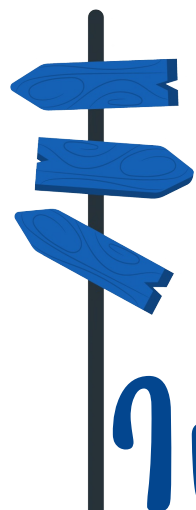




By @kakashi_copiador



Business Intelligence

Conceito

São técnicas, métodos e ferramentas de análise de dados p/ subsidiar processos de decisão de uma empresa.

Não substituem a intervenção humana. São apenas um apoio.

Habilidades

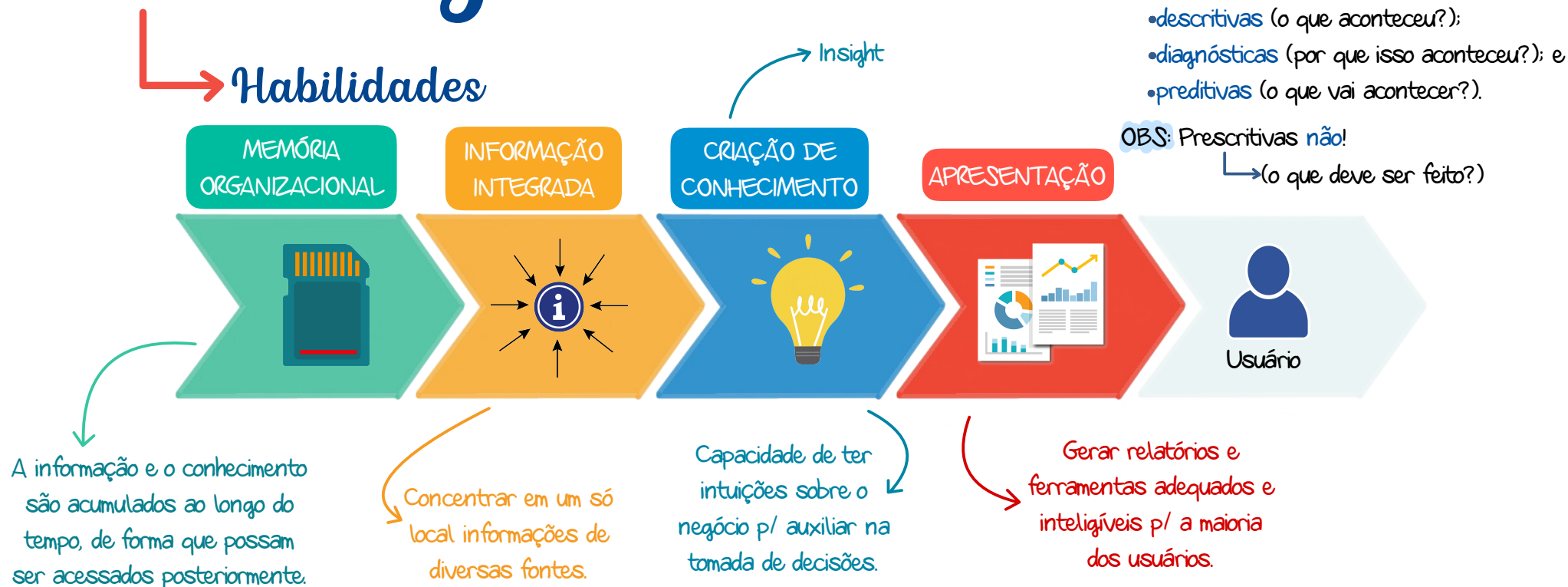
É realizado sobre dados históricos consolidados, provenientes de bases multidimensionais (como Data Warehouses).

É capaz de fornecer uma visão dos dados facilitando análises:

- descritivas (o que aconteceu?);
- diagnósticas (por que isso aconteceu?); e
- preditivas (o que vai acontecer?).

OBS: Prescritivas não!

↳ (o que deve ser feito?)



Conceito

- Data Warehouse = **armazém de dados**.
- Coleção de dados **orientados por assunto, integrados, variáveis com o tempo e não voláteis**, consolidados para dar suporte ao processo de tomada de decisão.

Características Essenciais

Orientado por assunto

- Trata de **temas específicos** e importantes p/ o negócio da organização.

Não-volátil

- Os dados, uma vez inseridos no DW, **não podem ser alterados**.
- **Somente leitura** p/ usuários finais (**consultas e relatórios**).

Integrado

- Os dados do DW são **integrados de múltiplas fontes** de forma consistente, padronizada e uniforme.

Histórico - variável com o tempo

- **Dados históricos** são mantidos em um DW.
- Em um banco de dados transacional, geralmente só os dados mais recentes são mantidos (ex: o último endereço de um cliente). Já o **DW pode conter todos os endereços já associados** a um cliente.

Espécies

Enterprise Data Warehouse (EDW)

- Armazém de Dados Corporativos ou Abrangentes.
- Repositório central **tático/estratégico**.
- Contém **todas as informações** de negócio de uma organização.
- Acessível **p/ toda a empresa**.

Operational Data Store (ODS)

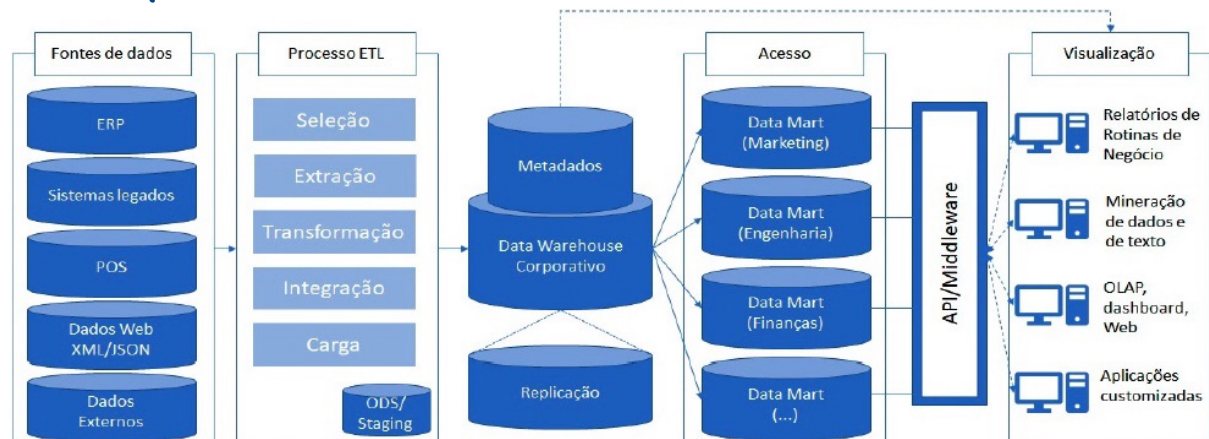
- Armazém de Dados Operacionais.
- Repositório **intermediário** de dados p/ suporte às **decisões operacionais**.
- Fonte de dados p/ o EDW.

Data Mart (DM)

- **Mercado de Dados**.
- **Subconjunto** de dados de um Data Warehouse.
- Contém dados **restritos a uma unidade** de negócio.
- Repositório **especializado**.

Data Warehouse

Componentes e Processo



Abordagem de Inmon

- Enterprise Data Warehouse.
- Abordagem **top-down** (de cima p/ baixo).
 - ↳ Construir o DW e dividi-lo em Data Marts.
- Orientação: **Orientado a dados**.
- Investimento Inicial **Alto**.
- **Normalizado** (3ª Forma Normal).
- Direcionado a **profissionais de TI**.
- Consultas realizadas no Data Mart.



Ambas possuem um único repositório integrado de dados corporativos que suporta a análise de informações em toda a organização.



Abordagem de Kimball

- Dimensional Data Warehouse.
- Abordagem **bottom-up** (de baixo p/ cima).
 - ↳ Construir vários Data Marts e depois integrá-los em um DW.
- Orientação: **Orientado a processos de negócio**.
- Investimento Inicial **Baixo**.
- **Desnormalizado** (não se importa com dados redundantes).
- Direcionado a **usuários finais**.
- Consultas realizadas no Data Warehouse.

Data Warehouse



ETL

Processo de **extrair** dados de fontes homogêneas e heterogêneas, **transformar** e **limpar** estes dados e **carregá-los** no DW.

A sigla ETL significa **Extract, Transform e Load**, mas o processo em si é mais complexo do que essas 3 etapas.

É o processo mais crítico e demorado na construção de um DW.

Dados trabalhados de forma errada podem gerar tomadas de decisões equivocadas na organização.



A etapa de extração é a mais demorada.

↳ Por isso é comum, quando um DW é criado, realizar uma **carga de dados inicial** e com o decorrer do tempo realizar **cargas incrementais**.

Stage Area é uma localização de **armazenamento temporário**, onde os dados permanecem até estarem prontos p/ serem carregados no DW.



P/ Kimball, a modelagem dimensional é a única técnica viável p/ bancos de dados que devem responder consultas em um DW.

