

Business Intelligence

Business Intelligence combina dados com ferramentas analíticas, de forma a disponibilizar informação relevante para a tomada de decisão. O objetivo é melhorar a disponibilidade e qualidade desta informação

O termo Business Intelligence tem substituído designações como a de sistemas de suporte à decisão, utilizada há mais de 40 anos.

Os sistemas de **Business Intelligence** combinam a recolha de dados, o armazenamento dos mesmos e a gestão de conhecimento com diversas ferramentas de análise que possibilitam a extração de informação útil, a partir dos dados armazenados. As suas tarefas são:

1. Elaborar previsões baseadas em dados históricos, nos desempenhos passados e atuais da organização
2. Criar cenários que evidenciem o impacto da alteração de diversas variáveis
3. Permitir o acesso ad hoc aos dados para responder a questões que não estão predefinidas
4. Analisar detalhadamente a organização, obtendo um conhecimento mais profundo da mesma.

Business Intelligence precisa de:

- Data Warehouse (onde os dados são armazenados) ->
- On-Line Analytical Processing (Técnicas de análise de dados) -> Informação e resultados
- Data mining (Técnicas de análise de dados) ->

Data warehouse

Um data warehouse consiste num conjunto de dados orientado por assunto, integrado, catalogado temporalmente e não volátil, que suporta os gestores no processo de tomada de decisão.

- Orientado por assunto

Num data warehouse, os dados são organizados em torno dos principais assuntos de uma organização, por exemplo, clientes, fornecedores, produtos, lojas, armazéns, etc. É um repositório que não está vocacionado para o armazenamento de informações associadas às operações do dia a dia da organização ponto visa análise de dados para o suporte à tomada de decisão

- Integrado

Um data warehouse é normalmente construído a partir de diversas fontes heterogéneas de dados por exemplo, a partir de diversas bases de dados de organização. Técnicas de limpeza e integração dos dados são utilizados para assegurar a consistência dos dados, exemplo ao nível dos nomes e significado dos atributos.

- Catalogado temporalmente

O objetivo data warehouse é fornecer informação sob uma perspetiva histórica. Estes sistemas armazenam dados que dizem respeito a um período de tempo que pode variar de 5 a 10 anos, enquanto os sistemas operacionais armazenam dados históricos limitados a horizonte temporal relativamente curto.

- Não volátil

Enquanto nos sistemas operacionais se realizam continuamente operações básicas sobre registos, tais como, consultas, inserções, atualizações e eliminações, no ambiente de um data warehouse existem apenas 2 operações carregamento inicial dos dados e acesso aos dados para processamento destes.

Base de dados

Objetivos operacionais
Acessos de leitura e de escrita
Acesso por transações predefinidas
Acessos a poucos registos de cada vez
Dados atualizados em tempo real
Estrutura otimizadas para atualizações

VS

Data Warehouse

Registo histórico
Acessos de leitura
Acessos por questões de ad hoc e relatórios
Acessos a muitos registos de cada vez
Carregamentos periódicos de mais dados
Estrutura otimizada para processamento

Data warehouse VS Data Marts

Um data warehouse contém informação relativa à organização no total e não parte dela. Sempre que o âmbito é mais restrito, por exemplo, departamental, o repositório de dados armazena um subconjunto específico de dados de organização, sendo designado por data marts e não por data warehouse

A organização deve procurar a arquitetura que melhor satisfaz as suas necessidades datawarehouse e o data marts. Pode optar por implementar um data warehouse organizacional, pela implementação de data marts independentes, ou, pela implementação de data marts dependentes do data warehouse organizacional

Dada a complexidade associada à concretização de um data warehouse **kimball e Ross** apresentam o ciclo de vida dimensional de negócio de desenvolvimento de projetos de data warehousing. Esta permite adicionar valor ao negócio e disponibilizar dados para o suporte à tomada de decisão.

De uma forma geral pode ser referido que esta metodologia assente em 4 princípios:

1. Foco no negócio
2. Foco nos dados
3. Abordagem incremental
4. Solução integrada

Identificados os requisitos do negócio, 3 atividades devem ser realizadas em paralelo:

1. a definição da arquitetura tecnológica da solução a desenvolver
2. a modelação multidimensional do data warehousing
3. e a propostas das aplicações de Business Intelligent, que servem de interface com os utilizadores.

Modelação multidimensional

É utilizada para projetar a estrutura de sistemas de data warehousing. É baseado em 2 pressupostos:

Produzir uma estrutura da base de dados fácil de compreender
E utilizar e otimizar o desempenho o processamento de questões

- Esquema em estrela

A **tabela de factos** corresponde ao assunto que se pretende analisar normalmente uma componente de negócio encomendas, vendas, compras, etc

Integra um conjunto de atributos numéricos e um conjunto de Chaves estrangeiras que relacionam a tabela de facto com as diversas dimensões que lhe estão associadas

Apresenta-se normalizada

Contém uma grande quantidade de registos ocupa normalmente 95% do espaço da data warehouse

As **tabelas de dimensão** possibilitam a análise da tabela de factos sobre diferentes perspectivas, permitindo responder a diversas questões, tais como: quem? Quando? Onde? Porquê? Como?

Existem tantas dimensões quantas vertentes pelas quais se pretende analisar os factos

Apresentam-se não normalizadas integrando na maioria dos casos uma grande quantidade de atributos

Contém poucos registos quando comparados com a tabela de factos

- Floco de neve

O esquema em floco de neve inclui a vantagem de explicitamente indicar a estrutura de cada uma das suas dimensões, ao contrário do esquema em estrela, onde as dimensões integram um conjunto não estruturado de atributos.

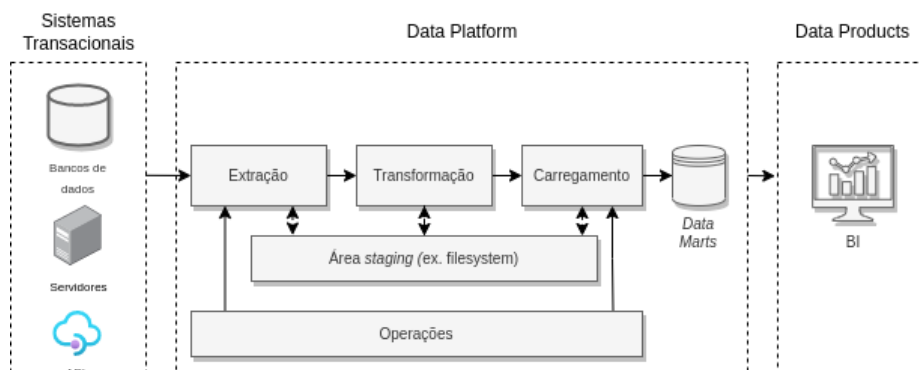
- Constelação

É um esquema que integra múltiplas tabelas de factos que partilham dimensões comuns.

ETL - Carregamento e refrescamento to um Data Warehouse

As ferramentas de Extração, Transformação e Carregamento (Extraction, Transformation, Loading – ETL) que permitem tratar da homogenização do dados, fazer a sua limpeza e o respetivo carregamento para o Data Warehouse.

As Ferramentas de ETL fazem a ponte entre os sistemas operacionais, sistemas-fontes de origem dos dados e o data warehouse que será posteriormente acedido através de aplicações próprias de Business Intelligence



Extração de dados, recolhendo dados de múltiplas fontes, que podem ser heterogéneas ou externas à organização

Transformação de dados, convertendo os dados do seu formato inicial para o formato em que irão ser armazenados no Data Waterhouse

Limpeza de dados, identificando erros nos dados e fazendo a sua correção (sempre que possível, como referido anteriormente)

Carregamento, transportando os dados para o Data Warehouse. O carregamento inclui tarefas de ordenação, agregação, consolidação e verificação da integridade dos dados, entre outras

Refrescamento, através do carregamento para o Data Warerhouse dos novos dados existentes nas fontes de dados, atualizando-o.

Octave

- Limpar: janela de comandos e o histórico e apagar qualquer editor com variáveis anteriores
- Criar variáveis numéricas:
 - 1D [1,1]
 - 1D [1,5] - introduz Manualmente , através de 1 ciclo for
 - 2D [300,500]
 - Operações aritem + x com variáveis 1D e 2D
- Concatenar variáveis com diferentes dimensões
- Várias variáveis Char
- Concatenar as variáveis Char
- 3D [150,250,100]