

## **Data manager pour la médecine génomique (Infrastructure CAD)**

**Employeur :** GIP CAD

**Date de début:** 01/07/2023

**Nature du contrat :** CDI ou CDD suivant le profil

**Rémunération:** entre 60k€ et 85k€

**Lieu de travail:** Paris et télétravail

**Nom et mel du contact:** [contact.pfmg2025@aviesan.fr](mailto:contact.pfmg2025@aviesan.fr)

**Mots-Clés:** Chief Data Officer, Data manager, gouvernance des données, données biologiques, réglementation, cycle de vie données

### **Contexte :**

Le Collecteur Analyseur de Données (CAD) est un Groupement d'Intérêt Public (GIP) récemment constitué dans le cadre du Plan France Médecine Génomique 2025 (PFMG2025 - <https://pfmg2025.aviesan.fr/>). L'objectif principal du CAD est de construire une infrastructure numérique sécurisée nationale ayant pour vocation d'héberger et de mettre à disposition de la communauté biomédicale les données de génomique humaine produites par les plates-formes de séquençage (Auragen et SeqOIA) et provenant des projets pilotes du PFMG2025 à des fins de soin et de recherche. Parmi ses missions, le CAD a pour but de développer une offre de services visant à permettre la réutilisation des données de génomique à des fins de recherche en médecine génomique (maladies rares & cancer). Cela se traduit par la mise en œuvre d'une infrastructure de stockage et de calcul conforme à la réglementation en vigueur (HDS) et par la construction et le déploiement d'environnements logiciels dédiés à l'analyse bioinformatique des données génomiques et cliniques. Le CAD souhaite définir sa politique de gouvernance opérationnelle des données ainsi que les processus de collecte et de la mise à disposition de ces données pour les projets de recherche et de soin. Dans ce contexte, le CAD souhaite recruter un Data Manager pour définir et organiser la stratégie de gestion des données du PFMG2025.

### **Missions :**

En lien avec les équipes bioinformatique, informatique et juridique du CAD, vous aurez pour missions de :

- Piloter la gouvernance opérationnelle des données

- Définir et organiser la stratégie de collecte et la mise à disposition des données issues du PFMG2025
- Mettre en place les méthodes permettant de garantir la qualité, la cohérence et la fiabilité des données
- Assurer la maîtrise du cycle de vie de la donnée et l'ensemble des opérations nécessaires sur les données (intégrité des données, stockage, sécurisation, effacement, réutilisation)
- Assurer et contrôler les opérations de sécurisation et la conformité de ces données avec les contraintes réglementaires (anonymisation et chiffréments)
- Participer à l'évaluation de la faisabilité technique entre les porteurs de projets scientifiques et les équipes techniques du CAD
- Assurer le suivi et la planification opérationnelle de la mise à disposition des données pour les projets de recherche soumis au CAD
- Auditer les processus de sécurisation des données du point de vue technique et organisationnel
- Contrôler le bon usage de la donnée au vu des différentes réglementations
- Contribuer à l'amélioration continue de l'offre de service logiciel du CAD
- Assurer une veille permanente sur les évolutions technologiques et l'état de l'art dans ce domaine
- Participer aux projets nationaux et européens impliquant le CAD

Pendant la phase initiale de lancement du CAD, vous aurez pour mission complémentaire de :

- Participer au cycle des projets de recherche opérés au CAD (identification du besoin, soumission et suivi du projet) au sein de la cellule d'accompagnement du CAD
- Assurer le suivi et la planification opérationnelle des projets de recherche soumis au CAD
- Etablir le cahier des charges des besoins logiciels et informatique avec les porteurs de projets
- Evaluer la faisabilité technique entre les porteurs de projets scientifiques et les équipes techniques du CAD
- Organiser la mise en œuvre et le suivi des déploiements logiciels avec les équipes techniques du CAD
- Assurer le reporting et présenter l'avancement des travaux aux parties prenantes de façon régulière
- Contribuer à l'amélioration continue de l'offre de service logiciel du CAD

#### **Compétences requises :**

- Formations et/ou qualifications : école d'ingénieur ou formation universitaire équivalente (Master, PhD) en bioinformatique et/ou informatique
- Maîtrise des réglementations sur la protection des données (RGPD) et des contraintes réglementaires spécifiques
- Très bonne connaissance des données et formats de données issus du séquençage haut-débit et des méthodes d'analyse bioinformatique en génomique
- Connaissance en programmation informatique (par ex : Python, PHP, R, Golang, shell-scripting)
- Expertise en bases de données (SGBD relationnel ou NoSQL) et langages orientés BDD (SQL)
- Compétences démontrées pour la gestion et l'analyse de données massives
- Compétences techniques sur les technologies de stockage objets (S3, CephFS, ...)
- Expérience validée en gestion de projet de recherche

- Connaissance dans l'administration, la modélisation et la sécurisation des données
- Maîtrise de l'anglais (écrit et oral)
- Connaissance dans les technologies de développement web (création de tableaux de bords) appréciée
- Connaissances en physiopathologie humaine (maladies rares et cancer) appréciées
- Expérience en encadrement appréciée

**Qualités requises :**

- Être capable de dialoguer avec les équipes informatique, bioinformatique et juridique pour conceptualiser et traduire les besoins en solutions techniques
- Sens de l'écoute et qualités de communication
- Sens du travail en équipe
- Aisance relationnelle
- Qualités organisationnelles
- Sens des priorités, capacités de synthèse
- Esprit d'initiative
- Force de proposition et de conviction
- Motivation pour la recherche en médecine génomique