



DIGITAL
INNOVATION
ONE

Aula 1: Introdução ao ETL

Fundamentos de ETL com Python

Objetivos

1. Definição
2. Por que precisamos
3. Visão geral
4. Algumas ferramentas / pacotes para Python
5. Para saber mais

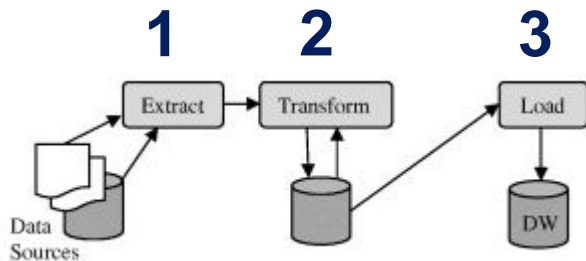
Aula 1 | Etapa 1:

ETL – Definição

Fundamentos de ETL com Python



ETL – Definição



1) **Extract:** os dados são extraídos de diferentes fontes de dados.

2) **Transform:** Propagados para a área de preparação de dados, onde são transformados e limpos.

3) **Load:** Carregados no data warehouse.

Aula 1 | Etapa 2:

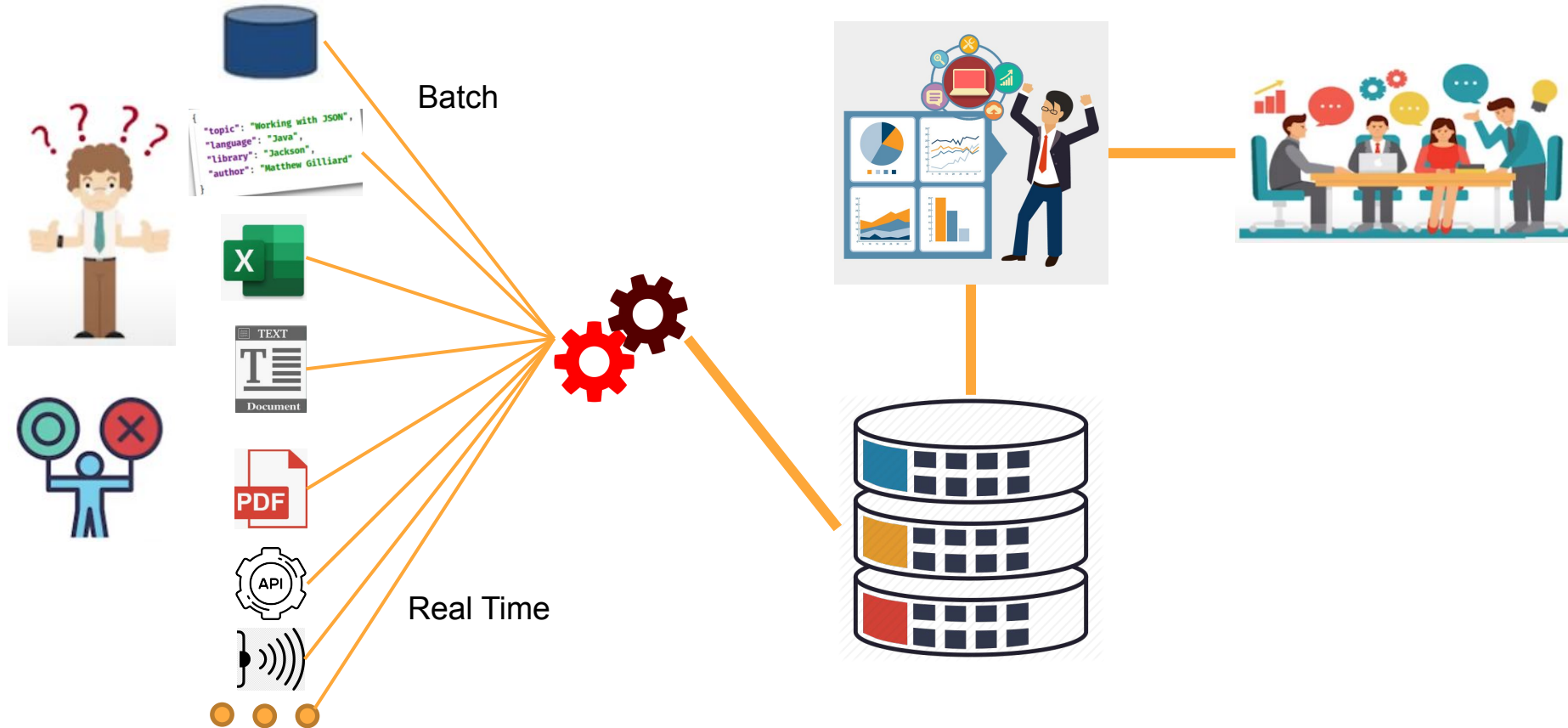
ETL - Por que precisamos?

Fundamentos de ETL com Python



DIGITAL
INNOVATION
ONE

ETL - Por que precisamos?



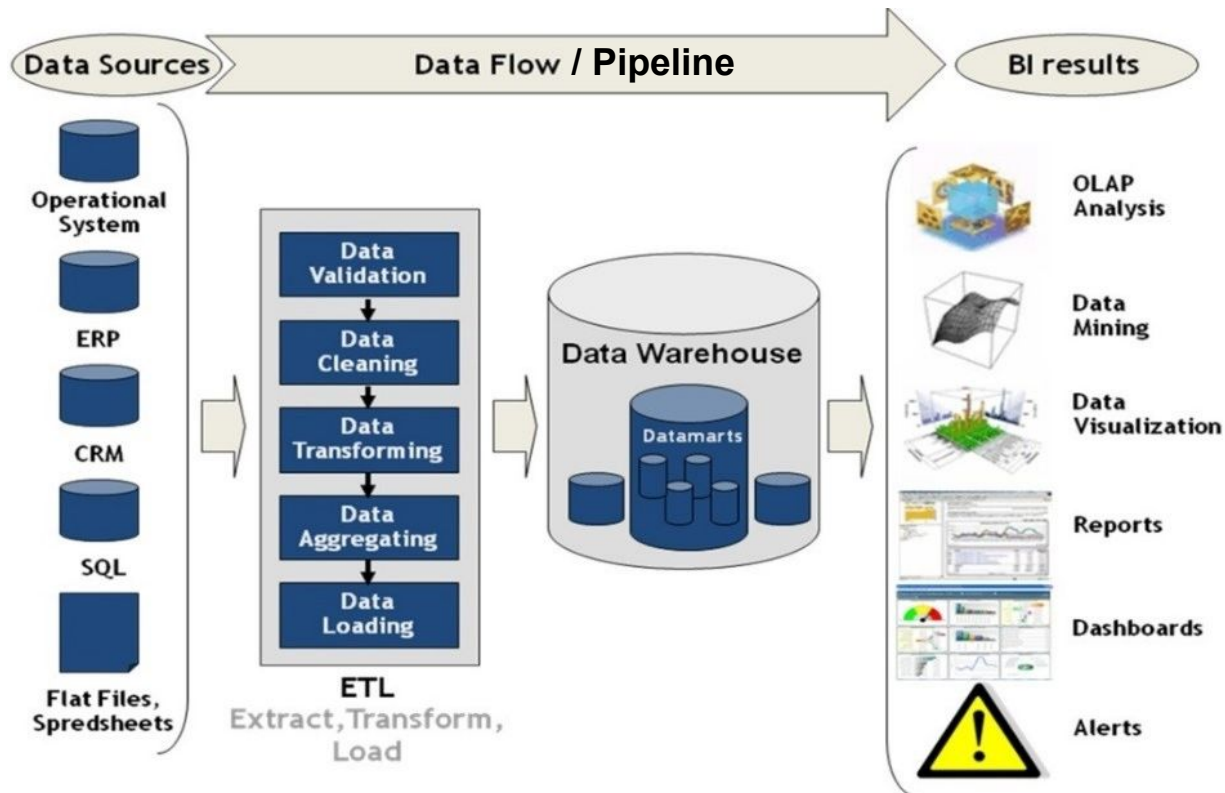
Aula 1 | Etapa 3:

ETL – Visão geral

Fundamentos de ETL com Python



ETL – Visão Geral



Aula 1 | Etapa 4:

ETL – Ferramentas/Pacotes

Fundamentos de ETL com Python



DIGITAL
INNOVATION
ONE

Ferramentas / Pacotes Para Python



Apache
Airflow



bonobo



bubbles

petl



pandas

Para saber mais

- <http://airflow.apache.org/>
- <https://luigi.readthedocs.io/en/stable>
- <https://www.bonobo-project.org/>
- <http://bubbles.databrewery.org/>
- <https://petl.readthedocs.io/en/stable/>
- [**https://pandas.pydata.org/**](https://pandas.pydata.org/)

Referências

- ALIEL-SAPPAGH, Shaker H.; AHMEDHENDAWI, Abdeltawab M.; BASTAWISSY, Ali HamedEl. A proposed model for data warehouse ETL processes. **Journal of King Saud University - Computer and Information Sciences**, [s. l.], ano 2011, v. 23, ed. 2, p. 91-104, Julho 2011.
- WILLIAM, Lourdes. **Why You Need ETL Testing and What You Need to Know**. [S. l.], 8 mar. 2021. Disponível em: <https://www.cigniti.com/blog/why-need-etl-testing-what-need-now/>. Acesso em: 16 maio 2021.

Dúvidas?

- > Fórum do curso
- > Comunidade online (discord)