

(Big) Data Engineering In Depth

From Beginner to Professional

Moustafa Alaa

Senior Big Data Engineer

 MoustafaAlaa  Moustafa Alaa  @Moustafa_alaa22

 Garage Education

 mustafa.alaa.mohamed@gmail.com

The Definitive Guide to Big Data Engineering Tasks

Videos classification

Watching Method / Audience	Computer	Mobile/Tablet	Just listening
Developer		●	
DevOps		●	
Business		●	

Table: Video classification

The green circle ● means short video.

The blue circle ● means medium video.

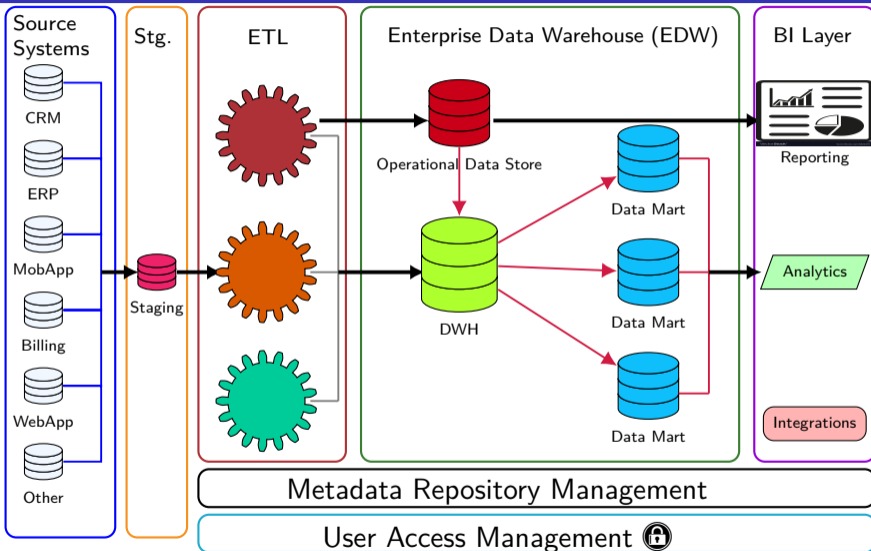
The red circle ● means long video

Section: DWH Characteristics

- The characteristics of DWH:
 - Integrated: *DWH is an integrated environment which allows us to integrate different source systems. Data are modeled (organized) in a unified manner.*
 - Time-Variant: *Data modeled (organized) based on periods (hourly, daily, weekly, monthly, quarterly, yearly).*
 - Subject-oriented: *DWH main target is to support business needs for the whole organization including (decision-makers, departments, and specific user requirements).*
 - Non-Volatile: *It refers to the data that erased or deleted (It could be archived and retrieved when needed). Data can be accumulated daily the new snapshots (refreshed at based on the source system interval. ↻ It could be updated daily, weekly, and monthly).*

Section: DWH Architecture

DWH Architecture Overview



DWH Architecture Layers

- DWH architecture contains the following layers:
 - Source system layer.
 - Extraction layer.
 - Staging Area.
 - Data Modeling.
 - ETL layer.
 - Storage layer.
 - Reporting (UI) layer.
 - Metadata layer.
 - System operations layer.

Videos classification

Watching Method / Audience	Computer	Mobile/Tablet	Just listening
Developer		●	
DevOps		●	
Business		●	

Table: Video classification

The green circle ● means short video.

The blue circle ● means medium video.

The red circle ● means long video