



# Curriculum Vitae

KHIARI Sana

INGENIEUR BIG DATA

# COMPETENCES

## Techniques

- **Big Data** : Hadoop(Cloudera), HDFS, Map Reduce, Redis, Spark, SparkML, Neo4j, Cassandra, MongoDB, ElasticSearch, AWS.
- **Business intelligence**: OLAP, Qlik Sense, Power BI, Score carding, DataViz(KPI).
- **Management des systèmes d'information**: Camunda, Talend, SoapUI, Activiti, MS Project, Esper.
- **Qualité logicielle**: SOLID, Design patterns, Jenkins, Sonar, Maven, TFS, Intégration et déploiement continue, Heroku, Firebase, Gulp, JShint, métriques logicielles, Github, TravisCI, mocha, Selenium, Docker.
- **Base de données**: SQL server 2016, MongoDB, Redis, Oracle (SQL,T-SQL).
- **Langage et framework**: Python, R, Nodejs ( LoopBack, Expressjs ), Angular4, React, Java, Javascript, C# .NET, web api, J2ee, Spring MVC, Spring Boot, C++, Devops, Unix.
- **Machine learning**: Réseau de neurones(Perceptron), Arbre de décision (Random forest), K-means, Naïve Bayes, plus proche voisin.
- **Analyse de données**: Analyse en composante principales (PCA), analyse factorielle des correspondances (AFC) avec R.
- **Méthodologies**: Agile (SCRUM).
- **Architectures**: REST, SOA, n-tiers, architecture parallèle (OpenMP).

## Fonctionnelles

- Accompagnement des experts métier dans la conception et leur expression du besoin fonctionnel
- Traduction des règles de gestion fonctionnelle en règles techniques
- Analyse des anomalies des règles métier des données EDF de ressources humaines
- Rédaction des spécifications techniques détaillées des demandes clients
- Réalisation de tests unitaires et d'intégration sur l'environnement de recette et production et la rédaction des fiches de tests

## Expériences professionnelles

### Novembre 2020 à Aujourd'hui

Contexte	<b>FEATWAY</b> <b>Client EDF</b>
Rôle	<b>Pilote d'un projet Big Data – PFE « Plateforme des Flux d'Echange »</b>
Mission	<ul style="list-style-type: none"> <li>Assurer les échanges avec le client EDF au travers d'instances de suivi du chantier.</li> <li>Réaliser la planification mensuelle des tâches de l'équipe et le suivi des livraisons des membres de l'équipe.</li> <li>Piloter les mises en production pour chaque lot de livraison et le suivi des incidents de production.</li> <li>Créer de nouveaux flux sur le projet intitulé « plateforme flux d'échange ». Le projet est un ensemble de flux sortants d'un Datalake de données RH des salariés d'EDF vers les systèmes d'information et les web services externes à EDF.</li> <li>Créer des tables partitionnés et vues Hive ou Spark.</li> <li>Réaliser des ordonnancements de workflow avec OOZIE.</li> <li>Réaliser des scripts Shell de chiffrement des données sortantes du Datalake avec le protocole GPG ou OpenSSL.</li> <li>Réaliser des scripts d'envoi des fichiers chiffrés avec le protocole SFTP.</li> <li>Réaliser des optimisations de code avec Hive/Spark</li> </ul>
Technologies	OOZIE, HIVE, GPG, OpenSSL, SFTP, SPARK

### De Mars 2019 à Octobre 2020

Contexte	<b>FEATWAY</b> <b>Client EDF</b>
Rôle	<b>Consultante Big Data</b>
Mission	<ul style="list-style-type: none"> <li>Concevoir et développer les briques d'un Datalake Big Data. Les solutions développées portent bien sur les données entrantes dans le Datalake ainsi que les données sortantes.</li> <li>Stocker les données partitionnées avec des tables Hive.</li> <li>Nettoyer les données avec Spark.</li> <li>Déchiffrer les données et leurs insertions dans des vues Hive.</li> <li>Transformer les données en table HBASE pour mise à disposition d'un web service EDF.</li> <li>Réaliser des tests unitaires et d'intégration sur l'environnement d'intégration et recette.</li> <li>Mise en production des évolutions de code sur l'environnement de production.</li> <li>Assurer le suivi des coordinateurs lancés en production avec l'interface</li> </ul>

	OOZIE. <ul style="list-style-type: none"> <li>Assurer la gestion devops (intégration continu) avec Jenkins depuis GITLAB sur l'environnement de production, recette et intégration.</li> </ul>
Technologies	Jenkins, Hive, Spark, HBase, Hadoop, Mobaxterm, visual studio code, gitlab, scoop

## De Mars 2019 à Octobre 2020

Contexte	<b>KONVERGENCE Business &amp; technologies</b>
Rôle	<b>Stagiaire Data</b>
Mission	<ul style="list-style-type: none"> <li>Implémenter une API de données avec le protocole oData</li> <li>Exposer les données dans une interface DATAVIZ en utilisant D3js et javascript.</li> <li>Implémenter une API se basant sur une architecture de microservice et les technologies Spring Boot, Spring Cloud, Netflix service de découverte eureka, kafka et redis.</li> <li>Stocker les données dans un conteneur Docker géré par Rancher.</li> <li>Connecter l'API de données aux logiciels de BI comme PowerBI et Tableau pour rendre les données disponibles en temps réel pour les clients. Les données sont de types financiers.</li> </ul>
Technologies	PowerBI, Tableau, Sprint Boot, Sprint Cloud, Kafka, Redis et Docker

## Juillet 2017 à Aout 2017

Contexte	<b>SMART INNOVATION TECHNOLOGY STARTUP</b>
Rôle	<b>Stagiaire en science de données</b>
Mission	<ul style="list-style-type: none"> <li>Développer un agrégateur et analyse de données.</li> <li>Réaliser l'extraction des offres de travail (Web Scraping ) avec Selenium en Python.</li> <li>Assurer le nettoyage avec Python et standardisation avec l'algorithme de Levenshtein.</li> <li>Gérer le stockage des données dans Redis puis MongoDB.</li> <li>Réaliser l'exploration des données avec R</li> <li>Réaliser un moteur de traitement et la REST API avec loopback en nodejs.</li> <li>Réaliser une interface Dataviz avec des graphes avec les statistiques des tendances technologiques avec Angular 4 et D3js.</li> </ul>
Technologies	Python, MongoDB, Nodejs, Angular 4 et D3js

## Diplômes et Formations

Année	Libellé
De 2013 à 2018	Ingénieur en Génie Logiciel Spécialité management des systèmes d'information – Option Big data