

Maxime GILLOT

Nantes, France.

www.linkedin.com/in/maximegillot/

29 ans

Permis B

Data Engineer

—

Développeur Big Data

FORMATION

2016 – 2017	Master 2 recherche informatique – Université d'Angers
2015 - 2016	Master 1 informatique - Université d'Angers
2011 - 2014	Licence scientifique mention INFORMATIQUE – Université d'Angers

EXPÉRIENCE PROFESSIONNELLE

Juillet 2020

-

En ce moment

Data Engineer AWS / Expert technique Spark – MyMoneyBank - freelance

Au sein de l'équipe Data publishing:

- Conception, Développement, Tester, Déployer et assurer le RUN de différents flux de données et applications afin de produire des reportings opérationnels.
 - Spark
 - Scala
 - Yarn
 - Postgresql
 - Java 8
 - AWS : S3, EMR, Athena
 - Sharepoint
 - Excel
 - Documentation Antora
- Intégration continue : Gitlab CI/CD, airflow
- Ordonnancement: Airflow / Python , Création de DAG.
- Optimisation job Spark.
- Soutient aux différentes équipes Spark: optimisation de job, debug, bonne pratique de développement, tests unitaires.
- Participation active aux développement des librairies communes MMB.
- Création de rapport Excel.
- Création de batch JAVA : Java / Spring.
- Rédaction de la documentation technique.
- Challenger le mapping fonctionnelles des données.

Environnement Technique :

AWS : S3, Athena, EMR, RDS, System Manager. Spark 2.6/3.1.1, scala/Java 8/11/14/17, gitlab Ci/CD, Kibana, Airflow, Sharepoint, Méthode agile : Scrum/Kanban

Mai 2020

-

Juillet 2020

Data Engineer / référent technique Spark – Sopra Steria - freelance

Dans le cadre de la création d'un starter kit Spark à la CNAM :

- Concevoir, développer, documenter, tester des Flux Spark :
 - Spark
 - hadoop
 - HBase
 - Java 8
 - Junit
- Développement dans un environnement sécurisé : Knox, Ranger, ranger KMS, Kerberos
- Dévops : jenkins, GitLab, artifactory
- Scripts linux/yarn/hadoop/hdfs : cycle de vie des applicatifs

Environnement technique :

Écosystème Hortonwork (Hadoop, Hdfs, YARN, Spark 2.3.2) Java 8, Linux, Jenkins, gitlab, Docker, IntelliJ, maven

Mars 2018

-

Avril 2020

Data Engineer – Sopra Steria - freelance

Dans le cadre d'un POC (Proof Of Concept), mise en place d'une filière Big Data à la CNAM (from scratch) :

- Concevoir, développer, documenter, tester et mettre en production différents flux pour alimenter un data lake :
 - Spark streaming
 - Batch Spark
 - Kafka stream, Kafka connect, custom SMT / Avro + schema registry
 - Flow Nifi, script Groovy, processors custom (Java 8)
 - Batch java/Hadoop
 - Montée de version applicative vers Hadoop 3 (Spark, Kafka, Hive)
- Dévops : Jenkins, Git/GitLab, Artifactory, Saltstack, Maven
- Scripts linux/yarn/hadoop/hdfs : cycle de vie des applicatifs
- Monitoring : création de dashboards Grafana (kafka,Yarn,nifi), configuration JMX, Prometheus, graphite exporter
- Scripts Hive, notebook Zeppelin (Livy, spark)
- Développement dans un environnement sécurisé : Knox, Ranger, ranger KMS, Kerberos
- Participation à la configuration des machines Spark, Hadoop, Nifi, YARN
- Mise en place d'un environnement de développement local (Docker, Hortonworks)
- Gouvernance de données avec Atlas, création linéage
- Méthode agile : SCRUM, iceScrum.
- Environnement de test : Junit5, hadoop-mini-cluster, Nifi mock framework.

Environnement technique :

Écosystème Hortonwork (Hadoop, Hdfs, YARN, Spark, Kafka (HDP et Confluent) , Nifi, Zookeeper, Hive, Ambari, ...), Avro, Java 8, Linux, Jenkins, gitlab, Docker, confluent, groovy, IntelliJ/Éclipse, maven, Sonar

Consultant Big Data – Bouygues Telecom

- Mise à jour et création de flux de données
 - Kafka stream
 - Flink
 - Hive
 - MongoDB
 - Oozie
 - Hdfs
 - Grafana
 - Java 8
- Création de script Hql et Teradata pour la migration, création et mise à jour de

Octobre 2017

-

Mars 2018

table.

- Monitoring : création de dashboards Grafana
- Déploiement/build : Maven, Git/GitLab, SVN
- Environnement de test : Junit 4
- Étude d'impact
- Méthode agile : SCRUM, Jira

Environnement technique :

Écosystème Cloudera, Hadoop/hdfs, Hive, impala, Mongoddb, Flink, oozie, Java 8, TeraData, git, svn

Mars 2017

Développeur Big Data – Sopra Steria

Dans le cadre de mon stage de fin d'études :

Septembre 2017

- Concevoir, développer, documenter, tester différents batches
 - Spark Batch et streaming
 - Cassandra
 - Hadoop (HDFS, YARN ...)
 - Hive
 - Java 8
 - Scala
 - Spring

- Translation de batch java/Spark vers Scala/Spark
- méthode agile : SCRUM, jira

Environnement technique :

Écosystème big data, (datastax/cloudera) Hadoop/Hdfs, Hive, Cassandra, git, java et scala, scrum.

*Autres projets
personnels*

- Création de formation sur Apache Kafka et Apache Nifi sur Udemy
- Création de jeux vidéos, développement blockchain, site internet ...
la liste disponible sur [mon site internet](#)

COMPÉTENCES

Langue

Anglais professionnel

Big Data

Spark, Kafka, Yarn, Nifi, Hdfs, Hadoop, Nifi, Avro, Schema Registry, Hive, Java/Scala ...

Environnement

Linux, Windows, Git/GitLab, Eclipse/Intellij, Maven, Jenkins, Docker, Scrum

Devops

Jenkins, Maven, artifactory, Ansible