We seek a senior data scientist with demonstrated data management skills to help develop and maintain an institutional healthcare data repository. The primary goal of the institutional data repository is to integrate the necessary data sources for research and to improve patient care. Our team integrates, organizes, and analyses the data of patients treated in the Jewish General Hospital network. The candidate will work with the existing team to develop and operate this unique infrastructure and will have the opportunity to contribute to improving the health of hundreds of thousands of patients. Integrating these various data from source systems and their analysis requires extensive experience with different heterogeneous data formats (Oracle, PostgreSQL, XML, HL7, PDF, images, DICOM, etc.). This position's primary responsibility includes analyzing, configuring, building and integrating data pipelines to provide data to the repository. This position requires collaboration with the project teams in all development lifecycle phases. The successful applicant will be required to perform the following functions:

Analysis

- Understand the current health landscape for relevant information and data models
- Collaborate on the definition and mapping between data sources and related FHIR data standards
- Kafka architecture, thorough knowledge of how Kafka components such as brokers, topics, partitions, consumer groups, and replication factors work together.

Design

- Design, implement, and maintain streaming data pipelines using Kafka, internal and external API and downstream data consumers.
- Support internal infrastructure teams on deployment and monitoring.
- Recommend solutions and propose architectures and solutions

Development

- Monitor and troubleshoot Kafka clusters, topics, producers, and consumers for performance and reliability.
- Contribute to the development of architecture processes, procedures, templates etc. and assistance
- Configuration, integration and development of code libraries pertaining to ETL process and FHIR data pipelines
- Stay current with emerging technologies and contribute to architectural improvements.

Quality Assurance

- Handle interoperability problems and ensure the solution functions as expected
- Identify, record, resolve or escalate issues as raised with team members

Implementation

- Monitor data pipeline activities and performance
- Provide support during the implementation
- Provide programming support, performance tuning and technical documentation

Qualifications/Skills

- A minimum of 10 years of experience in the field with the specified technology stack.
- Proficient in real-time processing and message queuing **Required:** Kafka, MirrorMaker, Apache Camel
- Demonstrated skills in database system design, integration, implementation, and management. **Required:** SQL, MS SQL, PostgreSQL
- Hands-on experience working with large databases, data warehouses, data lakes
- Demonstrated skills with scripted, programming languages, and orchestration frameworks **Required:** Python Java, JavaScript, Dagster
- Knowledge of distributed systems and large-scale data challenges
- Familiarity with cloud platforms and container-based deployment/development (such as Azure, Docker).
- Communication interface systems between systems with various formats that are not compatible. **Required:** APIs using REST, XML, JSON, familiarity with FHIR standard is an asset.
- Experience working with collaboration systems and version management (Confluence, JIRA, Bitbucket)
- Knowledge of HL7 messaging standards, terminologies and classifications (such as SNOMED-CT, Loinc, ICD), and familiarity with an integration tool is an asset
- Experience within health IT, research and health systems (SmileCDR, PACS, DICOM, etc.) is an asset

Other relevant skills:

- Strong intellectual curiosity and willingness to learn new methods and technologies stacks.
- Ability to balance autonomy and teamwork in a multidisciplinary context.
- Strong problem-solving skills and collaborative mindset.
- Excellent interpersonal and communication skills

Job Type: Full-time, Permanent

Salary: $80,000.00-$120,000.00 per year

Benefits:

- Company pension
- Dental care
- Disability insurance
- Employee assistance program
- Extended health care
- Life insurance
- On-site childcare

Schedule:

- 8-hour shift