

Data scientist spécialisé(e) en standardisation des données H/F

Direction des données du HDH

Contrat – CDI, temps plein

Rémunération – selon expérience et profil

Date de début – dès que possible

Localisation - Paris 15ème (télétravail partiel 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.

1)

Vous rejoindrez le pôle Data - "Valorisation et accompagnement des projets", directement rattaché à la directrice, qui a pour missions de concevoir une ou plusieurs offres de service à destination de l'écosystème des producteurs de données et/ou des utilisateurs de données, en lien avec les enjeux de documentation, d'interopérabilité et de qualité des données. Votre pôle sera constitué d'experts en analyse, science des données ou ingénierie des données.

DESCRIPTIF DU POSTE

Mission 1 - Participer à la standardisation des données de santé

Ayant pour mission de garantir un accès aisé et unifié, transparent et sécurisé aux données de santé pour améliorer la qualité des soins et l'accompagnement des patients, le Health Data Hub s'est positionné comme un acteur majeur de leur standardisation sur la scène française et internationale, notamment suite à la standardisation du Système National des Données de Santé sous format OMOP-CDM.

La standardisation s'incarne, entre autres, à travers la définition de :

1. **terminologies communes**, permettant d'assurer une compréhension partagées des notions manipulées ;
2. **structures de données communes**, permettant d'assurer une logique d'analyse stable ;
3. **protocoles d'échange communs**, assurant la simplicité de partage des données entre entités.

Ce positionnement prend forme par la conduite des activités suivantes, auxquelles le / la candidat(e) sera amené(e) à contribuer :

1. **Cartographier et étudier les standards de données existants** de manière à définir la doctrine du Health Data Hub en la matière ;
2. **Standardiser les jeux de données stratégiques**, potentiellement en partenariat avec les responsables des données ou des porteurs de projet, pour simplifier leur compréhension et utilisation ;
3. **Développer un ensemble d'outils** pour automatiser les activités de standardisation ;
4. **Promouvoir les travaux réalisés** auprès de la communauté nationale et internationale à l'aide de rendus ou lors d'événements.

Lors de ces activités, le / la candidat(e) sera notamment mobilisé(e) sur l'aspect technique de ces activités et plus particulièrement sur :

1. **la compréhension des modèles techniques** des standards ;
2. **le développement de pipelines industrielles** (ETL / ELT) pour mettre les données dans les formats des standards ;

3. **la réalisation de tests** pour s'assurer de la qualité des résultats obtenus ;
4. **la conduite d'analyse** pour étudier et valoriser le jeu standardisé de données produit ;
5. **le développement de bibliothèques** permettant d'automatiser la transformation et la vérification des données standardisées ;
6. **la documentation** du code et des données créés.

Mission 2 - Accompagner les partenaires à valoriser les données de santé à travers une mise à disposition

Le Health Data Hub soutient aussi ses partenaires en mettant à disposition des professionnels des données, tels que des experts SNDS, des data scientists ou des data engineers.

En tant que data scientist, vous rejoindrez donc les rangs de l'équipe du porteur de projet dans le cadre d'une mission d'une durée limitée pour réaliser les activités nécessaires à la préparation ou la réalisation de son étude. Ces activités peuvent être des natures suivantes :

1. **Fournir un regard critique ou aider le porteur de projet à définir l'approche algorithmique** nécessaire à la réalisation de son étude ;
2. **Développer les algorithmes** permettant de produire les résultats attendus (e.g., réalisation d'analyses statistiques, développement d'algorithmes de machine learning ou de deep learning) ;
3. **Rédiger des articles ou des publications** sur les travaux réalisés au nom du porteur de projet

Cette liste est non exhaustive ; le collaborateur pourra appuyer sa direction dans d'autres missions.

PROFIL RECHERCHÉ

Compétences techniques

- Expertise en science des données (analyse statistique, optimisation, machine learning, deep learning, reinforcement learning, etc.) ;
- Maîtrise en informatique et programmation ;
- Maîtrise du langage SQL ;
- Maîtrise du langage Python ;
- Connaissance du langage R ;
- Connaissance des processus d'ETL et du data modeling (DBT)
- Connaissance du traitement des données massives et des technologies Big Data ;
- Connaissance de standards dans le domaine de la santé (OMOP-CDM, FHIR, SNOMED-CT, etc.) ;
- Capacité à naviguer et appréhender des bases de données particulièrement complexes ;
- Appétence pour le secteur médical

Compétences relationnelles / personnelles

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

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

PROCESSUS DE RECRUTEMENT

- Un premier entretien avec un membre de l'équipe
- Une mise en situation à réaliser chez soi
- Un entretien avec le manager basé sur la mise en situation
- Un entretien avec la directrice du Health Data Hub
- Une rencontre RH

POSTULER

Retrouvez nous ici sur [Welcome To The Jungle](#) (CV et Lettre de motivation)