Data Curation Analyst
Job Description

Overview:

We are seeking a detail-oriented and analytical Data Curation Analyst to join our team. As a Data Curation Analyst, you will play a crucial role in ensuring the accuracy, completeness, and usability of our organization's data assets. You will be responsible for curating and managing data collections, performing quality assessments, and implementing best practices for data organization and documentation.

Responsibilities

Data Collection and Aggregation

- Identify, collect, and aggregate data from internal and external sources.
- Collaborate with data owners and stakeholders to understand data requirements and priorities.
- Develop processes and workflows for efficient data acquisition and ingestion.

Data Quality Assessment:

- Perform data quality assessments to identify errors, inconsistencies, and anomalies.
- Implement data validation checks and quality control measures to ensure data accuracy and reliability.
- Work with data producers to address data quality issues and improve data integrity.

Data Standardization and Enrichment

- Standardize data formats, naming conventions, and coding schemes for consistency.
- Enrich datasets with additional metadata, annotations, and contextual information.
- Enhance data documentation and descriptions to improve discoverability and usability.

Data Cleansing and Transformation

- Cleanse and preprocess data to remove duplicates, outliers, and irrelevant information.
- Transform data into standardized formats and structures for analysis and integration.
- Implement data enrichment techniques such as imputation, normalization, and feature engineering.

Data Cataloging and Metadata Management

- Maintain data catalogs, metadata repositories, and documentation libraries.
- Capture and document metadata attributes, including data lineage, provenance, and usage.
- Facilitate metadata discovery and search capabilities for end users and data consumers.

Data Governance and Compliance

- Ensure compliance with data governance policies, regulations, and best practices.
- Implement access controls, permissions, and security measures to protect sensitive data.

- Monitor data usage and access patterns to enforce governance policies and controls.

Collaboration and Communication

- Collaborate with cross-functional teams, including data engineers, analysts, and scientists.
- Communicate effectively with stakeholders to understand data requirements and priorities.
- Provide training and support to users on data curation best practices and tools.

Requirements:

- Bachelor's degree in Information Science, Library Science, Computer Science, or related field.
- Proven experience in data curation, data management, or information management roles.
- Strong understanding of data quality principles, standards, and methodologies.
- Proficiency in data manipulation tools and technologies (e.g., SQL, Python, R, etc.).
- Familiarity with data modeling concepts, metadata standards, and taxonomies.
- Excellent organizational, analytical, and problem-solving skills.
- Ability to work independently and collaboratively in a fast-paced environment.
- Attention to detail and commitment to data accuracy and integrity.