



MODELAGEM DW/DL

EXTRAÇÃO, TRANSFORMAÇÃO E CARGA (ETL)

INTRODUÇÃO

Sistemas ETL costumam demandar **tempo e esforço** para construir um ambiente DW/BI.

Variáveis que exercem influência e aumentam a complexidade:

- Necessidades do negócio;
- Realidade dos dados de origem;
- Orçamento;
- Habilidades do time.

Pensar nas **melhores práticas** pode ser relativo, mas Kimball sugere um caminho a seguir.



REQUISITOS

1. Necessidades do Negócio;
2. Compliance;
3. Data Qyality;
4. Segurança;
5. Data Integration;
6. Data Latency;
7. Arquivamento;
8. Interfaces de BI;
9. Habilidades disponíveis.



SUBSISTEMAS DO ETL

Extração: recebendo dados no DW

1. Data Profilling;
2. Change Data Capture System;
3. Extract System.



SUBSISTEMAS DO ETL

Cleaning and Conforming Data

1. Data Cleansing System;
2. Error Event Schema;
3. Audit Dimension Assembler;
4. Deduplication System;
5. Conforming System.



SUBSISTEMAS DO ETL

Delivering: Prepare for Presentation

1. Slowly Changing Dimension Manager;
2. Surrogate Key Generator;
3. Hierarchy Manager;
4. Special Dimensions Manager;
5. Fact Table Builders;
6. Surrogate Key Pipeline;
7. Multivalued Dimension Bridge Table Builder;
8. Late Arriving Data Handler;



SUBSISTEMAS DO ETL

Delivering: Prepare for Presentation

- 9. Dimension Manager System;
- 10. Fact Provider System;
- 11. Aggregate Builder;
- 12. OLAP Cube Builder;
- 13. Data Propagation Manager.



SUBSISTEMAS DO ETL

Managing the ETL Environment

1. Job Scheduler;
2. Backup System;
3. Recovery and Restart System;
4. Version Control System;
5. Version Migration System;
6. Workflow Monitor;
7. Sorting System;
8. Lineage and Dependency Analyser;



SUBSISTEMAS DO ETL

Managing the ETL Environment

- 9. Problem Escalation System;
- 10. Parallelizing/Pipeline System;
- 11. Security System;
- 12. Compliance Manager;
- 13. Metadata Repository Manager.



APRESENTAÇÃO DE CASO





Obrigado!

Professor Gustavo Dias
luizdias@univas.edu.br