

Opérateur données

Direction de la donnée

Contrat – Poste ouvert pour un CDI

Rémunération – Selon expérience et profil

Date de début – dès que possible

LE HEALTH DATA HUB

A la remise du rapport Villani, le Président de la République a affirmé sa volonté de faire de la santé un des secteurs prioritaires pour le développement de l'intelligence artificielle en France. Deux actions majeures ont été annoncées : l'élargissement du système national de données de santé (SNDS) aux données médicales et la création d'un « Health Data Hub » pour faciliter l'accès aux données de santé pour des projets d'intérêt général innovants sous hautes conditions de sécurité et dans le respect des droits des citoyens.

Faisant suite à plus de 100 auditions de l'ensemble de l'écosystème, le rapport de préfiguration, remis à la Ministre le 12 octobre, s'était appuyé sur une analyse des freins existants et des attentes des parties prenantes pour proposer une vision cible et une feuille de route de mise en œuvre pour le « Health Data Hub » autour de quatre grandes missions :

- Soutenir les producteurs dans la collecte et la consolidation du patrimoine national de données de santé ;
- Favoriser un accès simplifié, effectif et accéléré aux données à travers un rôle de « guichet unique » pour les porteurs de projets ;
- Mettre à disposition des porteurs de projets des capacités technologiques sécurisées et humaines à l'état de l'art pour permettre une réutilisation innovante et performante des données ;
- Soutenir l'écosystème et assurer le lien avec les citoyens et la société civile.

18 mois plus tard, la feuille de route ambitieuse a été respectée. Le 24 juillet, la loi relative à l'organisation et la transformation du système de santé a été promulguée. Son article 41 définit la Plateforme des données de santé autrement appelée Health Data Hub. En pratique, celle-ci fut effectivement créée le 30 novembre 2019. D'ici quelques semaines, la plateforme technologique

sera ouverte aux premiers projets pilotes dans la perspective d'une ouverture au plus grand nombre milieu 2020. Un premier catalogue de bases de données est en cours de définition qui permettra une mise à disposition facilitée des données jugées prioritaires aux chercheurs, associations de patients et citoyens, institutions, start-ups et plus généralement, aux différentes parties prenantes du secteur de la santé.

L'équipe du Health Data Hub compte aujourd'hui une vingtaine de personnes passionnées, aux compétences variées et rassemblées autour d'un projet qui fait sens. Face aux ambitions poursuivies, c'est autant d'experts scientifiques, médicaux, ingénieurs, chefs de projets qui sont en ce moment recherchés au sein des quatre directions : technique, données, guichet, fabrique.

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



DESCRIPTIF DU POSTE

En lien avec les producteurs de données, la Direction des Données a la charge de la collecte, de la mise en qualité et de la gestion du catalogue de données. Elle prépare les jeux de données requis pour la réalisation des projets par leurs porteurs, requérant une expertise fine des données pour la création d'extractions pertinentes sur le plan médical. Elle réalise les chaînages lorsque plusieurs sources de données sont mobilisées et doit être capable d'en documenter la qualité. Pour industrialiser ces procédures, elle met en place une gouvernance de la donnée. Elle apporte son soutien aux producteurs dans le processus de préparation des données et notamment la déidentification et la création d'échantillon synthétique. Elle élabore des services pour faciliter la compréhension des données disponibles au catalogue par les utilisateurs en mettant en place des outils de catalogue de métadonnées en ligne et de requêtage simple. Elle est force de proposition dans l'offre de formation et de développement d'outils à visée pédagogique sur le périmètre qui est le sien.

Au sein de la direction des données, l'opérateur données aura la charge, en lien avec les opérateurs de la direction technique.

De préparer les données mises à disposition des porteurs de projets

- De filtrer et préparer, sur la base des données disponibles au sein de la plateforme, les tables requises pour les projets
- De réaliser, le cas échéant, des appariement entre bases
- De développer, le cas échéant, des flux automatiques scalables optimisés
- De vérifier le contenu et la qualité des données utilisées

De contrôler le bon usage des données

- De s'assurer de la sécurité, de la stabilité et de la disponibilité des données de la plateforme
- De réaliser des audits a priori et a posteriori des données utilisées et exportées par les utilisateurs

De remonter les besoins techniques auprès de l'équipe plateforme

- D'apporter une expertise cloud et/ou en traitements distribués dans le cadre de la définition par le directeur technique de l'architecture et de l'implémentation de la plateforme et des services associés
- De remonter les besoins techniques éventuels

PROFIL RECHERCHE

Compétences techniques

- Expérience confirmée sur l'ensemble des techniques de data engineering en environnement cloud, en particulier :
 - gestion des conteneurs : Docker, Kubernetes, etc.,
 - stockage : bases relationnelles, NoSQL, HDFS, etc.,
 - traitement distribué : Hadoop, Spark, etc
- Maîtrise des langages de programmation python, scala
- Maîtrise des standards et bonnes pratiques de code et de développement (incluant versionning)
- Maîtrise des bonnes pratiques du développement agile
- La connaissance des données de santé, registre, entrepôts de données de santé ou SNDS est un plus
- Excellente maîtrise de la gestion de données de formats diverses (images, textes, massives, etc.)

Compétences managériales et relationnelles

- Vous aimez créer et vous avez un esprit entrepreneurial
- Vous prenez du plaisir à produire du code propre.

-
- Vous êtes exigeant et rigoureux et vous excellez dans ce que vous faites
 - Vous avez un excellent relationnel et un très bon sens de la communication
 - Vous aimez travailler en équipe
 - Vous faites preuve d'un leadership naturel et aimez entraîner vos pairs avec vous
 - Vous témoignez de qualité d'organisation, de planification et de rigueur
 - Vous souhaitez mettre vos compétences au service de la santé
 - Vous êtes curieux et vous témoignez de capacité d'adaptation au changement
 - Vous êtes sensible à la diffusion de la connaissance et souhaitez contribuer à l'opensource en partageant vos travaux

Pour postuler

Contact : jobs@health-data-hub.fr