

TCS AI Fridays Hackathon – Season 2

Title:

Data Engineering Metadata Catalog Description Enhancer

Problem Statement:

Metadata catalogs often contain sparse or uninformative dataset descriptions, complicating data discovery and governance. Enhancing descriptions manually is labor-intensive and inconsistent. A generative AI agent that automatically enriches metadata descriptions by synthesizing data schema, sample data, and usage context can improve catalog usability and data literacy.

Data Requirements:

Dataset schemas, sample records, and lineage metadata from data catalogs. Usage logs and query patterns. Existing metadata descriptions for reference. Domain ontologies and glossaries.

(Feel free to generate synthetic data as well from the provided access)

Expected Output:

The deliverable is an automated metadata description generator producing enriched text for catalog entries. Outputs include structured descriptions with sections for content, quality, and usage tips. The agent provides description exports in JSON or XML for catalog ingestion.

Be creative to go beyond as you solve the above needs