

# Hadrien Négros

13 rue Duc  
75018 Paris  
☎ 07 86 51 20 36  
✉ [hadrien.negros@gmail.com](mailto:hadrien.negros@gmail.com)  
30 ans

## Expériences

- Nov 2019 – **Senior Data Engineer**, *Ermes*.  
Fév 2020 Création et mise en place d'un cluster Spark sur un Kubernetes managé (OVH Public Cloud), avec Openstack Swift en Data Layer. Création du cluster, des outils d'administration de celui ci (playbooks ansible), transfert de workflows vers cette nouvelle plateforme, accompagnement des data scientist, mise en place de bonnes pratiques.
- Jan 2019 – **Senior Data Engineer**, *RelevanC*.  
Juin 2019 Administration d'un cluster MapR (20 noeuds). Création de workflows pyspark orchestrés par airflow. Gestion de l'infrastructure Data (~40 serveurs dédiés). Création de playbooks Ansible pour l'installation et la maintenance des serveurs. Installation de 2 clusters Elasticsearch. Installation d'un cluster Kubernetes on premise. Création d'une CI/CD via gitlab pour les projets Data, avec déploiement sur Kubernetes.
- Dec 2014 – **Data Engineer**, *EDF*.  
Dec 2018 **Pilote de projet** Prise en charge de la migration de l'outil de tracking web d'EDF (Comscore → Google Analytics) pour nos besoins analytiques internes: refonte complète du workflow d'intégration et évolution des DataMarts. Interface avec les APIs GCP (*Python*). Historisation, réconciliation avec les données Comscore et rapprochement avec les données client (*Hive*). Transformations et diffusion de l'identifiant client (*PySpark*).  
**Expert technique Hadoop pour les POCs de la Direction Numérique** : Mise en place d'un job Spark en java pour la génération d'un très grand nombre de graphiques personnalisés destinés à être envoyés par mail aux clients particulier. (passé en production) Création d'un job de classification de données textuelles avec la MLlib de Spark. (abandonné)  
**Développeur Hadoop pour le DataLake de la direction commerce d'EDF** : Divers travaux d'intégration de données nécessitant des développements Java (Map Reduce et Spark), Shell et Python pour la mise à disposition et l'enrichissement de données sur Hive et HBase. Ordonancement de ces travaux avec Oozie. Mise en place d'une stack ELK pour le monitoring applicatif.
- depuis Sept 2018 **Enseignant Vacataire**, *Telecom ParisTech*.  
2018 Enseignement du module "Introduction au Framework Hadoop" du mastère Big Data: Gestion et Analyse des Données Massives.
- depuis Fév 2018 **Enseignant Vacataire**, *Université Paris 1 Panthéon-Sorbonne*.  
2018 Construction et enseignement du cours "Introduction aux Big Data et Data Mining" du Master 2 "Management des systèmes d'information".
- Mar-Sep 2014 **IT Specialist Intern**, *IBM*.  
Exploitation des outils Big Data d'IBM pour faciliter la résolution des incidents dans un Cloud Data Center. Création d'une application BigInsights (La distribution Hadoop d'IBM) pour l'analyse de tickets d'incidents en langage naturel. Mise en place d'outils d'analyse de logs: Extraction, indexation, root cause analysis...
- Sep-Déc 2013 **Enseignant Vacataire**, *Université Montpellier 2*.  
Encadrement des Travaux Pratiques de l'Unité d'Enseignement "Concepts de base en informatique".

## Diplômes et Études

- 2014 **Master Données, Connaissances et Langage naturel**, *Université Montpellier 2*.  
Spécialisation dans les domaines de l'ingénierie des connaissances, du traitement automatique du langage naturel et de la fouille de données.
- 2011 **Licence d'Informatique**, *Université Montpellier 2*.

## Compétences

Langages	Python, Java, R, Bash...	Big Data	Hadoop ( <i>Mapr</i> , <i>Hortonworks</i> ), HDFS, YARN, Spark, Hive, Oozie, Hbase, Kafka, Nifi, Airflow
DevOps	Git, Docker, Kubernetes, CI/CD Gitlab	Data Science	Connaissances en Text-Mining et Machine Learning
Base de données	Modèles relationnels (Oracle, SQLServer, PostgreSQL), NoSQL (ElasticSearch, Neo4j, HBase)	Autre	Administration Linux. Architecture des systèmes d'information.

## Autre

