

## Health Data Hub

9 Rue Georges Pitard  
75015 Paris

# Stagiaire data engineer

## Equipe data

**Contrat – stage**

**Rémunération – à définir**

**Date de début – entre janvier et mars 2021**

## LE HEALTH DATA HUB

Le Health Data Hub (HDH) ou « plateforme des données de santé » est un Groupement d'Intérêt Public dont l'objectif est de faciliter les travaux de recherche sur le patrimoine des données de santé. Il interagit avec de nombreux acteurs tels que les Ministères, la Caisse Nationale d'Assurance Maladie, les hôpitaux. La transparence et la participation des citoyens sont des engagements clés du HDH.

Le Health Data Hub c'est :

- Une plateforme technologique à l'état de l'art disponible pour héberger les premiers projets de recherche
- Une équipe pluridisciplinaire de 37 personnes à haut potentiel constituée en moins d'un an, au service de l'intérêt général
- Près de 1 pétaoctet de données de santé traitées par an lorsque le HDH aura atteint son régime de croisière
- 28 projets pilotes accompagnés - Exemples :
  - Analyse du non recours au soin pendant le confinement à partir des résumés de passages aux urgences
  - Etude de l'impact et des traitements sur des données de vie réelle afin d'identifier les meilleurs schémas thérapeutiques pour le traitement du Sarcom
  - Evaluation de l'apport de l'IA dans le dépistage organisé du cancer du sein

Le HDH porte une mission de facilitation de la recherche médicale, et de ce fait est un vecteur d'amélioration de la vie des patients. Au coeur des enjeux du numérique et de la santé, il est porté par une équipe dynamique qui contribue au partage de la connaissance à travers l'open source.

Plus d'information sur le Health Data Hub et son équipe sur notre [site](#).

## DESCRIPTIF DU STAGE

Le stage se déroulera au sein de l'équipe data. Le Health Data Hub est un des lauréats de l'appel à projet organisé par EHDEN, le réseau européen des données de santé.

A travers ce nouveau partenariat, le Health Data Hub transforme des données médicales du SNDS (Système National des Données de Santé) vers un format qui a vocation à devenir une référence internationale : le format OMOP-CDM (Observational medical outcomes partnership - Common Data Model). La conversion des bases de données de santé vers le format OMOP-CDM renforcera leur interopérabilité et facilitera le partage d'outils, de méthodes, d'études et de programmes.

Le SNDS historique est une base de données médico-administrative comportant principalement les remboursements de soins de ville et les hospitalisations pour la quasi-totalité de la population française. De par son système de sécurité sociale, la France fait partie des quelques pays qui disposent de systèmes d'informations médico-administratives couvrant l'ensemble du territoire et de la population.

Aujourd'hui le SNDS, c'est plus de 180 tables, plus de 4500 variables, de nombreuses règles métiers, une évolution constante, 1.2 milliards de feuilles de soin par an, 11 millions de séjour hospitalier par an. Cela en fait une des plus grosses bases médico-administratives au monde : 450 To au total.

Le stage aura pour objectif d'accompagner l'équipe data dans ce projet de conversion des données SNDS sous le format OMOP-CDM.

Plus particulièrement :

- Ecriture de l'ETL
- Industrialisation de la transformation
- Contrôle qualité des résultats
- Documentation des données et métadonnées spécifiques à ces nouvelles tables
- Contribution aux outils spécifiques de l'équipe data (catalogue de métadonnées, outil de création de données synthétiques, etc.)
- Aide à l'implémentation de standards et normes

## PROFIL RECHERCHÉ

### Compétences techniques

- Maîtrise en informatique et programmation ;
- Maîtrise du langage SQL ;
- Maîtrise du logiciel Python ;
- Connaissance des processus d'ETL et du data modeling
- Connaissance du traitement des données massives et des technologies Big Data (en particulier, framework Spark)
- Capacité à naviguer et appréhender une base de données particulièrement complexe
- Appétence pour le secteur médical

## **Qualités humaines et relationnelles**

- Motivation ;
- Qualités d'organisation, de planification et de rigueur ;
- Sens du service et du travail en équipe ;
- Bon relationnel ;
- Curiosité et capacité d'adaptation ;

## **Durée du stage**

Stage de 6 mois

## **Pour postuler**

Contact : [jobs@health-data-hub.fr](mailto:jobs@health-data-hub.fr)