



- ✓ CDD ou CDI à temps plein



#### DIPLOÔME REQUIS :

Bac +5 ou doctorat en Biostatistiques, data science, santé publique ou épidémiologie ou diplôme d'ingénieur dans l'un de ces domaines



- ✓ Capacité de travail en équipe pluridisciplinaire
- ✓ Sens du relationnel
- ✓ Capacité d'adaptation



- ✓ Participation à un projet structurant d'exploitation des données médicales à des fins de recherche
- ✓ Travail en réseau avec les autres CHU impliqués

#### PERSONNES À CONTACTER :

M. Conrad Breuer  
[Conrad.breuer@chu-st-etienne.fr](mailto:Conrad.breuer@chu-st-etienne.fr)

Secrétariat DRI  
[Maggy.CINO@chu-st-etienne.fr](mailto:Maggy.CINO@chu-st-etienne.fr)

# PROFIL DE POSTE

## Data Scientist – Entrepôt de données de santé

### Présentation de l'établissement

Le CHU de Saint-Étienne est situé au cœur de région Auvergne Rhône Alpes, deuxième région de France. Il est l'établissement support du GHT Loire, un des plus grands GHT de France en nombre d'établissements.

Le CHU irrigue son territoire au travers un budget de plus de 730 millions d'euros. Il est le 20<sup>ème</sup> CHU des 32 CHRU de France.

Les activités du CHU sont articulées autour de 11 pôles et 68 services, disposant de 1.920 lits et places, répartis sur ses 3 principaux sites.

Les 8.350 professionnels de ces services prennent en charge annuellement plus de 170.000 patients en hospitalisation et séances et plus de 500.000 en consultations.

### Enjeux du poste

Le CHU de Saint-Etienne, dans le cadre d'un groupement avec les 3 autres CHU de la région Auvergne-Rhône-Alpes, déploie un entrepôt de données de santé qui vise à valoriser les données cliniques, biologiques et médico-administratives issues des systèmes d'information hospitaliers pour soutenir la recherche, l'évaluation des pratiques et l'innovation en santé. La gestion de cette activité est assurée par la Direction de la Recherche et de l'Innovation qui cherche à constituer une équipe pour exploiter cet entrepôt.

### Activité et profil recherché

Dans ce cadre, le Data Scientist contribue à la conception, à la mise en œuvre et à l'exploitation des modèles d'analyse avancée et d'intelligence artificielle appliqués aux données de santé. Le travail est réalisé en étroite collaboration avec les équipes du service d'informatique, les cliniciens et les chercheurs. Le data scientist est membre de la cellule d'ingénierie des données de santé et participe au comité de sélection et d'orientation des projets de recherche sur données de santé.

#### Missions stratégiques :

- Structurer l'exploitation des données de santé au CHU et développer des modèles d'analyse et d'IA pour la recherche clinique et l'épidémiologie
- Contribuer à la structuration régionale des données de santé en lien avec les équipes du GCS HOURAA
- Participer à l'animation du développement d'une politique de recherche sur données de santé structurée et lisible pour les investigateurs du CHU
- Participer au comité de sélection des projets de recherche sur données de santé
- Promouvoir en interne et en externe l'exploitation de l'entrepôt de données de santé

#### Missions opérationnelles :

- Extraction, préparation et analyse des données
- Ingénierie de pipelines analytiques
- Documentation et reproductibilité des analyses
- Interaction avec cliniciens et ingénieurs
- Structuration du lien avec l'éditeur pour faire évoluer les fonctionnalités et les données incluses dans l'entrepôt de données de santé
- Encadrement de plusieurs stagiaires issus des grandes écoles du bassin stéphanois sur la thématique de structuration des données

## Profil recherché

### Formation

- Diplôme d'ingénieur, master 2 ou doctorat en **data science, statistiques, bioinformatique, mathématiques appliquées** ou discipline équivalente

### Compétences techniques

- Maîtrise des langages de programmation : Python, R, SQL.
- Solide expérience en modélisation statistique, machine learning et deep learning.
- Connaissance des environnements de calcul distribués et des bases de données massives (Hadoop, Spark, SQL Server, PostgreSQL, etc.).
- Connaissances des standards de données en santé (OMOP, FHIR, CDISC, SNOMED CT...) appréciées.
- Notions en gouvernance et protection des données de santé (RGPD, hébergement HDS).
- Anglais scientifique et Vocabulaire médical

### Savoir-faire

- Analyser des données en élaborant des programmes informatiques et interpréter les résultats
- Aptitude à la synthèse afin de résumer les données et de rédiger des rapports statistiques
- Capacité à vulgariser les concepts techniques auprès d'interlocuteurs non spécialistes
- Aptitude à structurer son travail
- Aptitude à gérer plusieurs projets en parallèle
- Aptitude au travail en équipe, et à la communication
- Aptitude à s'adapter aux contraintes

## Expérience souhaitée

Expérience préalable souhaitée dans le domaine de la santé numérique, de la recherche clinique ou dans un environnement hospitalier sur des projets de conduite, déploiement et exploitation d'entrepôt de données de santé.

Une expérience de participation à des projets de recherche ou à des publications scientifiques est un atout.

## Conditions d'exercice

- Poste à temps plein, basé au CHU de Saint Etienne (Hôpital Nord)
- Rattachement hiérarchique : Responsable médical de l'EDS / Direction de la recherche et de l'Innovation
- Collaborations régulières avec des chercheurs hospitaliers, des partenaires extérieurs, les ingénieurs informatiques et les juristes de l'établissement
- Possibilité de télétravail selon la politique de l'établissement
- Rémunération selon l'expérience et la grille de rémunération de l'établissement