

Introduction To

Data Warehouse

What is a warehouse?



▶
A warehouse is a large building where goods and products are stored before they are distributed or sold. It serves as a central location for storing inventory and often includes facilities for organizing, tracking, and managing the flow of goods.

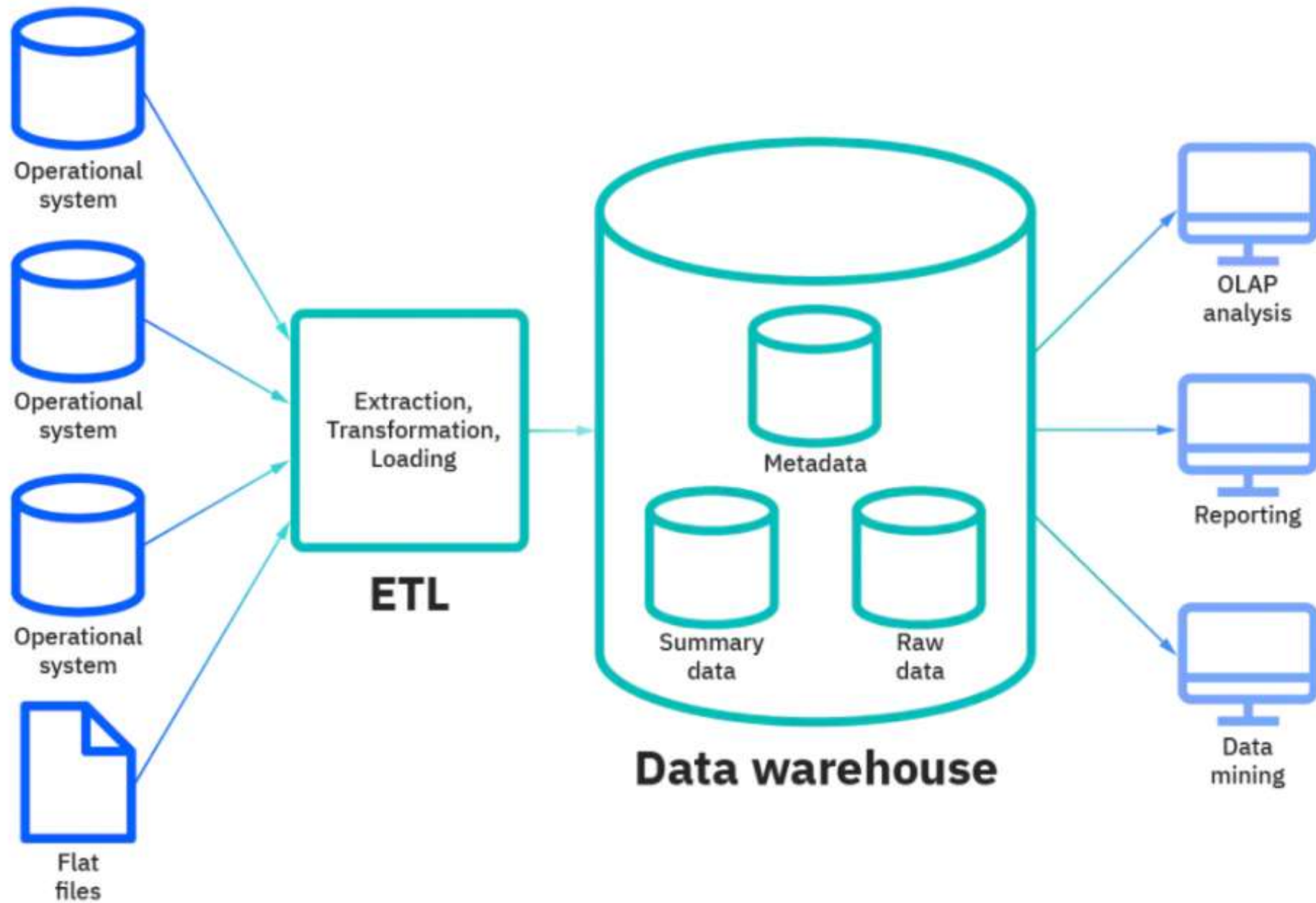


Then what is data
warehouse.....?



A data warehouse is a centralized repository that integrates data from various sources within an organization, designed for the purpose of analysis and reporting. It typically stores structured data in a format optimized for query and analysis, enabling users to efficiently extract insights and make data-driven decisions. Data warehouses often employ techniques such as data cleaning, transformation, and aggregation to ensure consistency and usability of the stored information.





ETL process



```
graph TD; ETL[ETL process] --> Extract((Extract)); ETL --> Transform((Transform)); ETL --> Load((Load));
```

The diagram illustrates the ETL process. At the top, a large blue oval contains the text "ETL process". Three arrows point downwards from this oval to three separate blue circles below it. The circles are labeled "Extract", "Transform", and "Load" from left to right, representing the sequential steps of the data integration process.

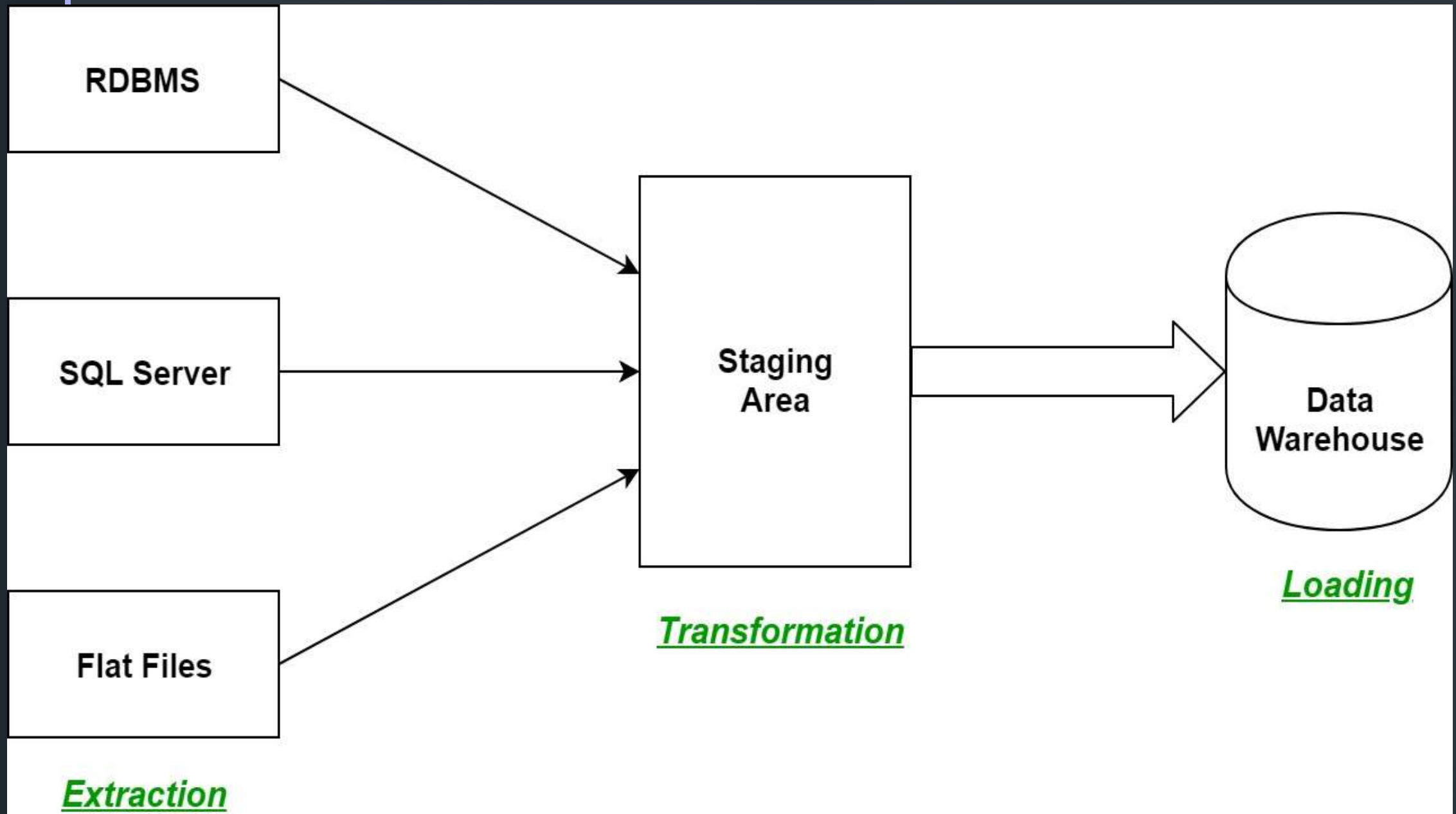
Extract

Transform

Load

Data Extraction

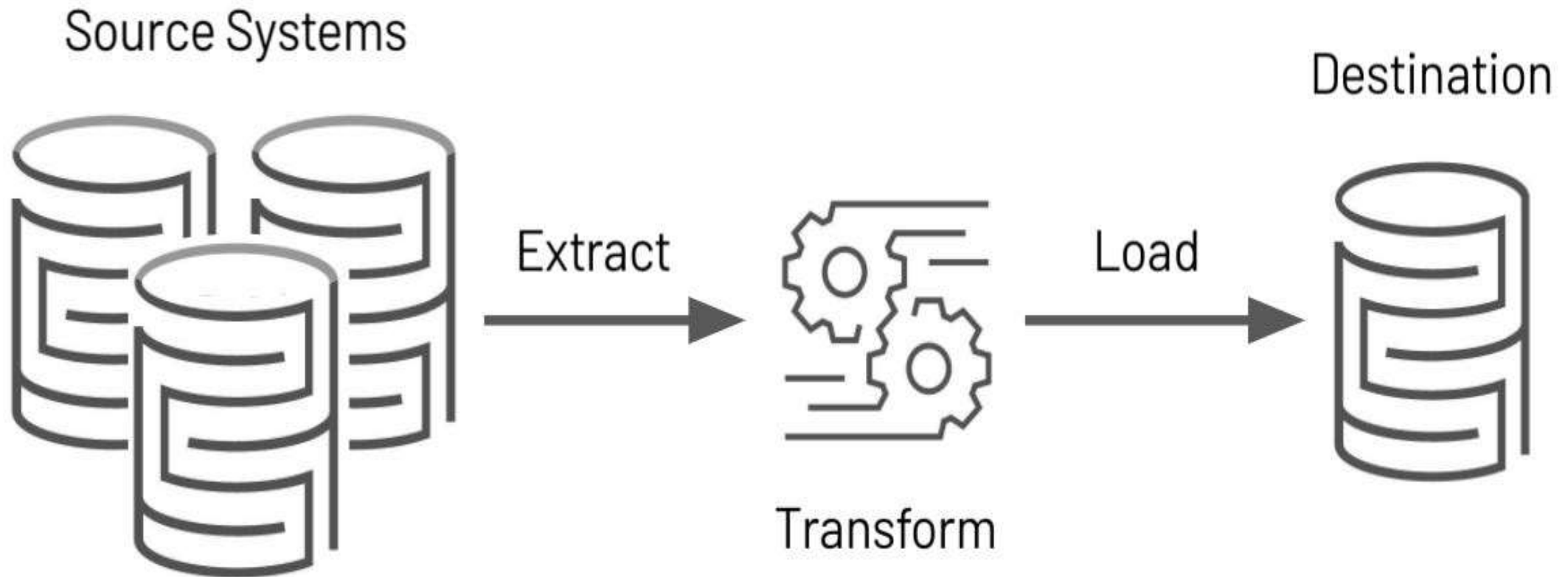
Data extraction in a data warehouse refers to the process of retrieving structured or semi-structured data from multiple sources, including operational databases, external systems, or other data repositories, and loading it into the data warehouse environment. This extraction process involves identifying relevant data, transforming it into a consistent format suitable for analysis, and moving it into the data warehouse's storage infrastructure.



Data Transform

Data transformation in a data warehouse refers to the process of converting and manipulating raw data from its original format into a structured, consistent, and analytically useful form. This process involves several steps, including data cleansing, normalization, aggregation, and enrichment.

ETL Process



Data Loading

Data loading in a data warehouse is the process of transferring structured or semi-structured data from external sources, such as operational databases, flat files, or other data repositories, into the data warehouse environment

Process



Data Warehouse



Thank
you

