

Grow With EDM

Database vs Data Warehouse vs Data Lake



Enterprise
Data
Management
Laboratory

DATA STORAGE

- Sebuah mekanisme atau tempat dimana data tersebut di simpan dengan konfigurasinya masing – masing.



DATABASES, DATA WAREHOUSES, DATA LAKES

GWE #2

Apaa Bedanyaa??

DATABASE

DATABASE

Jenis Database: Terbagi menjadi 2
(Relational DBMS dan NOSQL).

Karakteristik:

- OLTP
- real Time
- ACID Transactions
- Flexible Schema

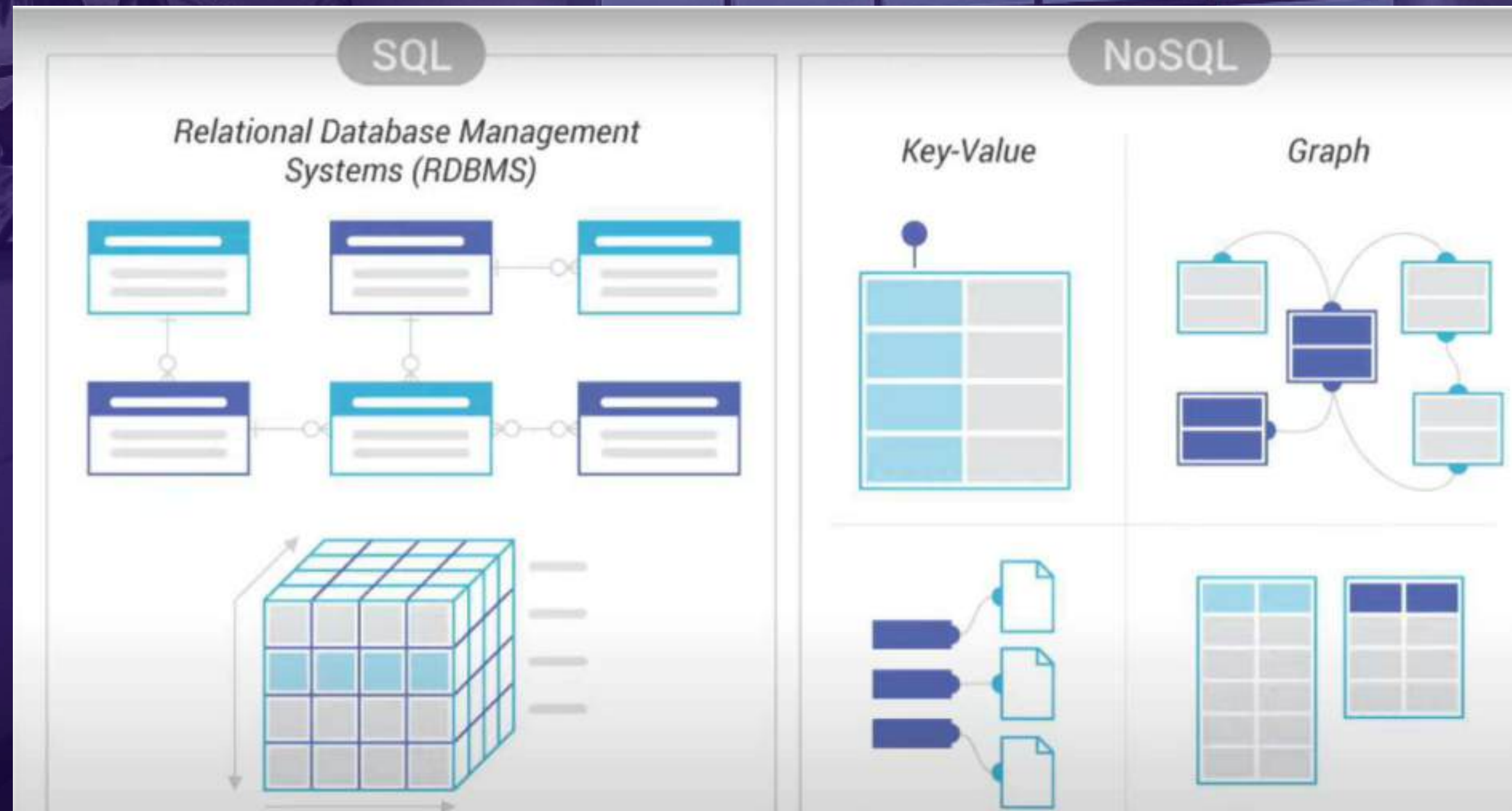
Pros:

- Efisien
- Security
- Fast Queries

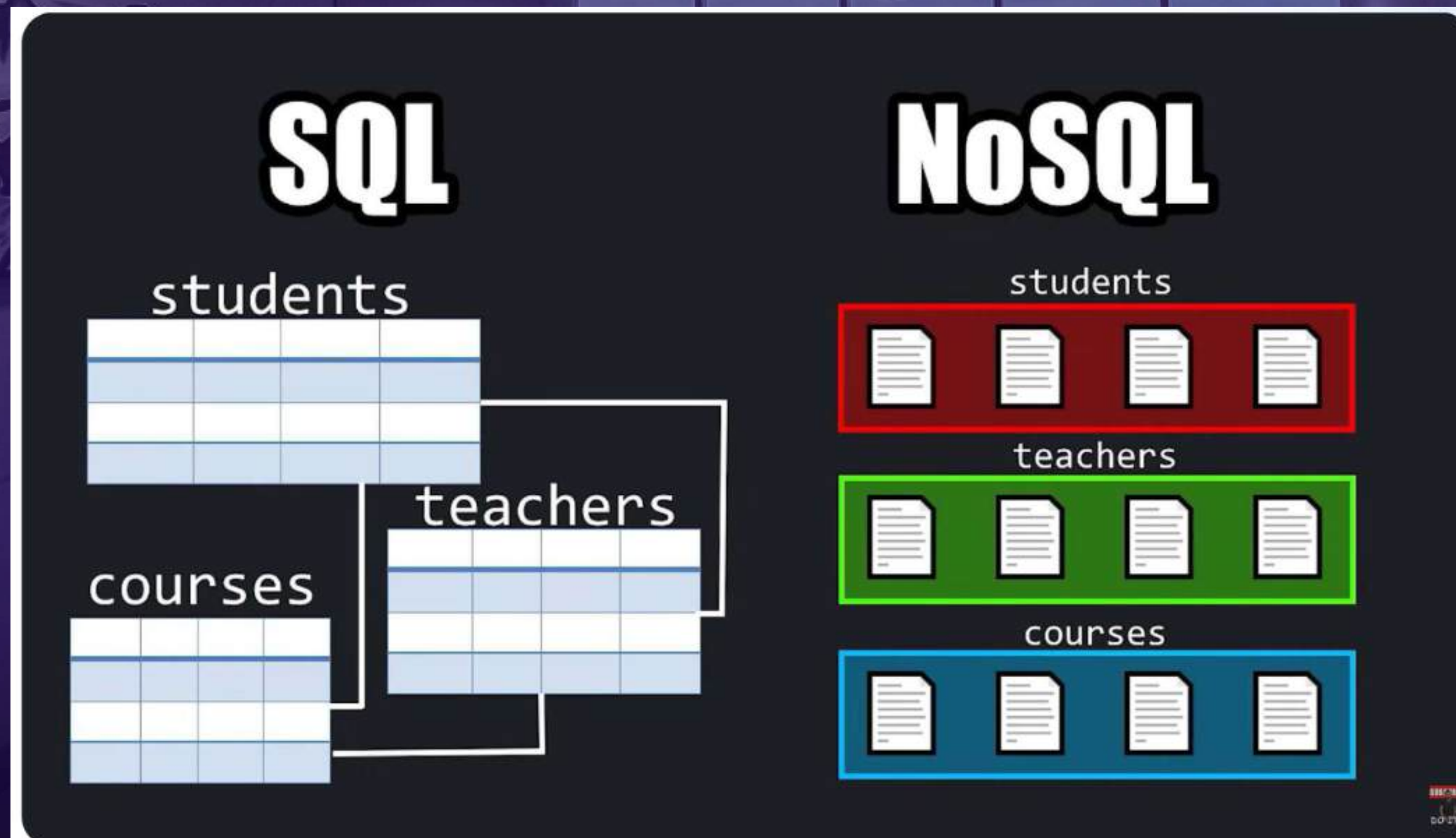
Cons:

- Keterbatasan Kapabilitas
- Kompleksitas
- Skalabilitas

DATABASE



DATABASE



DATABASE

SQL

name	age	gpa	fullTime
Spongebob	32	3.2	false
Patrick	38	1.5	false
Sandy	27	4.0	true

NoSQL

```
{  
  name: 'Spongebob',  
  age: 30,  
  gpa: 3.2,  
  fullTime: false,  
},  
{  
  name: 'Patrick',  
  age: 38,  
  gpa: 1.5,  
  fullTime: false,  
},  
{  
  name: 'Sandy',  
  age: 27,  
  gpa: 4,  
  fullTime: true,  
}
```



DATABASE

SQL

- + Query data relational lebih mudah
- + Data lebih terorganisir dan terstruktur,
- + Transaksi yang aman

- Prepared structure
- Hard to scale (horizontal scaling)

NOSQL

- + Flexible
- + Penyimpanan cepat untuk data tidak terstruktur

- Kurang cocok untuk aplikasi yang memerlukan konsistensi tinggi
- Query kompleks

DATA WAREHOUSE



Enterprise
Data
Management
Laboratory

DATA WAREHOUSE

- Digunakan untuk mendukung Online Analytical Processing (OLAP)
- Rigid Schema
- Data dipindahkan dari sumber aslinya ke warehouse melalui proses ETL.
- Dapat dihubungkan ke alat Business Intelligence (BI) seperti Tableau, Power BI, atau Looker.
- Bersifat Historical



DATA WAREHOUSE

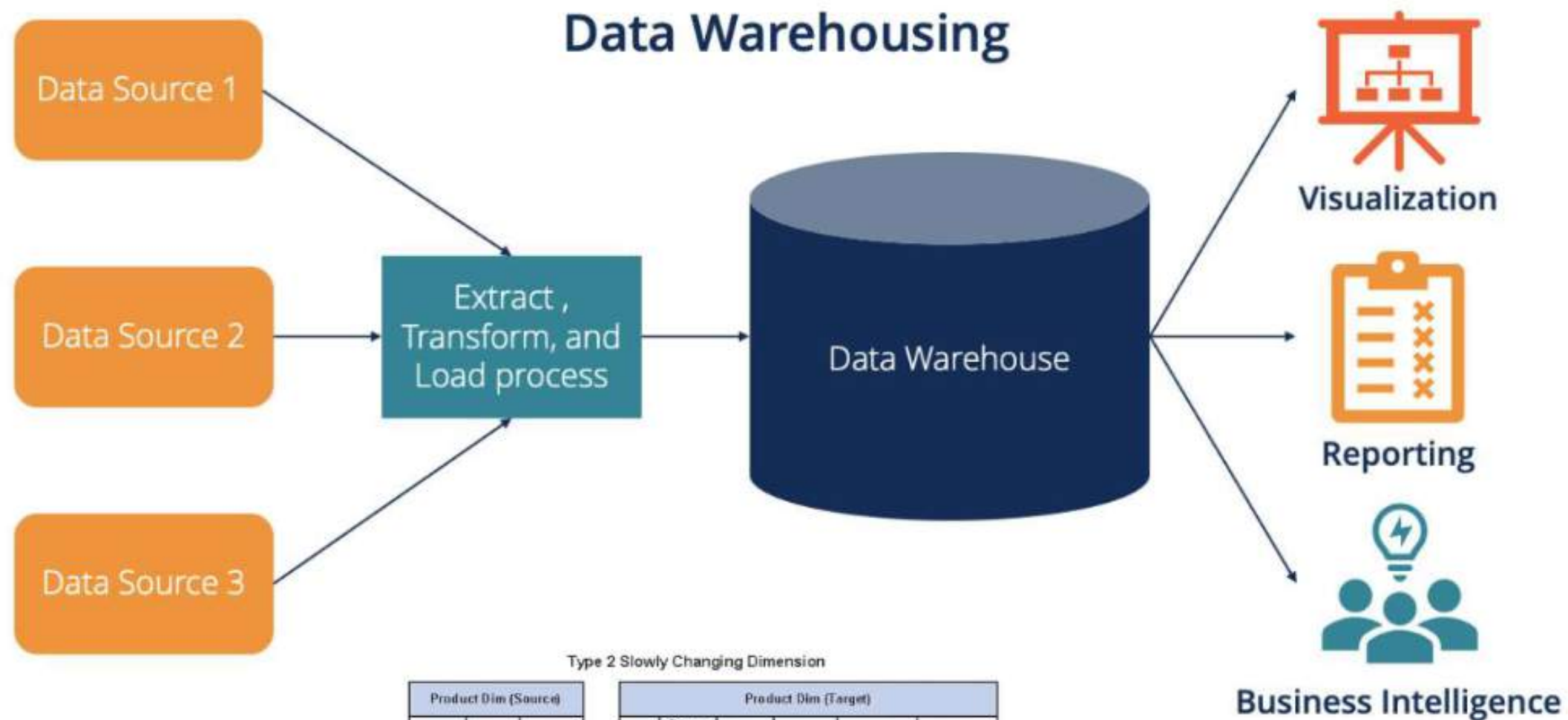
PROS

- + Mendukung Analisis Historis
- + Kinerja Optimal untuk OLAP
- + Data Konsisten dan Terorganisir
- + Mendukung Pengambilan Keputusan

CONS

- Tidak Mendukung Real-Time Data
- Tidak untuk Operasi Transaksional
- Kompleksitas Implementasi
- Kebutuhan Ruang Penyimpanan Besar

DATA WAREHOUSE



DATA LAKE



Enterprise
Data
Management
Laboratory

DATA LAKE

- Designed to capture raw data
- AI/ML Purposes
- Schema-on-read
- Fleksibilitas
- Volume Data Besar



GWE #2

DATA LAKE

PROS

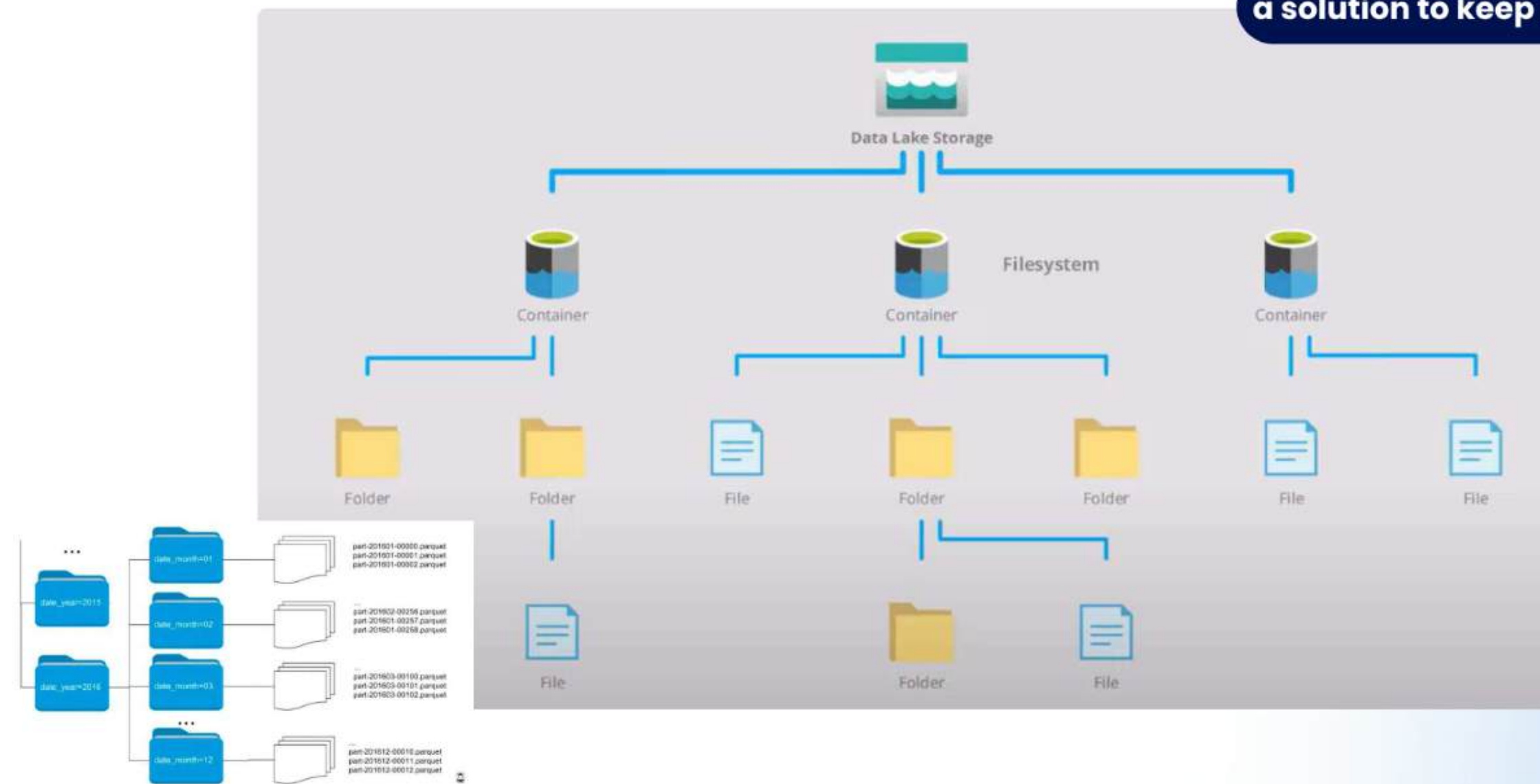
- + Mendukung berbagai format data
- + More Security and Governance
- + Mendukung Machine Learning dan analisis prediktif yang kompleks.
- + Mendukung Pengambilan Keputusan

CONS

- Data Swamp.
- Tidak seefisien data warehouse untuk analisis data yang terstruktur.
- Tidak dioptimalkan untuk analisis data real-time atau kebutuhan transaksi.

DATA LAKE

a solution to keep raw data





Enterprise
Data
Management
Laboratory

APA BEDANYA SAMA DATA WAREHOUSE???????

YG SATU GUDANG, YG SATU DANAU XIXIXI

GWE #2



Data Lakes



Raw

Data Lakes contain unstructured, semi structured and structured data with minimal processing. It can be used to contain unconventional data such as log and sensor data

Large

Data Lakes contain vast amounts of data in the order of petabytes. Since the data can be in any form or size, large amounts of unstructured data can be stored indefinitely and can be transformed when in use only

Undefined

Data in data lakes can be used for a wide variety of applications, such as Machine Learning, Streaming analytics, and AI

Data Warehouse



Refined

Data Warehouses contain highly structured data that is cleaned, pre-processed and refined. This data is stored for very specific use cases such as BI.

Smaller

Data Warehouses contain less data in the order of terabytes. In order to maintain data cleanliness and health of the warehouse, Data must be processed before ingestion and periodic purging of data is necessary

Relational

Data Warehouses contain historic and relational data, such as transaction systems, operations etc

CONCLUSION

	DATABASE	DATA WAREHOUSE	DATA LAKES
WORKLOADS	TRANSACTIONAL	ANALYTICAL	ANALYTICAL
DATA TYPES	STRUCTURED	STRUCTURED	STRUCTURED, SEMI STRUCTURED, UNSTRUCTURED
SCHEMA FLEXIBILITY	FLEXIBLE	RIGID	SCHEMA-ON-READ
DATA UPDATES	REAL TIME	BASED ON PROCESS	BASED ON PROCESS
USERS	DEVELOPER	DATA ANALYST, BUSINESS ANALYST	DATA SCIENTIST, AI/ML ENGINEER



Enterprise
Data
Management
Laboratory

Let's try

Q & A



Enterprise
Data
Management
Laboratory

Quiz



Enterprise
Data
Management
Laboratory

BARCODE

ABSENSI

EDM Laboratory

GWE #2

Presensi GWE #2

<https://tel-u.ac.id/presensigwe>



GWE #2



Terima Kasih



GWE #3



@lifeatedmlab



@186mcgwc