



RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité



Durée et période : La durée de l'intervention est au minimum 1 jour par semaine sur 2 mois ou deux fois une semaine qui peuvent être non consécutives.

Lieu de la mission : La mission se fait à distance ou en télétravail.



CONTEXTE

Le LABO BI et Big Data a été associé au projet du ministère de la santé plus précisément la cellule de crise du COVID pour mettre en place la plate-forme interministérielle destinée aux professionnels de santé et aux acteurs de la gestion de

la crise : <https://covid-19.sante.gouv.fr>

La plateforme comporte 4 modules. Il s'agit pour cette mission de collaborer au module de veille scientifique pour faciliter l'accès aux publications scientifiques (Waked-co).

COVID 19 : Faciliter la veille scientifique pour les chercheurs et le corps médical

Une mission proposée par :

Ministère des Armées/ Délégation de la transformation et de la performance ministérielles (DTPM)



LA MISSION

En fonction des compétences et appétences du talent, une des missions suivantes peut lui être confiées notamment en matière de construction de modèle NLP :

- réalisation d'un algorithme permettant d'extraire des informations dans le corpus documentaire en l'occurrence les publications scientifiques et les essais cliniques
- réalisation d'un algorithme permettant d'opérer des traitements (indexation) sur du texte des publications scientifiques afin d'y retrouver plus facilement des mots et d'augmenter la pertinence des recherches
- réalisation d'un algorithme pour la recherche d'objets ou d'images dans un texte
- réalisation d'un algorithme de traduction des publications (anglais/français, chinois/anglais,...)
- réalisation d'un algorithme pour ingérer de nouvelles sources de données. Cela peut être l'utilisation d'une API, de flux json ou autre méthode respectant les conditions générales d'utilisation (CGU) de l'éditeur des publications :

- Recherche de nouvelles sources intéressantes
- Étude des CGU pour utilisation de leurs données
- Étude des outils de récupération des données à disposition (API, flux, web scrapping...)
- Ingestion des données
- Gestion des erreurs et anomalies de traitement

Tous les codes développés et méthodes utilisés seront décrits et déposés dans le Gitlab du projet.



PROFIL ET EXPERTISES RECHERCHÉS :

Le talent devra détenir les compétences suivantes :

- Data scientist ou Data engineer
- Python : web scrapping, utilisation d'API
- Connaissance de GitLab. Option : Connaissance de la stack ELK (elastic)



ENVIRONNEMENT DE TRAVAIL

Équipe mixte constituée d'agents de la cellule de crise du COVID (ministère de la Santé) et du labo BI et Big Data du ministère des Armées partageant un Gitlab public.

Cette mission vous intéresse ? **Proposez vos compétences !**

→ bit.ly/talents-numeriques

**Partagez
vos talents
numériques**

Un dispositif piloté par
la direction interministérielle
du numérique (DINUM)