

Merci de compléter l'ensemble des champs afin que la fiche de poste puisse être validée

Date de la demande : 05/12/2020

Date d'entrée souhaitée : 02/03/2020

Flexibilité :  1 mois  2 mois  > 3 mois

Durée du stage : 6 mois

Nombre de poste à pourvoir : 1

### INFORMATIONS GENERALES

Entité : Crédit Agricole SA

Direction d'accueil (*pas de sigle*) : Marketing et Communication

Pôle (*sigle*) : DCI

Maître de stage : Hugo Hamad – hugo.hamad@credit-agricole-sa.fr

Type de métier : *Marketing et Communication*

Type de métier complémentaire si nécessaire : Choisissez un élément.

Lieu de travail : Montrouge

Y avait-il un autre stagiaire précédemment sur les mêmes missions ? Si oui, précisez le nom et la période de présence ? Non

### DESCRIPTION DE LA MISSION

Intitulé du poste : Stagiaire Data & AI Engineer – DataOps (DevOps)

Présentation du service :

**Crédit Agricole, Agir chaque jour dans l'intérêt de nos clients et de la société**

Nous pensons que la Data et l'IA doivent avant tout être au service de l'humain.

Au service de nos clients tout d'abord, en leur offrant un service personnalisé dans les moments qui comptent.

Au service de nos collaborateurs bien sûr, en leur permettant de gagner du temps pour être davantage disponibles pour nos clients.

Nos principales missions :

- Concevoir en partenariat avec les entités du Groupe Crédit Agricole des solutions d'IA internes et innovantes,
- En accélérer l'industrialisation en étroite coopération avec nos producteurs informatiques,
- Animer le partage d'expérience au sein de notre vaste communauté d'experts Data pour favoriser une montée en compétences collective.

Pour mener à bien ses missions, le DataLab dispose de Data Scientists, d'infolinguistes, de Data & AI Engineers et de Chefs de Projet expérimentés qui travaillent avec ses partenaires au sein de squads pluridisciplinaires.

Dans le cadre de ses missions, des thématiques scientifiques à forte valeur ajoutée sont étudiées : Apprentissage Automatique, Auto-ML, Traitement du Langage Naturel, Process Mining, Time Series Mining, Deep Learning, Géomatique, etc. Ces activités sont menées conjointement avec des partenaires internes : les Caisses Régionales, les Entités du Groupe et les Producteurs Informatiques. Le DataLab développe également un réseau de partenaires externes lors de missions industrielles (Editeurs de logiciels, startup, SSII, etc.) ou de collaborations universitaires.

**Dans le cadre de ce stage, vous rejoindrez l'équipe Data & AI Engineering du DataLab afin de contribuer à la valorisation de la Data et participer à la mise en place d'approches innovantes au service des Clients du Groupe.**

### **Descriptif de la mission : Améliorer et optimiser le cycle de vie des projets de DataScience**

#### Contexte et objectifs du stage :

Une des missions principales de l'équipe Data & AI Engineering au DataLab est d'accélérer l'industrialisation des solutions IA développées en interne.

Pour réaliser cette mission, l'équipe met en place des pratiques de développement, d'automatisation et d'agilité et cherche à répondre aux challenges qui découlent des spécificités des projets Data Analytics (Gérer le code et les données sur toute la chaîne analytique, automatiser la création des environnements et assurer la collaboration entre les différents rôles/compétences).

L'objectif de ce stage est de mettre en place des pratiques DataOps pour optimiser et automatiser la chaîne de traitement d'un projet Data.

Le DataOps est un cadre qui s'inspire du monde du DevOps mais la génération et la livraison de code n'est qu'une partie.

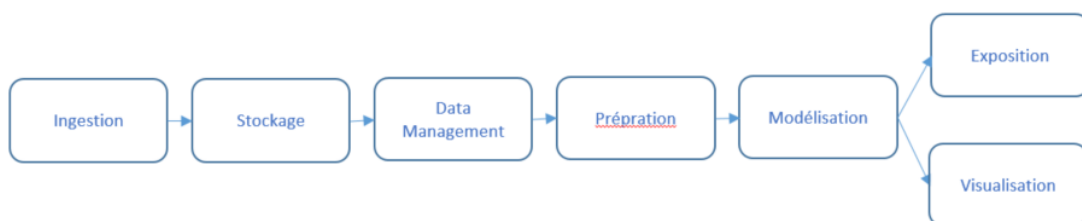
**DataOps** cherche à réduire le temps de cycle de bout en bout de l'analyse de données. Le cycle de vie des données dépend des personnes en plus des outils. Pour que DataOps soit efficace, il doit gérer la collaboration et l'innovation. À cette fin, **DataOps introduit le développement agile dans l'analyse des données afin que les équipes de données et les utilisateurs travaillent ensemble plus efficacement et plus efficacement.**

Parmi les leviers nous retrouvons les pratiques suivantes :

- Automatisation (Intégration continue / Déploiement continu)
- Tests unitaires
- Gestion d'environnements
- Gestion de version
- Monitoring

Donc, plusieurs étapes se succèdent pour automatiser la conception, le déploiement et la gestion des flux de livraison de données et des modèles.

### Pipeline Data



Ce pipeline est orchestré/automatisé pour garantir sa bonne mise à l'échelle jusqu'en production, grâce à diverses pratiques de CI/CD.

Le stagiaire prendra en charge la mise en place des pratiques DataOps en utilisant les technologies OpenSource de l'état de l'art (MLFlow, KubFlow, Kubernetes, Kafka, ElasticSearch, Kibana, etc.).

## PROFIL RECHERCHE

Niveau d'études préparé :

Bac + 2    Bac + 3    Bac +4    Bac +5

Formation :    Université    Ecole de commerce    Ecole d'ingénieur    Autre précisez :

Ecoles ou formations dans lesquelles vous souhaitez éventuellement diffuser l'annonce :

X, IMT, EPITA, ENS, Ecole des Mines, Telecom Paris, Ecole Centrale, INSA, UPMC,42, Université Paris Dauphine, Université de Lyon

Spécialisation :

Big Data, Système d'information et Génie Logiciel, Data Engineer, Intelligence Artificielle

Compétences techniques ou spécifiques au poste:

- Architecture logicielle
- Développement Python, Java
- CI/CD
- DevOps/DataOps
- La conception des bases de données

Compétences générales et transverses :

- Autonomie et communication
- Rigueur
- Innovation/R&D et force de proposition
- Travail en équipe

Outils informatiques :

Python, Flask, MongoDB, MLFlow, KubFlow, Kubernetes, Kafka, ElasticSearch, Kibana, NGinx, Celery, Redis, ReactJS, Docker, etc.)

Langues 1 : Français

Niveau requis : *Opérationnel (usage régulier)*

## ASPECTS BUDGETAIRES

### **PARTIE A COMPLETER OBLIGATOIREMENT**

CGB de facturation :

Unité d'affectation opérationnelle (UO) :

Votre correspondant budgétaire

Avez-vous reçu l'accord de votre correspondant budgétaire ?  OUI    NON

Signature et nom du Manager :

Signature et nom du correspondant budgétaire :