

Overview of Data Repositories



IBM Developer



Introduction

Data Repository is a general term used to refer to data that has been collected, organized, and isolated



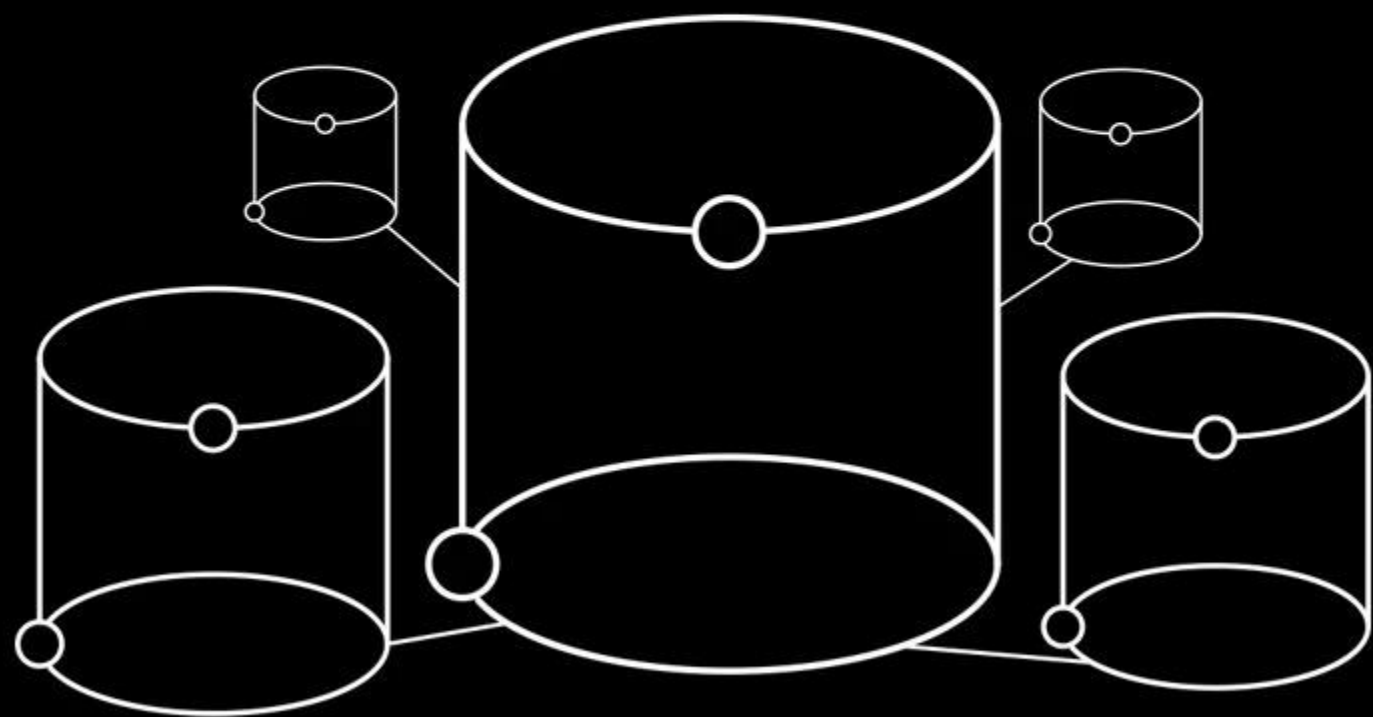
for use in business
operations



mined for reporting
and data analysis



Introduction



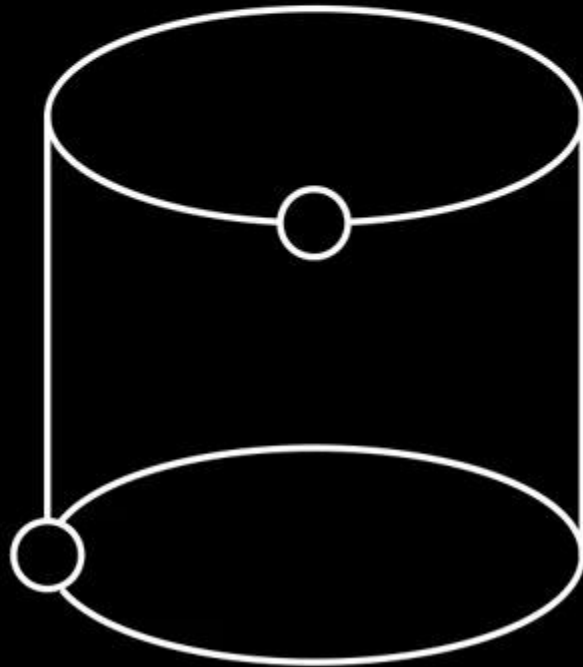
Types of data repositories include:

- Databases
- Data Warehouses
- Big Data Stores



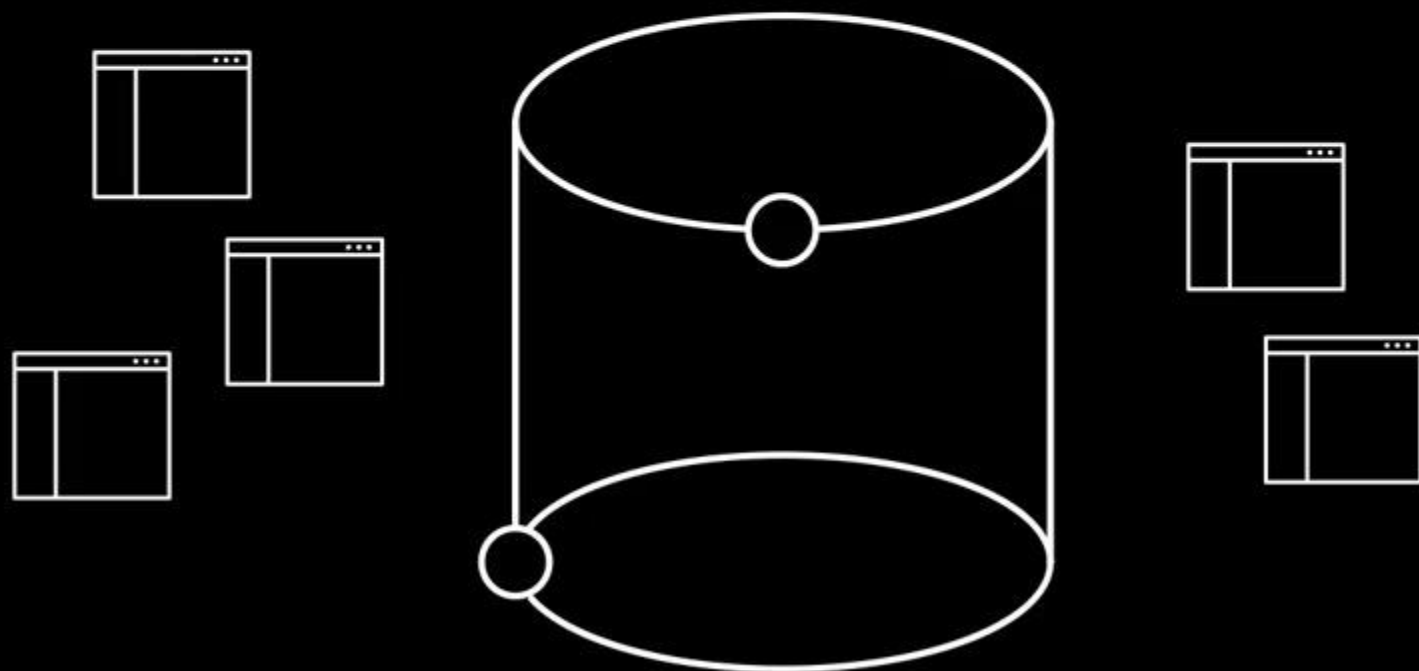
Databases

Collection of data for input, storage, search, retrieval, and modification of data.



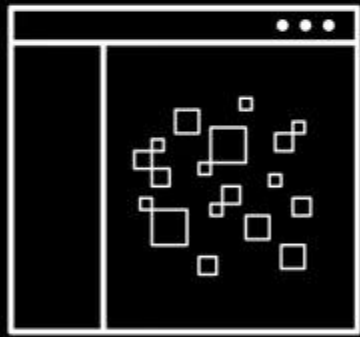
Databases

Set of programs for creating and maintaining the database, and storing, modifying, and extracting information from the database.

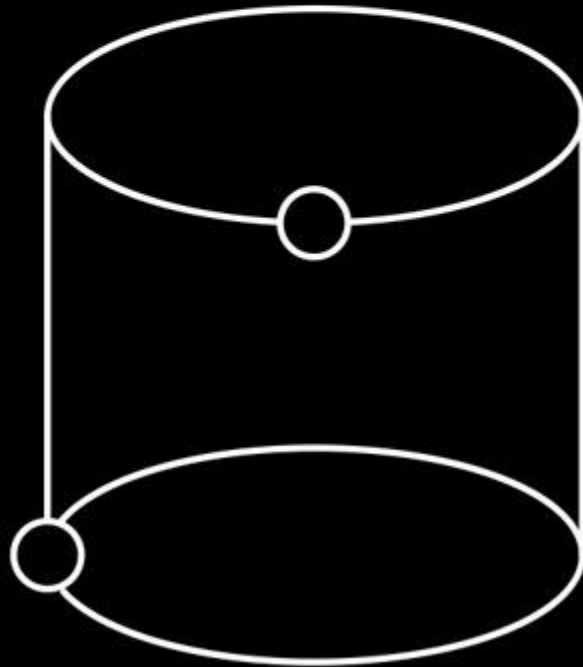


Databases

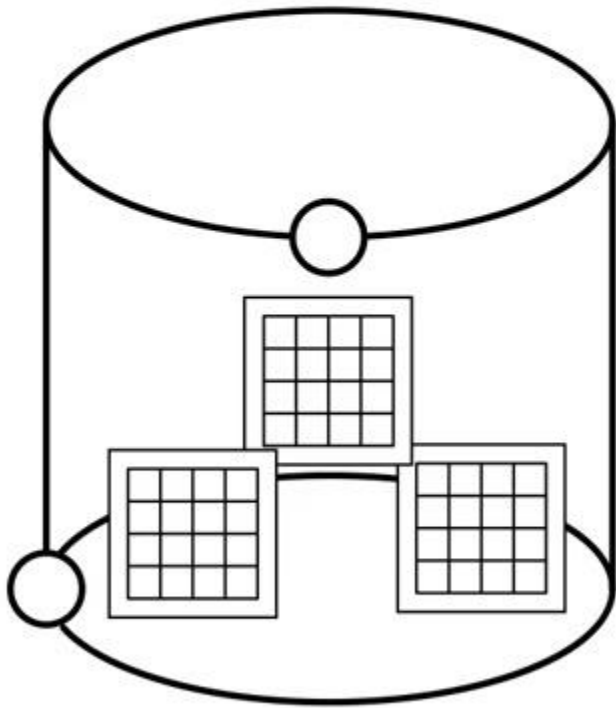
Even though a database and DBMS mean different things the terms are often used interchangeably.



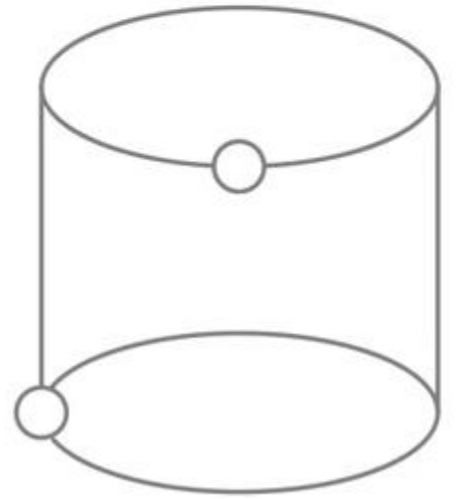
Querying



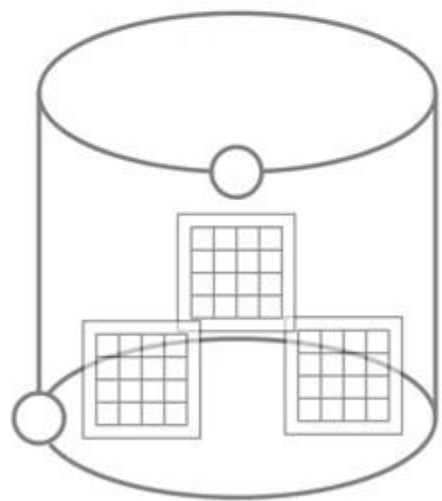
Relational Databases



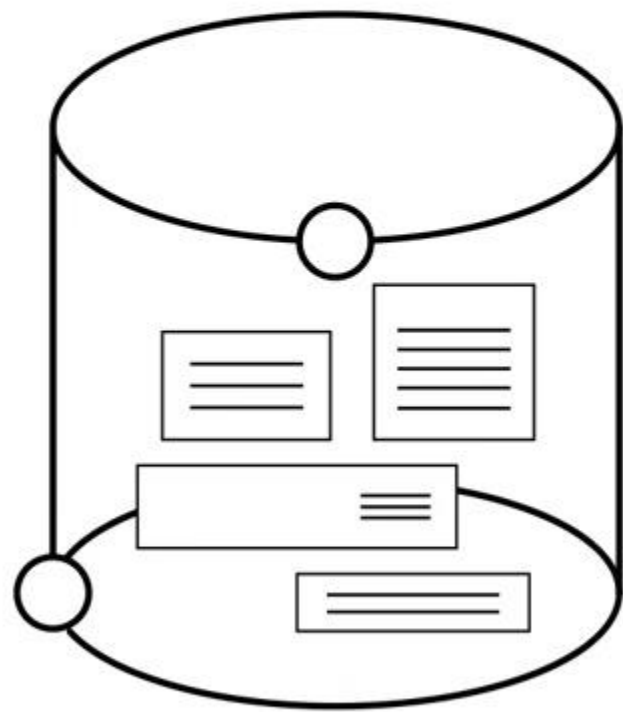
- Data is organized into a tabular format with rows and columns
- Well-defined structure and schema
- Optimized for data operations and querying
- Use SQL as the standard querying language



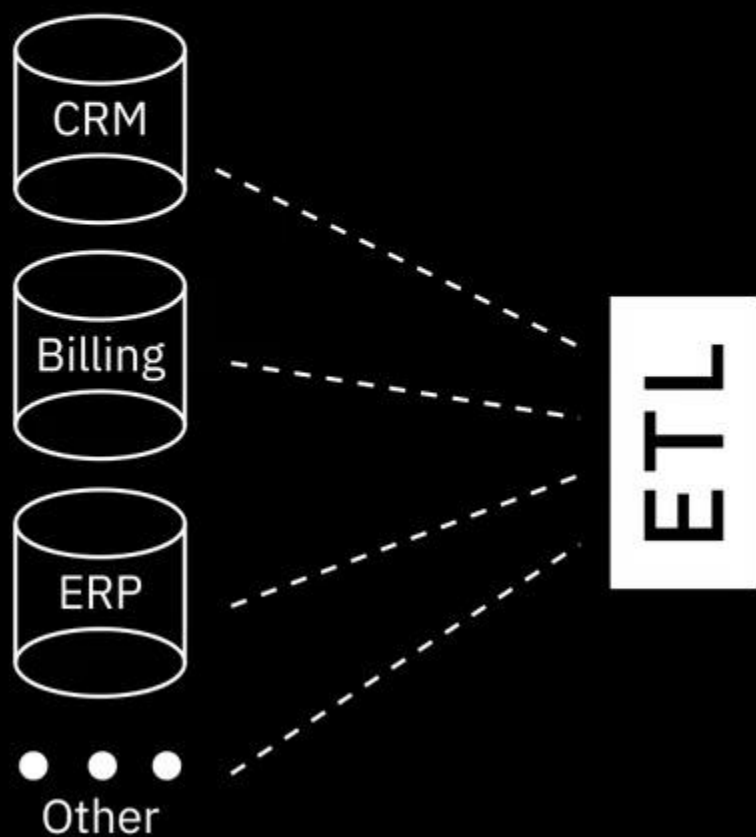
Non-Relational Databases



- Emerged in response to the volume, diversity, and speed at which data is being generated today
- Built for speed, flexibility, and scale
- Data can be stored in a schema-less form
- Widely used for processing big data



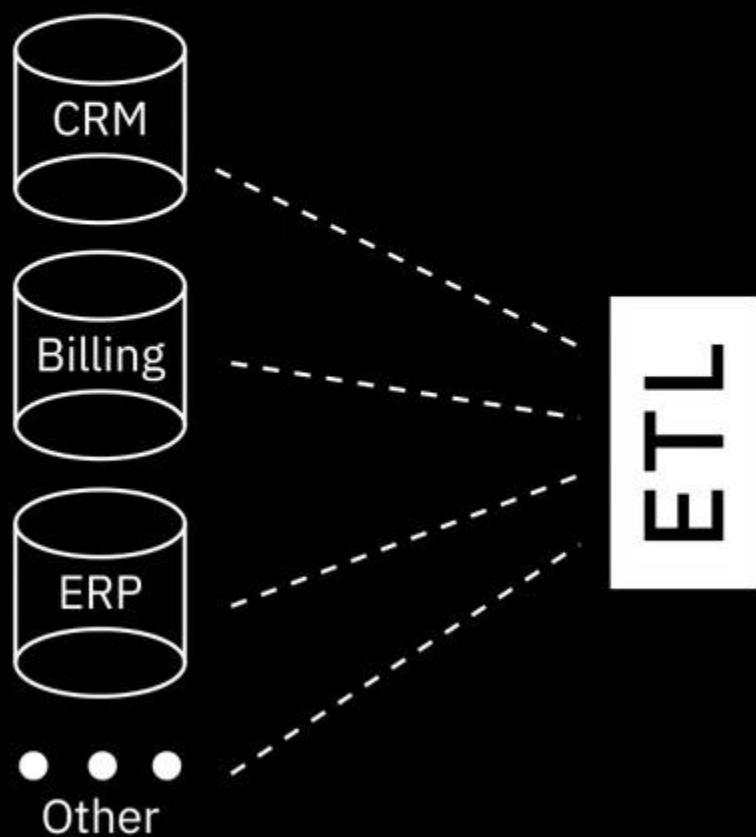
Data Warehouse



Consolidates data through the extract, transform, and load process, also known as the ETL process, into one comprehensive database for analytics and business intelligence.



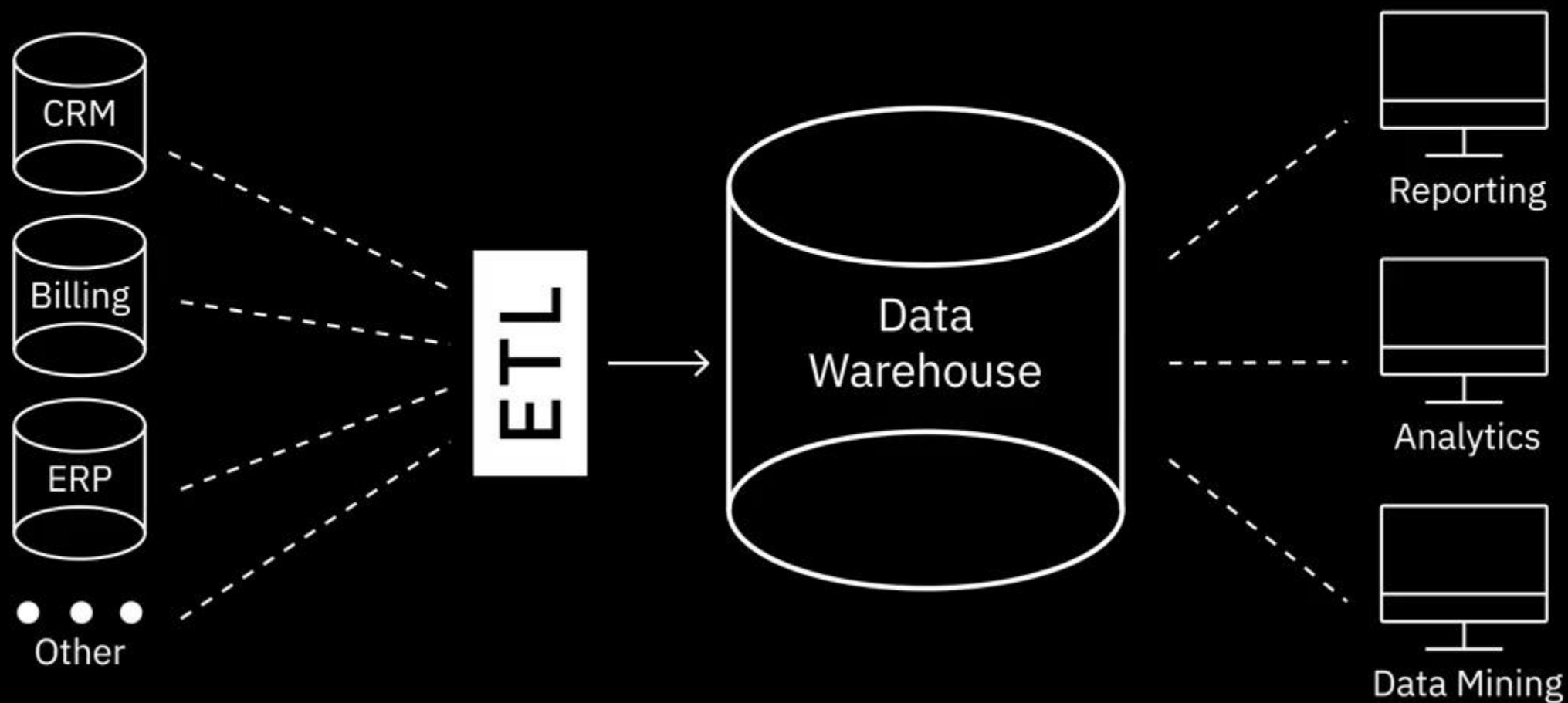
Data Warehouse



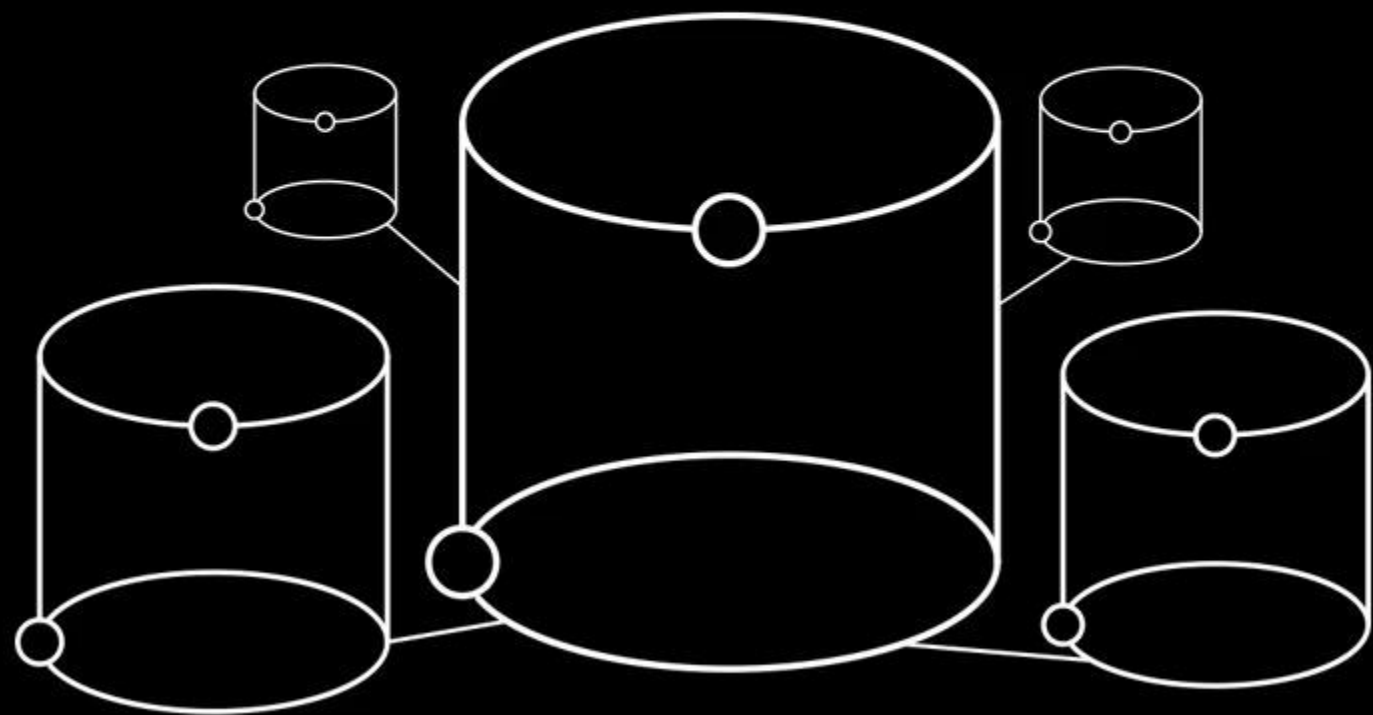
- Extract data from different data sources
- Transform the data into a clean and usable state
- Load the data into data repository



Data Warehouse



Big Data Stores



Distributed computational and storage infrastructure to store, scale, and process very large data sets.



Summary

Data repositories help to isolate data and make reporting and analytics more efficient and credible while also serving as a data archive.

