

Bridging Data Engineering and Generative AI: Building the Intelligent Data Fabric
By Subhash Tatavarthi

Abstract

Enterprises everywhere are exploring how Generative AI can transform their operations, but few realize that its success depends on a strong data foundation. Large language models (LLMs) and other GenAI tools need high-quality, well-governed, and context-rich data to produce meaningful outcomes. This paper explains how the traditional world of data engineering must evolve to support this new class of intelligent systems. It introduces the concept of an Intelligent Data Fabric — a dynamic architecture that connects data, models, and business context into one adaptive ecosystem.

1. Introduction

The rise of Generative AI has changed how enterprises use data. Traditional data warehouses and pipelines, once sufficient for business intelligence, now fall short in serving the dynamic, context-driven demands of AI models. The role of data engineers is expanding — from building static pipelines to designing intelligent, adaptive data ecosystems.

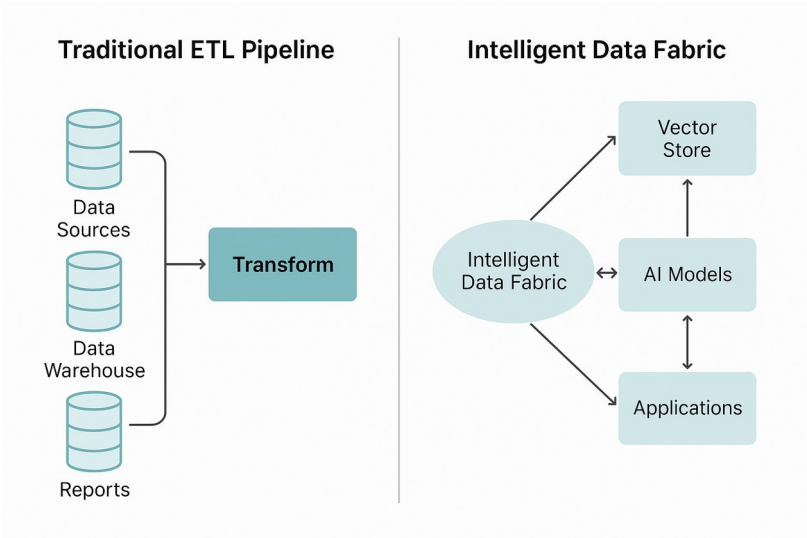


Figure 1 – Traditional vs Intelligent Data Flow

2. From Data Pipelines to Data Fabrics

A data fabric is an architectural approach that weaves data across multiple systems, clouds, and applications into a unified, intelligent layer. It automates data discovery, integration, and governance using metadata, semantics, and AI. Unlike traditional ETL systems, a data fabric adapts dynamically to new data sources and business needs.

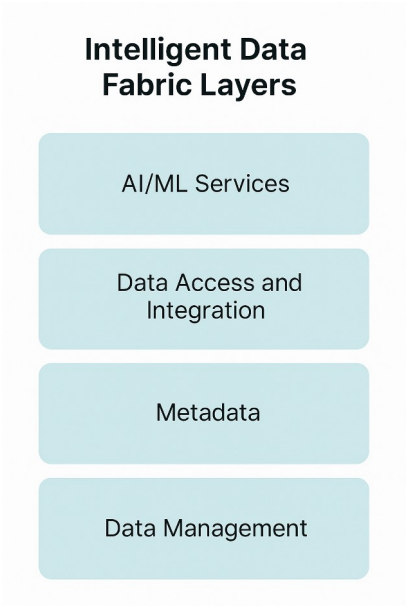


Figure 2 – Intelligent Data Fabric Layers

3. Integrating Generative AI into the Fabric

In an Intelligent Data Fabric, Generative AI acts as both a data consumer and collaborator. It uses the data fabric to access real-time, structured, and unstructured data while also improving it through feedback loops. Model outputs are continuously logged, analyzed, and used to enhance data quality — forming a cycle of learning and optimization.

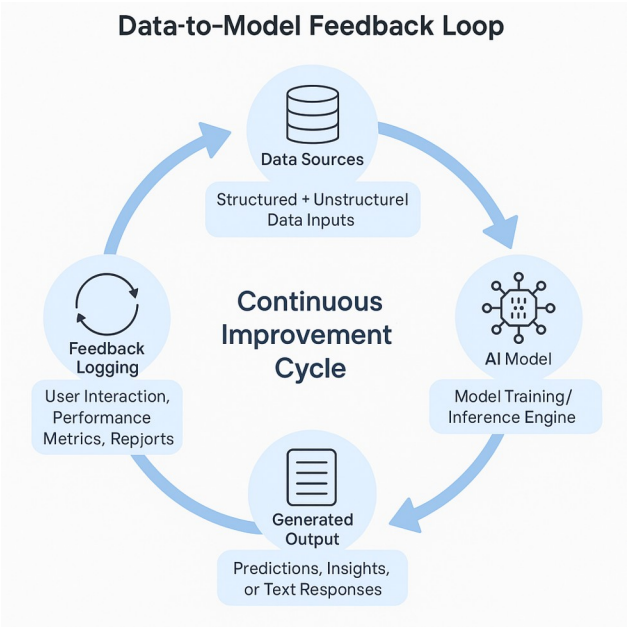


Figure 3 – Data-to-Model Feedback Loop (placeholder for visual diagram)

4. Phased Implementation Roadmap

Enterprises can adopt the Intelligent Data Fabric in phases:

- 1. Discovery – Map data assets and ownership.
- 2. Modernization – Adopt hybrid or cloud-native infrastructure.
- 3. Governance – Establish metadata and lineage controls.
- 4. Intelligent Integration – Enable semantic and AI layers.
- 5. Continuous Learning – Close feedback loops between data and models.

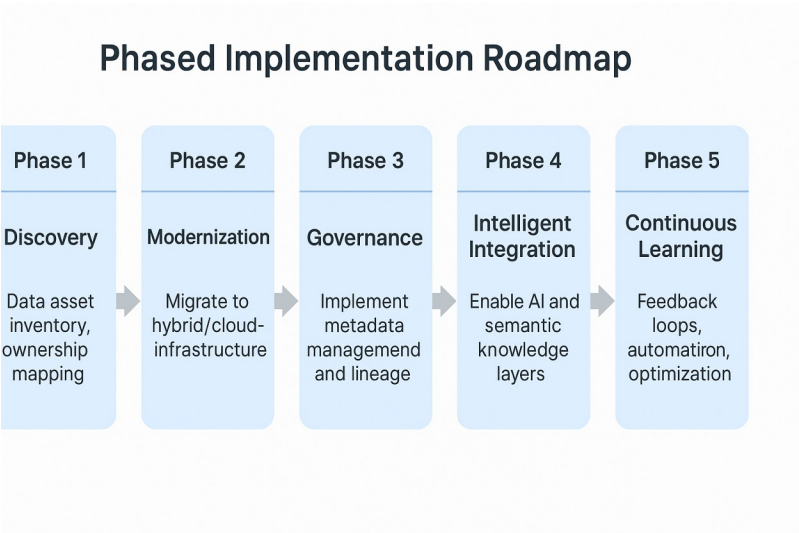


Figure 4 – Phased Implementation Roadmap (placeholder for visual diagram)

5. Conclusion

Generative AI depends on the strength of its data foundation. Building an Intelligent Data Fabric allows enterprises to unify governance, quality, and intelligence into one cohesive system. By bridging data engineering and AI, organizations can turn raw information into trusted, actionable knowledge — a key step toward sustainable decision intelligence.