

Data Engineer H/F

Direction Data du HDH

Contrat – CDI

Rémunération – compétitive

Diplômes – Bac+5/Master

Date de début – dès que possible

Localisation – Paris 15ème (télétravail possible)

LE HEALTH DATA HUB

Comment **améliorer les dépistages** et faire en sorte que les patients soient pris en charge le plus tôt possible ?

Comment leur **proposer les meilleurs traitements** sur le long cours ?

Comment **appuyer les professionnels de santé** dans un contexte clinique qui se complexifie ou en cas de crise sanitaire ?

L'Intelligence Artificielle et les **données de santé** font partie de la réponse. Elles sont incontournables pour la **recherche** et **l'innovation en santé**, par exemple, pour prévenir des insuffisances cardiaques à partir de données issues d'appareils connectés, pour accélérer le dépistage du cancer du sein à partir d'analyses automatiques des examens de mammographies ou même pour réunir assez d'informations afin d'améliorer la prise en charge des maladies rares.

Et pour ça, la France a la chance de disposer de **bases de données extrêmement riches!**

Mais ces données sont souvent sous exploitées car dispersées. Grâce à des **solutions innovantes** telles que l'IA, l'objectif du Health Data Hub est justement de permettre d'accéder de manière **facilitée, unifiée, transparente** et **sécurisée** à un catalogue de bases de données de santé françaises.

Comment ?

Le Health Data Hub a mis en place une **plateforme technologique** qui met à disposition des porteurs de **projets d'intérêt public**, dans un environnement technologique sécurisé et à l'état de l'art, les données de santé **pseudonymisées** des français. Ces porteurs de projets vont mobiliser des sources de données très **volumineuses**, les croiser entre elles, et utiliser une puissance de calcul pour faire tourner des algorithmes de recherche complexes. Il s'agit par exemple de projets de start-up pour améliorer des logiciels d'aide au professionnel de santé, de projets permettant **d'améliorer la prise en charge des patients** en comparant l'efficacité de prise en charge, de projets portés par les administrations pour éclairer les politiques publiques.

Notre offre technologique, en constante évolution, peut être consultée [ici](#).

Les défis sont de taille pour traiter ces données de santé sensibles, volumineuses de natures et formats variables. La plateforme doit être un **levier d'innovation** dans l'écosystème de la donnée de santé français.

En résumé, avec le Hub, nous **accompagnons des porteurs de projets innovants** qui contribuent à trouver les solutions de demain pour améliorer la santé de tous les citoyens.

DESCRIPTIF DU POSTE

Direction des données :

Pour mener à bien les missions qui lui ont été confiées, le Health Data Hub a formé la direction des données dont les objectifs principaux sont de :

- 1) **Définir des stratégies novatrices** sur la gestion, l'exploitation et le partage de données de santé, permettant de réaliser la vision du HDH ;
- 2) **Partager et mutualiser les outils et les connaissances** nécessaires à l'analyse des données de santé, dans le cadre d'une démarche open source.
- 3) **Gérer et mettre à disposition les données** qui lui sont confiées aux porteurs de projet au sein de la plateforme technologique du Health Data Hub ;
- 4) **Soutenir les projets d'intérêt public** que le HDH accompagne, aussi bien sur la compréhension des données de santé que sur leur exploitation via des experts des données de santé, des data scientists et des data engineers.

Les missions :

Accompagner le producteur dans la collecte des données cliniques :

- Soutien des producteurs à la collecte et à l'exploration des données ;
- Structuration des données et constitution de tables en lien avec les informaticiens ;
- Mise en qualité et documentation des données en lien avec le corps médical ;
- Mise en place d'outils permettant l'exploration des données
- Mise en place d'outils spécifiques permettant la gestion des données
- Mise en place d'outils permettant l'enrichissement des données (outils d'annotation, etc.)
- Sélection et anonymisation des données

Construire les flux de données issus des producteurs pour alimenter le « Health Data Hub » :

- Mettre en place des flux de collecte de données (extraction des données issues des logiciels métiers, sélection et ingestion de sources de données, mise en place d'outils d'interfaçage, d'indexation, etc.)
- Réaliser les activités de préparation des données (nettoyage, standardisation, ciblage, anonymisation) ;
- Industrialiser le code de génération du flux de données et assurer sa performance globale ;
- Aide à l'implémentation de standards et normes de mise à disposition des données.

Collecter, analyser et intégrer les retours des producteurs de données ;

Contribuer à l'animation de l'écosystème des producteurs : en poste chez un producteur de données impliqué dans un projet pilote soutenu par le "Health Data Hub", le

Data Engineer mènera une mission clé car de l'existence de jeux données, qui plus est de qualité, dépendront les avancées de la science et de la médecine personnalisée. Tel un ninja, le data engineer viendra soutenir les efforts d'équipes petites ou grandes dans un établissement hospitalier, une clinique, une structure responsable d'une cohorte de recherche ou de l'agrégation de données médicales de diverses natures. Il viendra toutefois de temps en temps travailler dans le "Health Data Hub" sur le développement d'une doctrine commune de mise à disposition des données et pourra partager son expérience avec les autres data engineers et le reste de l'équipe. Il pourra venir appuyer différents producteurs au cours de son contrat.

Dans le cadre de vos fonctions de data Engineer vous serez susceptible de travailler en étroite collaboration avec les partenaires du Health Data Hub et pourrez être amené(e) à travailler dans leurs locaux pour la bonne réalisation de vos missions

PROFIL RECHERCHE

Compétences techniques :

- Grande rigueur, application, goût de l'effort
- Maîtrise en informatique et programmation ;
- Excellentes compétences en codage (Python, SQL, R, etc.) ;
- Maîtrise des outils ETL (Talend, Sqoop...), d'informatique décisionnelle et des méthodes de data warehouse (OLTP, RDBMS...)
- Connaissance des standards d'interopérabilité du domaine de la santé (FHIR, OMOP, CDA, HL7, CIM, Snomed, LOINC...)
- Bonne connaissance du traitement des données massives et des technologies Big Data (Hadoop, Spark, Elastic Search, NoSQL, etc.) ;
- Bonne connaissance et pratique des webservices (SOA, REST) et des technologies d'intégration à l'aide d'interfaces de programmation (API)
- Expérience de travail au sein d'une équipe de développement agile ;
- Appétence pour le secteur médical

Compétences managériales et relationnelles :

- Flexibilité et goût de l'aventure ;
- Excellentes compétences organisationnelles et souci du détail ;
- Aptitude à travailler en équipe et communiquer ;
- Curiosité, goût du changement et capacité d'adaptation ;
- Pédagogie, clarté de présentation.

POURQUOI CHOISIR LE HEALTH DATA HUB?

Vous êtes motivé(e) à rejoindre une équipe impliquée dans un projet ambitieux, qui a du sens et une finalité d'intérêt public ? Rejoignez-nous !

Notre récente structure a besoin de talents créatifs, autonomes et proactifs pour continuer de grandir ! Ensemble, nous nous sommes engagés à :

- Accompagner les porteurs de projet visant à analyser les données de santé pour le bien commun.
- Construire et opérer une plateforme technologique pour leur offrir les meilleurs outils avec un très haut niveau de sécurité à respecter.
- Réunir et mettre en forme les données au plus grand potentiel pour la recherche et l'innovation.
- Promouvoir le partage des connaissances, des expertises et du savoir et diffuser une culture de la donnée de santé auprès de tous.

Bon à savoir:

 Rejoindre le HDH c'est surtout participer à un projet enrichissant humainement qui a du sens, avec un fort impact sociétal

 Au HDH on favorise la prise d'initiative, dans une ambiance de challenge perpétuel

 Ici la bonne humeur et l'esprit d'équipe règnent

Process de recrutement

- Un entretien avec un membre de l'équipe
- Un test à réaliser chez soi et à présenter devant un membre de l'équipe / directeur des données
- Un entretien avec la directrice du Hub
- Un entretien RH

Pour postuler: CV et lettre de motivation obligatoire sur WTTJ [ICI](#)