O que é ETL?

ETL é um tipo de **integração de dados com três etapas**, usado para combinação de dados de fontes diversas. A sigla ETL resume 3 processos de tratamento de dados:

* Extração;
* Transformação;
* Carregamento.

Portanto, entendemos por ETL os **softwares** que realizam a sistematização, **tratamento e limpeza de dados**. Esses dados podem ser extraídos de diversos sistemas para a inserção em um [Data Warehouse](https://www.significados.com.br/data-warehouse/) ou Data Mart.

Como funciona?

**Extração**

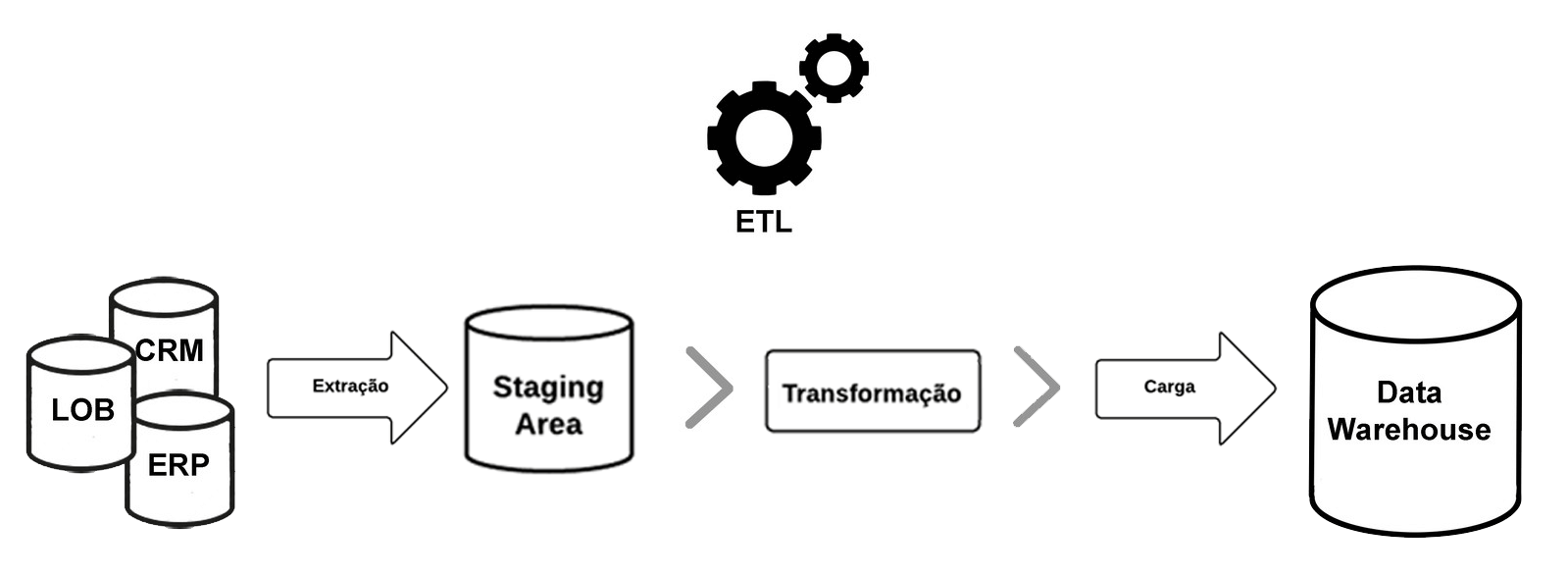
A primeira etapa é extrair os dados dos diversos sistemas de uma organização (OLTP). A partir disso, os dados são conduzidos para uma staging area, ou seja, área de transição. Assim os dados extraídos serão **convertidos para um único formato.**

**Transformação**

Em segundo lugar, a etapa de transformar envolve a limpeza dos dados. Nessa fase os dados extraídos serão tratados, padronizando e **corrigindo erros e inconsistências**. Dessa forma, os dados passam por uma transformação de acordo com as regras do negócio, para que **possam ser analisados**.

**Carregamento**

Assim que os dados são devidamente tratados na etapa de transformação, podemos partir para o carregamento. Então, nessa fase os dados são transmitidos, realizando a carga de dados no Data Warehouse.



Softwares ETL

Jasper

Talend

Apache Camel