Data Management in AI Projects
Key Concepts in Data Management

Learning Objective

By the end of this lesson, learners will be able to:

- Understand the importance of data management in AI projects.
- Identify key challenges in managing AI data (governance, security, quality, scalability).
- Recognize core strategies for handling AI data effectively.

Why Data Management Matters in AI

AI models are only as good as the data they are trained on. Poor data management leads to unreliable AI outcomes, compliance risks, and inefficient workflows. Organizations need strong data governance, quality control, and security measures to scale AI effectively.

Key Data Management Challenges

Challenge	Description
Volume	Managing large-scale data storage and processing efficiently.
Variety	Handling diverse data types (structured, unstructured, semi-structured).
Velocity	Ensuring real-time or fast processing of streaming data.
Quality	Maintaining clean, consistent, and unbiased data for AI models.
Labeling	Efficiently annotating and categorizing data for supervised learning.
Security	Protecting sensitive data from unauthorized access and breaches.

Core Strategies for Handling AI Data
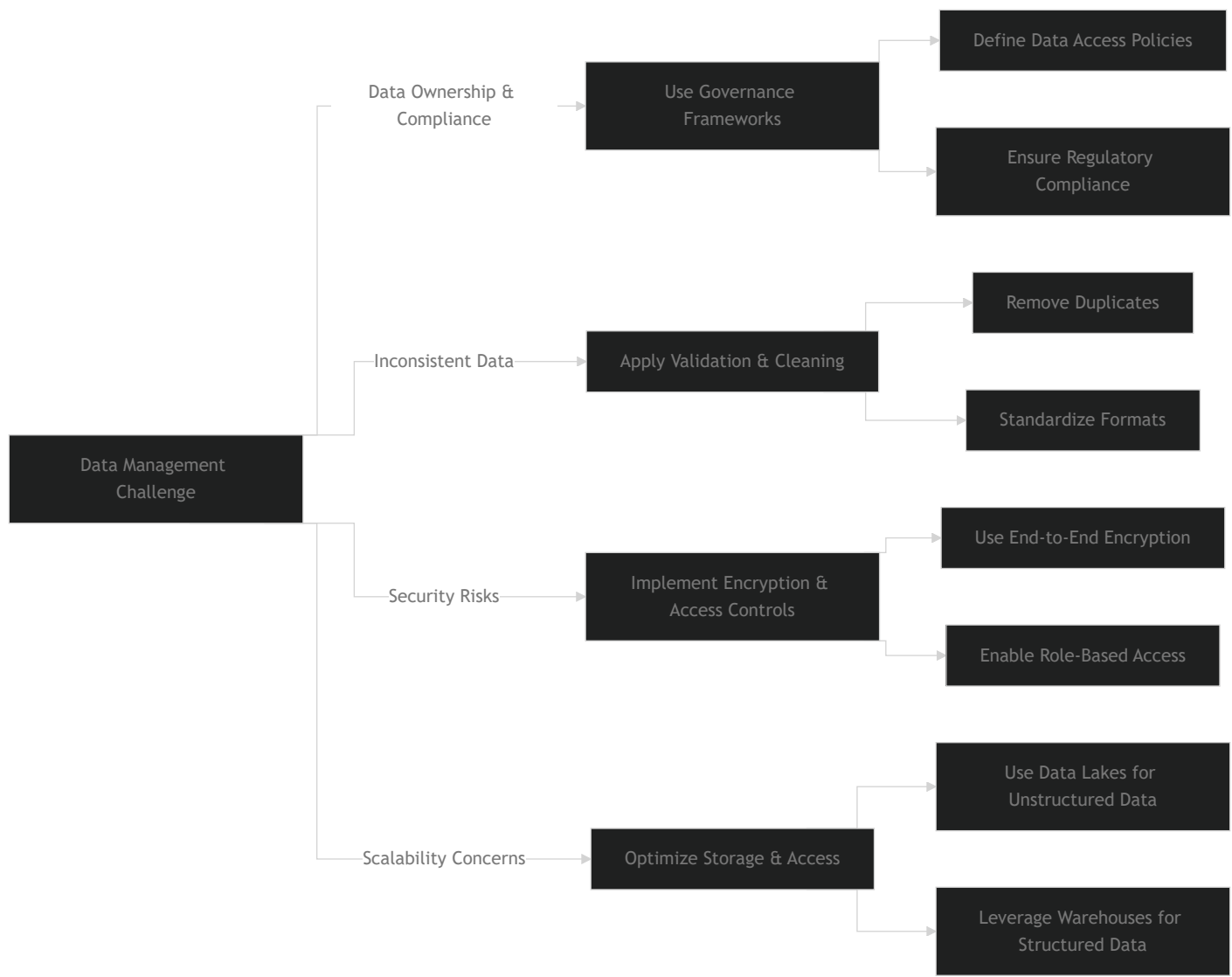
Group Activity: Data Risk Assessment

Scenario: Your team is training an AI model on customer transaction data for a large retail client. Identify three potential risks related to:

- Data quality

- Privacy and security

- Scalability

Discuss your findings and suggest mitigation strategies - be prepared to discuss with the class!

< Previous

Next >