

Nous recherchons en présentiel un scientifique de données sénior avec des compétences démontrées en gestion de données pour aider à développer et à maintenir un référentiel de données de soins de santé institutionnel. L'objectif principal du référentiel de données institutionnel est d'intégrer les sources de données nécessaires à la recherche et d'améliorer les soins aux patients. Notre équipe intègre, organise et analyse les données des patients traités dans le réseau de l'Hôpital général juif. Le candidat travaillera avec l'équipe existante pour développer et exploiter cette infrastructure unique et aura l'opportunité de contribuer à l'amélioration de la santé de centaines de milliers de patients. L'intégration de ces différentes données issues des systèmes sources et leur analyse nécessite une grande expérience des différents formats de données hétérogènes (Oracle, PostgreSQL, XML, HL7, PDF, images, DICOM, etc.). La principale responsabilité de ce poste comprend l'analyse, la configuration, la construction et l'intégration de pipelines de données pour fournir des données au référentiel. Ce poste nécessite une collaboration avec les équipes de projet dans toutes les phases du cycle de vie du développement. Le ou la candidat(e) retenu devra remplir les fonctions suivantes :

Analyse

- Comprendre le paysage actuel de la santé pour obtenir des informations pertinentes, des modèles de données
- Collaborer à la définition et à la cartographie entre les sources de données et les normes de données FHIR associées
- Architecture Kafka, connaissance approfondie du fonctionnement des composants Kafka tels que les brokers, les topics, les partitions, les groupes de consommateurs et les facteurs de réplication.

Conception

- Concevoir, implémenter et maintenir des pipelines de données en flux continu utilisant Kafka, des API internes et externes, ainsi que des consommateurs de données en aval.
- Apporter un soutien aux équipes d'infrastructure internes pour le déploiement et la supervision.
- Recommander des solutions et proposer des architectures et des solutions logicielles.

Développement

- Surveiller et modifier les clusters Kafka, les sujets, les producteurs et les consommateurs afin d'en optimiser les performances et la fiabilité.
- Contribuer au développement des processus, procédures et modèles d'architecture, et apporter son soutien.
- Configurer, intégrer et développer les bibliothèques de code relatives aux processus ETL et aux pipelines de données FHIR.
- Se tenir informé des technologies émergentes et contribuer à l'amélioration de l'architecture.

Assurance qualité

- Gérer les problèmes d'interopérabilité et s'assurer que la solution fonctionne comme prévu
- Identifier, enregistrer, résoudre ou escalader les problèmes soulevés avec les membres de l'équipe

Mise en œuvre

- Surveiller les activités et les performances du pipeline de données
- Fournir un soutien lors de la mise en œuvre
- Fournir un support de programmation, un réglage des performances et une documentation technique

Qualifications/compétences

- Un minimum de 10 ans d'expérience dans le domaine avec l'échelonnement technologique spécifié.
- Maîtrise du traitement en temps réel et de la mise en file d'attente des messages. Request : Kafka, MirrorMaker, Apache Camel
- Compétences démontrées dans la conception, l'intégration, la mise en œuvre et la gestion de systèmes de bases de données. Requis : SQL, MS SQL ; PostgreSQL
- Expérience de travail avec de grandes bases de données, des entrepôts de données, des lacs de données ou d'autres domaines jugés équivalents
- Compétences démontrées avec les scripts et d'autres langages de programmation. Requis : Python, Java, JavaScript, Dagster
- Connaissance des systèmes distribués et des défis liés aux données à grande échelle
- Bonne connaissance du déploiement/développement basé sur des conteneurs. Requis : Azure, Docker
- Systèmes d'interface de communication entre systèmes de différents formats non compatibles. Requis : APIs utilisant REST, XML, JSON, les standards de FHIR seraient un atout.
- Expérience de travail avec les systèmes de collaboration et la gestion des versions (Confluence, JIRA, Bitbucket)
- Connaissance des normes de messagerie HL7, des terminologies et des classifications (telles que SNOMED-CT, Loinc, ICD), la familiarité avec un outil d'intégration est un atout
- Expérience dans le domaine de l'informatique de la santé, de la recherche et des systèmes de santé (SmileCDR, PACS, DICOM, etc.) est un atout

Autres compétences pertinentes :

- Forte curiosité intellectuelle et volonté d'apprendre de nouvelles méthodes et technologies.
- Capacité à équilibrer autonomie et travail d'équipe dans un contexte multidisciplinaire.
- Excellentes compétences interpersonnelles et de communication

Types d'emploi : temps plein

Salaire : 80 000 \$ à 120 000 \$ par année

Avantages :

- Pension d'entreprise
- Soins dentaires
- Assurance invalidité
- Programme d'aide aux employés
- Assurance-maladie complémentaire
- Assurance-vie
- Garderie sur place