- Document - analyser
- Config
- Model - Loader

Central Repo.

Rag (Archive)
Dataversioning → every execution

today →
Previous data
(Delete)

Embedd →

VectorDB

(Newdata)

tomorrow →

Ingestion

Data redundancy
↓
data duplication
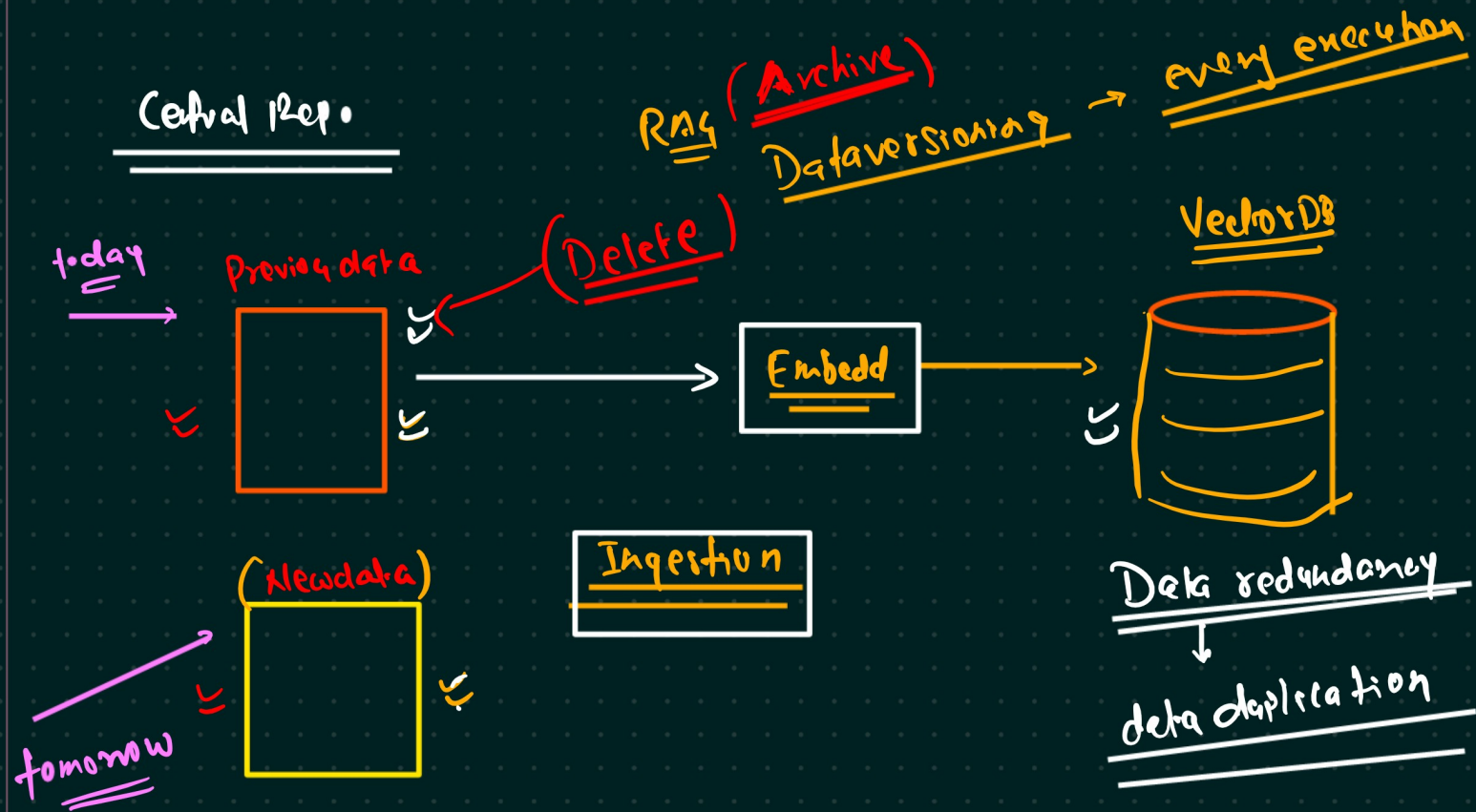
every ⟹ execute ⟹ (Session - Used)
                     ↳ Deta =

Data pipeline

Prod.

Period time

S3 ← Storage clan ← Athena
                      Aws clue
1year ⟵ hot. Storage   Redshift
        Cold-Storage
        warm. Storage
        Archival

Data

| PDF Files |

(Load this PDF file) → ⌣ Save. file   ⌐(Load document.)

                                ↓
                        | Session folder | | Load the Content. of PDF |

                        | Archival |

←——————————— Data insertion ——————————→

Content of PDF File ⌣ —— (Analysis this Content. using LLM capability)→ | Desired O1P |
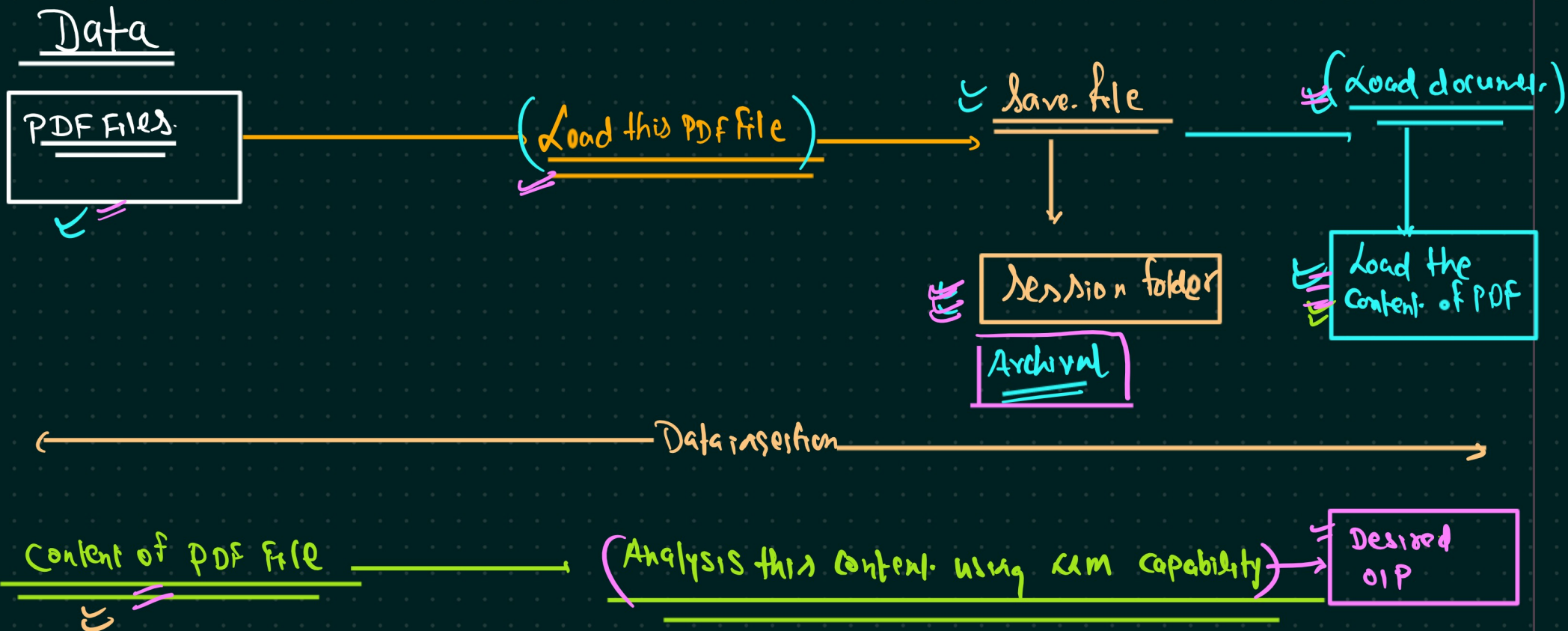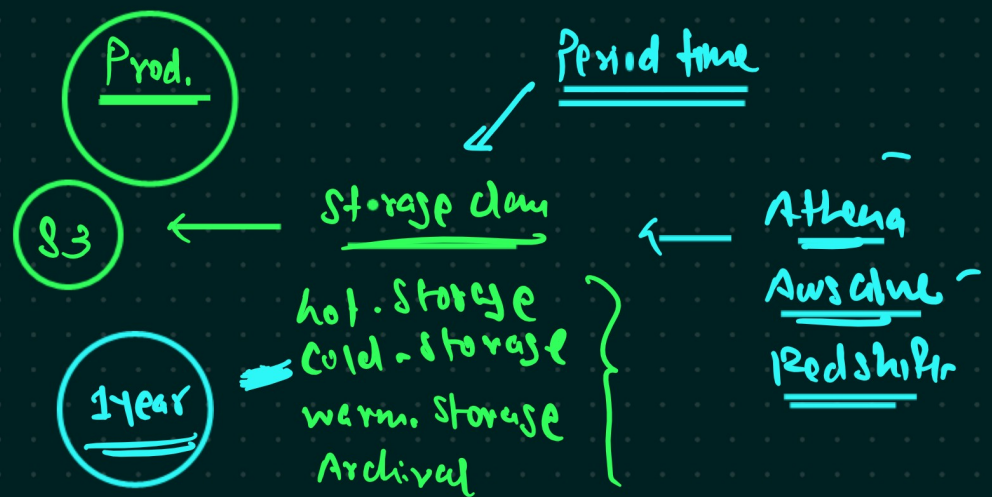
Scalable

Prompting -
Pydantic data validation

↓ Chain