

**Lien pour postuler :** <https://www.groupecreditagricole.jobs/fr/nos-offres-emploi/2020-52853/idOrigine/502>

Date de la demande : 01/11/2020

Date d'entrée souhaitée : 01/03/2021

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*) : DataLab Groupe

Pôle (*sigle*) : ITD

Maître de stage : Aymen SHABOU ([aymen.shabou@credit-agricole-sa.fr](mailto:aymen.shabou@credit-agricole-sa.fr))

Type de métier : *Autres*

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 ?

### DESCRIPTION DE LA MISSION

**Intitulé du poste :** Assistant Data Scientist – Génération de documents artificiels

**Présentation du service :**

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

Au sein du pôle Informatique et Transformation Digitale, les missions du DataLab Groupe visent à maximiser la contribution de la Data et de l'IA au développement et au fonctionnement du Crédit Agricole :

- Concevoir en partenariat avec les entités du Groupe Crédit Agricole des solutions d'IA internes innovantes,
- En accélérer l'industrialisation et la diffusion en étroite coopération avec les métiers et nos producteurs informatiques,
- Animer notre vaste communauté Data pour favoriser la montée en compétences collective.

Pour cela, il dispose d'une équipe de Data Scientists, d'infolinguistes, de Data & AI Engineers et de Chefs de Projet expérimentés travaillant avec ses partenaires internes et externes en Squads pluridisciplinaires dont les réalisations s'appuient sur des technologies à forte valeur ajoutée : Machine Learning, Auto-ML, Natural Language Processing (NLP), Process Mining, Time Series Mining, Deep Learning...

**Dans le cadre de ce stage, vous rejoindrez l'équipe Data Science.**

## Descriptif de la mission :

### Contexte et objectifs du stage :

La digitalisation des services bancaires accroît les volumes de documents numérisés à traiter par les équipes du Crédit Agricole. Il est nécessaire d'améliorer leur analyse pour des fins d'archivage, de traitement automatisé et d'audit. Plus particulièrement, l'archivage de ces documents en vue de les exploiter régulièrement pour entraîner en continue des modèles d'intelligence artificielle ou pour alimenter de nouveaux cas d'usages, se heurte souvent à la réglementation qui limite la durée de stockage des données à caractères personnel qu'ils contiennent.

L'objectif du stage est de développer des IA de génération de documents artificiels qui préservent au maximum les aspects visuels d'un document et certaines zones d'intérêt du texte. Cette IA multimodale (visuelle/textuelle) reposera sur les récentes avancées des réseaux adverses génératifs (GANs) qui ont montré des résultats très encourageants.

Elle servira en interne pour :

- Archiver des documents anonymisés « by design » pour l'amélioration continue des modèles d'IA en production et la réalisation des tests de non régression tout au long du cycle de vie d'un modèle ;
- Augmenter les bases internes peu volumineuses.

### Organisation et livrables :

Pour ce faire, le stage se déroulera en plusieurs étapes, sous l'encadrement de Data Scientists :

- Réalisation d'un état de l'art sur les réseaux adverses génératifs appliqués à l'image et au texte ;
- Sélection et implémentation des algorithmes les plus adaptés à la problématique ;
- Intégration des développements dans le produit interne du DataLab ;
- Publication scientifique si les travaux aboutissent à de nouvelles approches plus performantes que l'état de l'art.

Des interactions fréquentes avec les équipes de Data & AI Engineers du DataLab et des experts métier du Groupe auront lieu.

## 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, MVA, ENSTA, Mines, Telecom, Ecole Centrale-Supelec, ENSAI, ENSAE, Sorbonne Université

Spécialisation :

Formation comportant une dominante vision par ordinateur, NLP, Machine & Deep Learning, développement logiciel en python.

Compétences techniques ou spécifiques au poste (5 maximum) :

- Machine & Deep Learning
- Vision par ordinateur
- NLP
- python

Compétences générales et transverses (5 maximum) :

- Autonomie, rigueur et communication
- Force de proposition
- Travail en équipe

**Outils informatiques :**

- Keras, Tensorflow, opencv, pytorch

**Langues 1 :** Français

*(Si aucune langue souhaitée indiquer N/A)*

**Niveau requis :** *Opérationnel (usage régulier)*

**Langues 2 :** Anglais

*(Si aucune langue souhaitée indiquer N/A)*

**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 :