



Yahya KHALID

Data Engineer

93260 Les Lilas, Île-de-France | +33 7 87 67 33 05 | yahyakkhalid@outlook.fr

Ingénieur Data spécialisé en Data pipelines, ETL et monitoring, avec expertise en traitement et analyse des données, technologies Big Data et conception de pipelines de données. Je cherche à contribuer à des projets innovants au sein d'une entreprise axée sur l'excellence technologique et la croissance professionnelle.

EXPÉRIENCES PROFESSIONNELLES

Showroomprivé

Saint-Denis, Île-de-France

Stagiaire Data Engineer

03 Avril 2023 – 29 Septembre 2023

- Implémentation d'une plateforme de traitement de données.
 - Implémenter les composants de la plateforme tout en assurant simultanément la scalabilité horizontale et une haute disponibilité.
 - Établir un système de surveillance et d'alerte pour assurer le suivi continu de la plateforme.
- Visualisation des données.
 - Traiter et analyser les données du Data Warehouse Redshift pour les rendre exploitables pour la visualisation.
 - Élaborer des tableaux de bord contenant des visualisations alimentées par des requêtes SQL, permettant à l'équipe de prévisualiser les données et d'identifier ainsi d'éventuels problèmes liés à la qualité des données.
 - Établir des seuils pour déclencher des alertes en cas de problèmes liés à la qualité des données.
- **Technologies Utilisées:** AWS (S3, EC2, RDS, Athena, Glue, Redshift), Airflow, PostgreSQL, Grafana, Prometheus

NanoStrategy

Casablanca, Maroc

Stagiaire Data Engineer

01 Mars 2022 - 31 Août 2022

- Implémentation d'une plateforme de traitement de données dans le cadre de la migration vers le cloud.
 - Élaborer des scripts Terraform et Ansible visant à automatiser la configuration et le déploiement des infrastructures cloud.
 - Implémenter les composants de la plateforme tout en assurant simultanément la scalabilité horizontale et une haute disponibilité.
 - Établir un système de surveillance et d'alerte pour assurer le suivi continu de la plateforme.
- Concevoir et déployer des pipelines ETL pour extraire des données, les transformer et les charger dans le Data Warehouse en mode batch et temps réel.
- **Technologies Utilisées:** Ansible, Terraform, Airflow, Spark, Kafka, PostgreSQL, Grafana, Prometheus

FORMATION

Université Paris-Saclay

M2 en Data Management in large-scale Distributed Systems

Février 2024

- Extraire, analyser et exploiter l'information et la connaissance stockées dans une infrastructure de gestion de données à large échelle.
- Développer et déployer des applications de gestion de données orientés services.

Université Hassan 2, Maroc

Master Professionnel en Big Data and Cloud Computing

Octobre 2022

- Concevoir, mettre en œuvre et administrer des solutions Big Data et Cloud Computing pour répondre aux besoins croissants de capacité de calcul et de stockage des entreprises.

COMPÉTENCES

Python, Bash, SQL, PostgreSQL, MongoDB, Hive, Spark, Kafka, Airflow, AWS (S3, RDS, Redshift, Athena, Glue), Terraform, Ansible, Grafana, Prometheus, Loki, Git, Jira, Confluence.

LANGUES

- **Anglais :** Professionnel