Jules Carpentier

Big Data Engineer

jules.carpentier@example.com

+33 6 405 72 31

Grenoble

"Big Data Engineer spécialisé dans télémétrie réseau distribuée. Je construis des produits de données fiables avec Kafka, Spark, Hadoop, Hive, Kubernetes et relie les choix techniques à des résultats mesurables."

Education

Grenoble INP

Master informatique · Data engineering et systèmes logiciels

2015.09 - 2019.06

Skills

TECHNOLOGIES

Kafka Spark Hadoop Hive Kubernetes

LIVRAISON

Modélisation des données

Qualité des données CI/CD

Observabilité Documentation

COLLABORATION

Ateliers de cadrage

Réponse aux incidents Mentorat

Livraison agile

Languages

Français — Langue maternelle

Anglais — Niveau professionnel

Certifications

Cloudera Data Platform Generalist

Organisme de certification professionnelle

2024.05

Data Reliability Workshop

Data Engineering Guild

2023.08

Work Experience

Big Data Engineer

Alpes Réseaux — 2022.01 - Present

• Conçu des pipelines Kafka et Spark pour télémétrie réseau distribuée chez Alpes Réseaux, avec les équipes produit et plateforme sur 8,7 Po d'historique; la part des traitements terminés à l'heure est passée de 92 % à 99,6 %. (rôle actuel, point 1)

• Refondu les modèles Hadoop et le partitionnement du projet Observatoire du réseau en encadrant une équipe de 9 ingénieurs; le temps médian des requêtes est descendu de 24 à 4,1 secondes. (rôle actuel, point 2)

• Instrumenté les contrôles de schéma, de fraîcheur et de lineage sur 38 tables de production pour télémétrie réseau distribuée; avec 14 analystes, les incidents de données ont baissé de 63 %. (rôle actuel, point 3)

• Automatisé les backfills, alertes et runbooks des jobs Hive à Grenoble pour 12 marchés; le support d'astreinte a économisé 86 heures par trimestre. (rôle actuel, point 4)

Analytics Engineer

Alpes Réseaux Data Studio — 2019.04 - 2021.12

• Conçu des pipelines Kafka et Spark pour télémétrie réseau distribuée chez Alpes Réseaux, avec les équipes produit et plateforme sur 8,7 Po d'historique; la part des traitements terminés à l'heure est passée de 92 % à 99,6 %. (rôle précédent, point 1)

• Refondu les modèles Hadoop et le partitionnement du projet Observatoire du réseau en encadrant une équipe de 6 ingénieurs; le temps médian des requêtes est descendu de 24 à 4,1 secondes. (rôle précédent, point 2)

• Instrumenté les contrôles de schéma, de fraîcheur et de lineage sur 38 tables de production pour télémétrie réseau distribuée; avec 5 équipes applicatives, les incidents de données ont baissé de 63 %. (rôle précédent, point 3)

• Automatisé les backfills, alertes et runbooks des jobs Hive à Grenoble pour 4 unités métier; le support d'astreinte a économisé 86 heures par trimestre. (rôle précédent, point 4)

Projects

Observatoire du réseau

Création d'un produit de données documenté pour télémétrie réseau distribuée, avec transformations testées et livraison prête pour l'analyse.

Kafka

Spark

Hadoop

Hive