State the current issue with the data

Why do we need data cleaning

Global schema, structure of data source   
Any problem with conflict in data, missing data

Mapping data to the local schema to the global schema , what is the problem

data cleaning đầu vào là gì , đầu ra là gì

Xem xét trùng nhau của 2 tập dữ liệu tên của công việc giữa 2 tập hợp công việc là không giống nhau

nếu có 1 danh sách các thuộc tính cần chuẩn hóa thì làm thế nào, chuẩn hóa từ rupee sang usd

Cần nêu vấn đề cần giải quyết

vấn đề gặp phải là gì , cách giải quyết

dữ liêu như thế nào, có đặc điểm gì

là điều cần thiết để đi đến cách xử lý

phân tích nguồn dữ liệu cần tích hợp -> vấn đề cần giải quyết là gì

cần làm rõ flow chart cần cso sự không đồng nhất

để có thể dọp dẹp dữ liệu , chỉnh sửa cấu trúc ,

lý do về việc chuyển đối dữ liệu

cần nêu tường minh vấn đề tích hợp , nguyên nhân cần giải quyết cuả bài toán tích hợp

tại sao chọn quá trình xử lý

và nhận được cái gì tại sao tôi làm cái này, cách thức tôi làm có hiệu quả hay không

tôi nhận được cái gì

phân tích cấu trúc dữ liệu của nguồn -> tôi sẽ làm cái gì -> cách tôi làm có tác dụng trong việc xử lý gì không

—----------

star rock

câu hỏi đặt ra là dữ liệu 2 chiều đẩy vào thì có khác gì nhau không

Minio

cần biết biến đổi dữ liệu, xử lý như thế nào

dùng framework nên không nhìn thấy xử lý như thế nào

dữ liều đầu vào là gì

có 1 luồng dữ liệu tại sao cần làm sạch dữ liệu

mục tiêu là gì, với nguồn dữ liệu -> giải quyết gì

mô phỏng tổ chức lưu trữ thế nào

tìm ý nghĩa của việc xử lý dữ liệu

luồng xử lý dữ liệu, đâu là vấn đề cần giải quyết

—--------