GCP (Services 2023, Training 2024, EMEA 2025), spécialisée dans le développement de logiciels et l'accompagnement technologique des entreprises.</p> <p>En charge de recrutements de profils techniques: Data Engineers, Data Scientists, ML Ops, Devops, Fullstack</p> <p><i>Compétences :</i> GCP (Google Cloud Platform)</p>	

Jan. 2023

Vision — Data Consultant pour Engie

Paris, FR

Jan. 2025

Contexte : Engie est le premier fournisseur de gaz naturel et d'électricité verte en France.

Data Engineer faisant partie de l'équipe agile (5 personnes) développant les produit Mon Bonus Engie / Ecodefi+, ICE (Impact Carbone Elec) et Ecogestes Dynamiques évoluant dans le framework SAFe

- Ecodefi+ (Mon Bonus Engie) est un service gratuit proposant aux particuliers d'être récompensé sur leur sobriété énergétique lors de challenges (**300k clients**)
- ICE (Impact Carbone Elec) permet aux clients d'**analyser les pics de consommation carbone dans le futur et d'adapter leur consommation** en fonction (**1M clients**).
- Ecogestes Dynamiques contribue à expliquer les impacts des objets électriques dans les foyers (**1M clients**)

Missions

- Implémentation et développement des produits
- Gestion des données à grande échelle (bases des données, orchestration des flux...)
- Elaboration des KPIs techniques et fonctionnelles
- Monitoring et Alerting des solutions
- Création d'APIs REST (100k appels/jour), latence < 1sec
- Participation à l'architecture des micro-services AWS et des plateformes digitales
- Optimisations (mémoire, temps d'exécution, refactoring...)
- Optimisations des flux bases de données (structure, index...)
- Reporting automatisé via SNS, Gitlab CICD
- Rédaction de roadmaps dans le cadre de la migration d'une technologie centrale au sein du train SAFe
- Participations aux architectures technico-fonctionnelles
- Scaling d'algorithmes de Deep Learning (DeepAR, Temporal Fusion Transformer...)

Compétences : c3ai · JavaScript · Jenkins · AWS Lambda · Amazon Kinesis · Databricks · Elastic Stack (ELK) · Python · AWS S3 · Terraform · Amazon CloudWatch · Splunk · MLflow · Amazon DynamoDB · TIDB · Apache Airflow · SQS · SNS

Jan. 2023

Vision — Data Consultant

La Défense, FR

Jan. 2025

Contexte : Vision est une entreprise de **conseil** spécialisée dans les projets utilisant la donnée à grande échelle. Ma présentation vidéo de l'entreprise et du métier: [Youtube](#)

- Recrutements de profils techniques (Data Ingénieurs, Data Scientists, Devops, Fullstack)
- **Direction du projet Pysquad**, développant en interne des projets digitaux
- Coordination de l'implémentation générique d'un **projet d'IA générative** pour différentes tâches d'inférence avec Azure AI Studio
- **Création d'une data plateforme** en vue d'aider les sales à récupérer les données de missions, prospects, clients, consultants etc

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

Août 2021

Aqua Digital Rising — Lead Data Scientist

Paris, FR

Oct. 2022

Contexte : Aqua Digital Rising est une **startup fintech** qui ambitionne de révolutionner le pari sportif.

Gestion de la donnée

- Gestion du stockage (MongoDB, SQL server, HDF5) à partir de Sport Radar et Opta
- Structuration des flux de données en Python et optimisation de la vitesse de lecture grâce aux index SQL Server
- Création d'une API REST hautement configurable pour l'équipe front-end
- Sauvegardes temporaires via Redis et mise à jour en direct avec RabbitMq
- Analyses, data visualisation et algorithmes de machine learning (clustering)

Autres missions

- Configuration divers services Azure (Functions, Web Services, Artifacts, Pipelines, Keyvault)
- Analyses quantitatives de crypto-monnaies avec InfluxDb pour alimenter la stratégie d'investissement
- Coordination de 2 ingénieurs (un data scientist et un développeur front-end)

Compétences : Microsoft SQL Server · Scikit-learn · Redis · JIRA · RabbitMQ · Bases de données · InfluxDB · Bash · MongoDB · Microsoft Azure

Fév. 2021

Orange Business Services — Data Scientist

Paris, FR

Juil. 2021

Contexte : Société de **conseil** dans la transformation digitale dans des projets de communication.

- Analyses et diagnostics des anomalies du réseau VoIP via machine learning (clustering)
- Création d'un graphe de réseau VoIP dynamique avec Kibana puis Neo4J intégré dans une page web
- Analyses des métriques Devops d'un cluster de serveurs Elastic
- Automatisation de la suppression des logs Elastic avec log4j

Compétences : HTML · Scikit-learn · Elastic Stack (ELK) · CSS · Python · Neo4j · JavaScript · Bases de données · Bash

Août 2020

Digiberries — Software Engineer

Paris, FR

Mars 2021

Contexte : Agence de référencement web, améliorant les stratégies SEO des clients.

- Développement d'un script d'automatisation de création de rapports excel (Google Ads)
- Crédit d'un outil d'analyse SEO et rapport automatisé PDF avec Lighthouse et NodeJS
- Automatisation d'envoi de mails

Compétences : Lighthouse · Python · Google Ads · Node.js

Avr. 2020

Eudonet CRM — Data Scientist

Courbevoie, FR

Août 2020

Contexte : Développe des logiciels CRM dans un context B2B.

- Mise en place de la suite Elastic (ELK + Filebeat) pour l'analyse de logs Windows IIS
- Création d'un outil de monitoring et algorithme Python de détection de baisse d'utilisation (FBProphet)
- Création d'une plateforme web (Javascript, Node.js) de visualisation dynamique
- Visualisations PowerBI interne et création de scripts powershell pour la gestion d'applications IIS
- Requêtes SQL pour rapports et modification de paramètres, sprints agiles hebdomadaires

Compétences : Microsoft SQL Server · HTML · PowerShell · ELK · Python · Power BI · JavaScript · Bash

Juil. 2019

Veesion — Software Engineer

Paris

Sep. 2019

Contexte : Scale-up française de détection automatique de vols via IA.

- Crédation d'une plateforme interne de visualisation et labellisation vidéo avec Django
- Classification de données pour l'apprentissage du modèle de computer vision

Compétences : HTML · CSS · Python · JavaScript · Amazon S3 · Django

PROJETS PERSONNELS

- **Intelligence Artificielle** : Création automatisée de contenus digitaux multimodaux
- **Finance** : Algorithmes et visualisations d'aide à l'analyse financière (Python)
- **Algorithmes de tri** : Plateforme de visualisation d'algorithmes de tri (React)
- **NLP** : Agent conversationnel lyrics de chanson (Chatbot Messenger)
- **Computer Vision** : Transformation des gestes en équations (Tensorflow, OpenCV)
- **Deep Learning** : Algorithme de scoring santé de recettes cuisine (Tensorflow)

CERTIFICATIONS

2025

GCP - Associate Cloud Engineer / Professional Data Engineer

EDUCATION

Sep. 2018

ESILV (École Supérieure d'Ingénieurs Léonard De Vinci)

Courbevoie, FR

Juin 2021

- Master en Informatique, Big Data et Objets Connectés - Spécialisation en Data Science

Sep. 2019

UCR (University Of California Riverside)

Riverside, CA, USA

Jan. 2020

- Data Mining Techniques (CS 235), Machine Learning (CS 171), Python programming (CS 009)

Sep. 2016

CPGE Fénelon Sainte-Marie

Paris, FR

Déc. 2018

- MPSI / MP (Classe Préparatoire aux Grandes Écoles)

LANGAGES

English: Fluent

Français: Natif