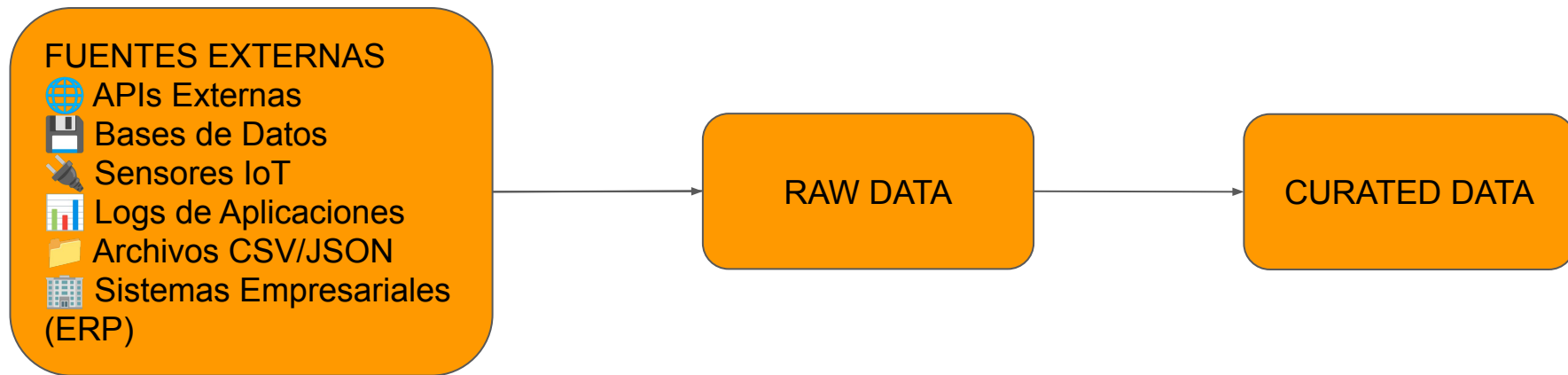


Raw Data → Curated Data →  
Feature Store →  
Experimentation Environment



# RAW DATA: FUENTE INICIAL DEL PIPELINE

Fuente inicial de datos SIN PROCESAR que marca el INICIO del ciclo de vida MLOps. Datos no estructurados que requieren procesamiento.





# RAW DATA: PUNTO CRÍTICO DEL PIPELINE MLOPS

## ♦ NATURALEZA DEL FLUJO

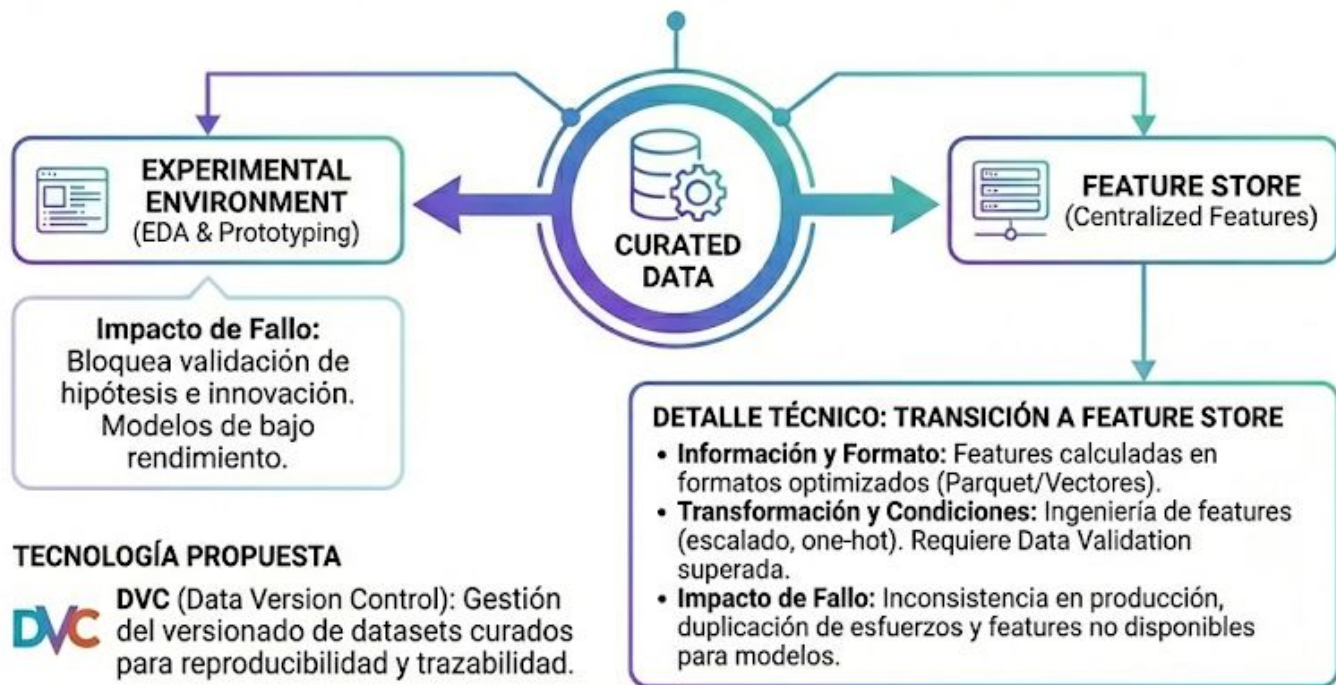
- Raw Data → Curated Data es un MOVIMIENTO DE DATOS
- Transfiere información física (archivos, tablas, gigabytes)
- SI puede contener metadatos (Timestamps, versiones de datasets)

## ♦ PUNTO ÚNICO DE FALLO

⚠ Raw Data es el INICIO del pipeline → Si falla, TODO se para.

# CURATED DATA

**Curated Data:** Datos procesados, limpiados y estructurados a partir de fuentes crudas, optimizados para ML.



## EJEMPLO DE FALLO Y SOLUCIÓN

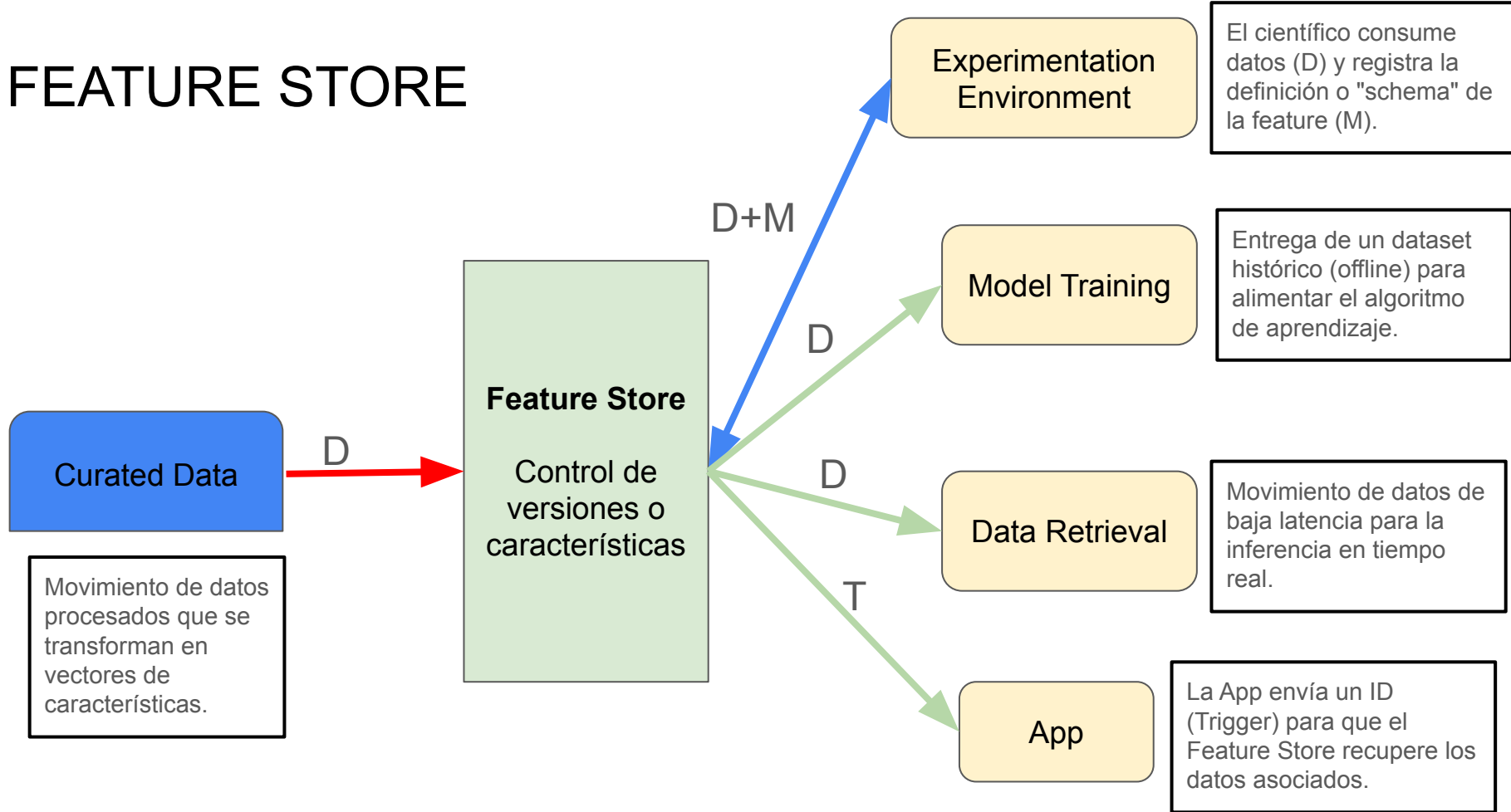


**Posible Fallo:**  
Inconsistencia (ej. valores nulos inesperados) antes del pipeline.



**Solución:**  
Implementar Data Validation con reglas automáticas de calidad.

# FEATURE STORE



# Experimentation Environment

