

COORDONNÉES



eliseetegue@gmail.com



06 02 39 25 62



<https://www.linkedin.com/in/elise%C3%A9e-tegue-yombi-a1455720a/>



<https://eliseetegue.vercel.app/>



<https://github.com/el-tegy/>

COMPÉTENCES GÉNÉRALES

- Expérience en écriture de code industrialisable et réutilisable pour créer des pipelines de données de qualité
- Expérience dans l'ingestion, le stockage, et les transformations de données sur le cloud (AWS, Azure & GCP) avec Databricks
- Solides compétences interpersonnelles et communication, écrites et orales (quart-de-finaliste d'un concours d'éloquence)

COMPÉTENCES TECHNIQUES

- Big Data : Apache Kafka, Apache Spark, Databricks, Elasticsearch, Apache Nifi, dbt
- Langages : Python 3.9, SQL, PySpark, SparkSQL, NoSQL, HCL
- Bases de données : MongoDB, Cassandra, DynamoDB, Neo4j, Teradata
- Frameworks : Sphinx, Scrapy, Pytest, Pylint, MyPy, Black, Isort, Poetry
- DevOps : Terraform, Github Actions, Gitlab CI, python-semantic-release
- Méthode : Scrum
- Certifications : 6 x Databricks, dbt fundamentals

MÉDAILLES GITHUB

- Pull Shark (x2),
- Quickdraw,
- Pair Extraordinaire,
- Yolo

TEGUE YOMBI Elisée

Recherche d'un CDI en Data Ingénierie à partir de septembre 2024.



FORMATIONS ACADÉMIQUES

CYCLE INGÉNIEUR BIG DATA - ESIGELEC, Rouen, France **2021-2024**

CLASSES PRÉPARATOIRES - PrépaVogt, Yaoundé, Cameroun **2019-2021**



EXPÉRIENCES PROFESSIONNELLES

Apprenti Data Engineer – ENGIE, Paris

09/2022 – 09/2024

Principales missions :

- Contribution à l'implémentation d'une librairie Python interne mettant des connecteurs au Data Lake à disposition des autres équipes ;
- Mise en œuvre d'un pipeline de données de production et de consommation d'électricité des clients en mode CDC avec Spark 3 et Databricks ;
- Rédaction de documentation technique avec Confluence ;
- Conception d'un package Python et d'un Git flow dans le cadre de la MEP du code d'une équipe de Data Scientists ;
- Refactoring de code pour une adhérence aux standards Pylint (note > 7/10) ;
- Contribution à l'élaboration des tâches de sprint avec Jira (Scrum) ;
- Modélisation, optimisation du Système d'Information et suivi applicatif.

Environnement technique : Python 3.9, PySpark 3, Pylint, Pytest, Databricks, AWS, Spark SQL 3, Airflow, Gitlab, Teradata, Debian 11

Business Analyst en mission chez ENGIE Belgique **04/2023 – 07/2023**

Principales missions :

- Suivi des applications SAS de production avec l'outil Visual TOM (VTOM) ;
- Préparation et modélisation de données ;
- Conversion de rapports du format BEX au format Analysis for Office.

Environnement technique : SAP BW, SAS, VTOM, Power BI, Azure



PROJETS ACADÉMIQUES & PERSONNELS

Conception et réalisation d'une solution de streaming de données

Reddit en Python

12/2023 – 01/2024

Principales missions :

- Mise en place d'un cluster Kafka ayant 03 brokers à l'aide de Docker ;
- Utilisation de Praw API pour récupérer les informations pertinentes des commentaires Reddit (auteur, contenu, ...) et envoi au producer Kafka ;
- Traitement en temps réel de ces informations avec Spark Streaming ;
- Ecriture batch des données traitées dans Cassandra (moving average) ;
- Création d'un dashboard Grafana montrant l'évolution du score d'analyse des sentiments pour chaque subreddit ;
- Orchestration de la chaîne complète avec Kubernetes et Terraform.

Environnement technique : Docker, Spark 3, Terraform, Zookeeper, Python 3.9, GitHub

COORDONNÉES

- @ eliseetegue@gmail.com
- 06 02 39 25 62
- <https://www.linkedin.com/in/eliseetegue-yombi-a1455720a/>
- <https://eliseetegue.vercel.app/>
- <https://github.com/el-tegy/>

COMPÉTENCES GÉNÉRALES

- Expérience en écriture de code industrialisable et réutilisable pour créer des pipelines de données de qualité
- Expérience dans l'ingestion, le stockage, et les transformations de données sur le cloud (AWS, Azure & GCP) avec Databricks
- Solides compétences interpersonnelles et communication, écrites et orales (quart-de-finaliste d'un concours d'éloquence)

COMPÉTENCES TECHNIQUES

- Big Data : Apache Kafka, Apache Spark, Databricks, Elasticsearch, Apache Nifi, dbt
- Langages : Python 3.9, SQL, PySpark, SparkSQL, NoSQL, HCL
- Bases de données : MongoDB, Cassandra, DynamoDB, Neo4j, Teradata
- Frameworks : Sphinx, Scrapy, Pytest, Pylint, MyPy, Black, Isort, Poetry
- DevOps : Terraform, Github Actions, Gitlab CI, python-semantic-release
- Méthode : Scrum
- Certifications : 6 x Databricks, dbt fondamentaux

MÉDAILLES GITHUB

- Pull Shark (x2),
- Quickdraw,
- Pair Extraordinaire,
- Yolo



PROJETS ACADÉMIQUES & PERSONNELS

Réalisation d'une solution de traitement puis d'analyse de données en temps réel

02/2024

Principales missions :

- Utilisation de data generator pour générer les données de streaming et envoi de celles-ci à Apache Kafka ;
- Traitement des données brutes avec PySpark et utilisation d'HDFS comme système de stockage ;
- Contrôle des ressources avec Apache Hadoop YARN ;
- Lecture des données de streaming depuis le topic Kafka l'aide de Spark Streaming ;
- Ecriture des données de streaming dans Elasticsearch et visualisation de celles-ci avec Kibana ;
- Ecriture des données de streaming dans MinIO (Object Storage) ;
- Utilisation d'Apache Airflow pour orchestrer la pipeline complète.

Environnement technique : Apache ZooKeeper, Apache Kafka, Apache Hadoop YARN, Apache Hadoop HDFS, Apache Airflow, Apache Spark, Elasticsearch, Kibana, MinIO, Docker, GitHub

Mise en œuvre d'un système similaire à cookiecutter basé sur une base de données graphe

11/2023

Principales missions :

- Implémentation des fonctions CRUD relatives aux bases de données MongoDB et Neo4j pour le back-end ;
- Développement de formulaires dynamiques basés sur Neo4j ;
- Intégration du système de recommandation de template ;
- Configuration et gestion de l'infrastructure du projet dans GCP avec Terraform (un pare-feu et 04 VMs ayant chacune un disque de stockage externe) ;
- Gestion des dépendances avec Poetry ;
- Maintien de la qualité du code avec Pylint, Black, Isort, MyPy et Darglint en utilisant pre-commit ;
- Rédaction de la documentation des différents modules du back-end avec Sphinx et déploiement sur Read the Docs ;
- Réalisation des tests sur les différents composants du back-end avec Pytest ;
- Contribution à l'élaboration des tâches de sprint (Scrum) ;
- Rédaction d'un cahier d'activité pour chaque sprint.

Environnement technique : GCP, Python 3.9, Poetry, Pre-Commit, Pytest, Pylint, MyPy, Isort, Darglint, Terraform, Sphinx, Read the Docs, GitHub

Implémentation d'un comparateur de prix entre deux sites de vente d'iPhones avec Docker

10/2023

Principales missions :

- Utilisation de Scrapy pour récupérer les informations pertinentes des articles (nom, prix, lien vers l'article, ...) des deux sites (Pixmania et Idealo) ;
- Conservation des informations dans une base de données MySQL se trouvant dans un conteneur Docker ;
- Ajout d'un conteneur comportant un serveur web Apache pour visualiser les informations de la base de données ;
- Création d'une UI pour afficher pour chaque article le site le moins coûteux.

Environnement technique : Docker, Docker Compose, Python 3.9, Pylint