



# Marc Giannuzzi, Software Engineer

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

LIENS

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

## EXPERIENCE

Janvier 2023 — *present*

### Visian — Machine Learning Engineer

La Défense, FR

*Contexte* : Visian est une entreprise de **conseil** spécialisée dans les projets utilisant la donnée à grande échelle. Ma présentation vidéo: [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'Intelligence Artificielle (IA)** générative pour différentes tâches d'inférence (textual, vision, multimodal etc...) avec l'utilisation des services ML/DL Azure et 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

Janvier 2023 — *présent*

### Engie — Consultant Data

Paris, FR

*Contexte* : Engie est le premier fournisseur de gaz naturel et d'électricité verte, ainsi que le troisième plus grand groupe mondial dans le secteur de l'énergie.

**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
- Scaling d'algorithmes de Deep Learning (DeepAR, Temporal Fusion Transformer...)
- 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 utilisée depuis plusieurs années au sein du train SAFE
- Participations aux architectures technico-fonctionnelles

*Compétences* : c3ai · JavaScript · Jenkins · AWS Lambda · Amazon Kinesis · Databricks · Elastic Stack (ELK) · Python (langage de programmation) · AWS S3 · Terraform · Amazon CloudWatch · Splunk · MLflow · Amazon DynamoDB · TIDB · Apache Airflow · Amazon Simple Queue Service (SQS) · Amazon Simple Notification Service (SNS)

*Contexte* : Aqua Digital Rising est une **startup fintech** qui ambitionne de révolutionner le pari sportif en analysant toutes les statistiques des joueurs et des équipes.

#### Gestion de la donnée

- Gestion du stockage (création, structuration, nettoyage) de plusieurs bases de données (MongoDB, SQL server, HDF5) à partir de données récupérées par les fournisseurs Sport Radar et Opta
- Structuration des flux de données en Python à travers les différents composants de l'architecture et optimisation de la vitesse de lecture grâce aux index SQL Server (plusieurs centaines de Go)
- Création d'une API REST hautement configurable permettant à l'équipe front-end de récupérer les agrégations de la base de données et d'autres informations statiques
- Sauvegardes temporaires des résultats de requêtes de l'API grâce à Redis
- Mise à jour en direct des données statistiques avec l'utilisation de RabbitMQ
- Analyses et data visualisation
- Algorithmes de machine learning (clustering) et visualisations pour l'analyse des statistiques des joueurs et des équipes

#### Autres missions

- Configuration de divers services Azure (Azure Functions, Azure Web Services, Azure Artifacts, Azure Pipelines, Azure Keyvault) en lien avec la stratégie Cloud de l'entreprise
- 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

*Contexte* : Orange Business Services est une société de **conseil** dans la transformation digitale dans des projets de communication.

- Analyses et diagnostics des anomalies du réseau VoIP, notamment avec du machine learning (clustering)
- Création d'un graphe de réseau VoIP dynamique avec Kibana puis Neo4J pour une visualisation plus facile, intégré dans une page web
- Analyses des métriques Devops (utilisation CPU, stockage...) d'un cluster de serveurs Elastic
- Automatisation de la suppression des logs Elastic avec log4j

**Note:** VoIP = Voice over Internet Protocol est un réseau qui permet les appels téléphoniques via Internet

*Compétences* : HTML · Scikit-learn · Elastic Stack (ELK) · Feuilles de style en cascade (CSS) · Python (langage de programmation) · Neo4j · JavaScript · Bases de données · Bash

*Contexte* : Digiberries est une **agence de référencement web**, améliorant les stratégies SEO des clients.

- Développement d'un script permettant l'automatisation de création de rapports excel à partir de données Google Ads
- Création d'un outil d'analyse SEO et de création de rapport automatisé PDF avec Lighthouse et des packages NodeJS
- Automatisation d'envoi de mails

*Compétences* : Lighthouse · Python (langage de programmation) · Google Ads · Node.js

*Contexte* : Eudonet développe des logiciels CRM dans un contexte B2B.

- Mise en place de la suite Elastic (ElasticSearch, Kibana, Logstash + Filebeat) pour l'analyse de fichiers logs Windows IIS
- Création d'un outil de monitoring avec des dashboards dynamiques dans Kibana
- Création d'un algorithme Python pour la détection de baisse d'utilisation du CRM des clients. Utilisation de la librairie FBProphet
- Création d'une plateforme web (Javascript, Node.js) dynamique permettant une visualisation des résultats de l'algorithme précédent
- Visualisations PowerBI pour la gestion interne des projets des employés
- Création de scripts powershell pour la gestion d'applications IIS
- Requêtes SQL pour la création de rapports sur l'utilisation des clients et la modification de paramètres
- Sprints agiles hebdomadaires

*Compétences* : Microsoft SQL Server · HTML · Windows PowerShell · Elastic Stack (ELK) · Feuilles de style en cascade (CSS) · Python (langage de programmation) · Microsoft Power BI · JavaScript · Bases de données · Bash

Veesion est une scale-up française voulant **détecter automatiquement les vols** dans des supermarchés.

- Création d’une plateforme interne permettant de visualiser et labeliser (afin d’améliorer l’algorithme de détection de vols de l’entreprise) en direct des vidéos avec Django
- Classification de données permettant l’apprentissage du modèle de computer vision Veesion est une scale-up française voulant détecter automatiquement les vos dans des supermarchés.
- Création d’une plateforme interne permettant de visualiser et labeliser (afin d’améliorer l’algorithme de détection de vols de l’entreprise) en direct des vidéos avec Django
- Classification de données permettant l’apprentissage du modèle de computer vision

*Compétences* : HTML · Feuilles de style en cascade (CSS) · Python (langage de programmation) · JavaScript · Amazon S3 · Django

PROJETS

2024	<b>Intelligence Artificielle</b>
	Création automatisée de contenus digitaux multimodaux
2024	<b>Finance</b>
	Algorithmes et visualisations d’aide à l’analyse financière (Python)
2021	<b>Algorithmes de tri</b>
	Plateforme de visualisation d’algorithmes de tri (React)
2021	<b>Sémantique du Web</b>
	Site affichant la disponibilité des vélos/trains (RDF)
2020	<b>NLP (Natural Language Processing)</b>
	Agent conversationnel affichant les paroles d’une chanson (Chatbot Messenger)
2020	<b>Computer vision</b>
	Algorithme de computer vision transformant les gestes du doigt en équations via une webcam (Tensorflow, OpenCV)
2019	<b>Deep Learning</b>
	Algorithme de scoring santé de recettes cuisine (Tensorflow)

EDUCATION

Sept. 2018 — Juin 2021	<b>ESILV</b>	Paris, FR
Sept. 2016 — Déc. 2018	<b>CPGE Fénélon Sainte-Marie (MPSI-MP)</b>	Paris, FR

LANGAGES	<b>English:</b> Fluent	<b>Français:</b> Natif
----------	------------------------	------------------------