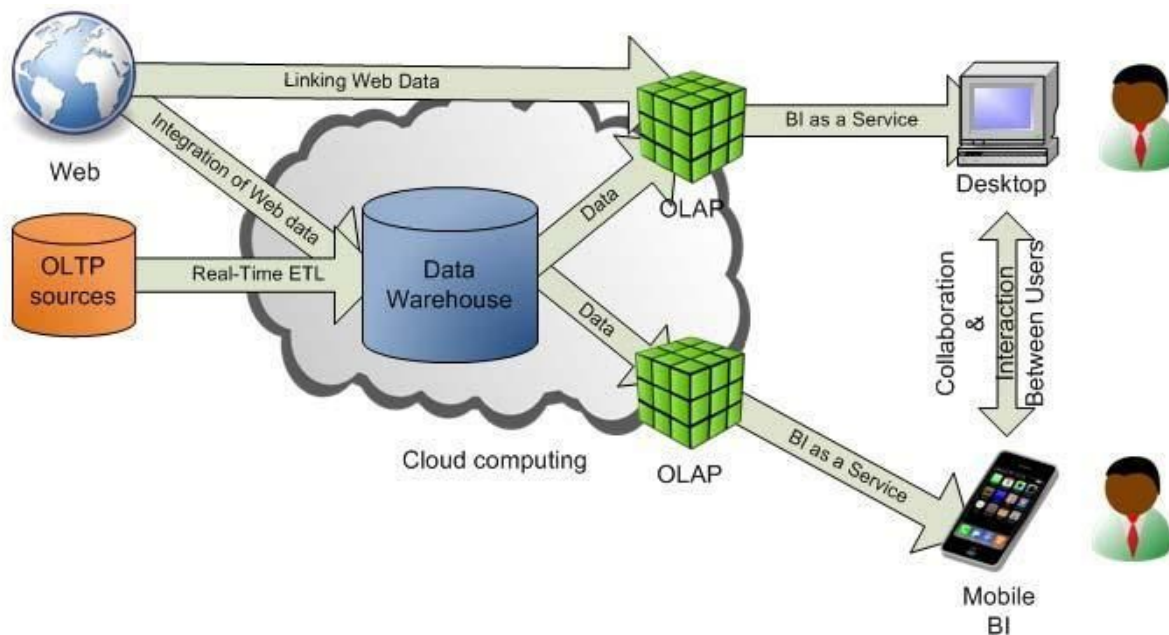


Componentes do Business Intelligence

Os componentes de Business Intelligence são: fontes de dados, Data Warehouse (DW), *Online Analytical Processing* (OLAP), Data Mining, Extract, Transform, Load (ETL) e ferramentas de visualização.



Fontes de dados

A produção de dados está em todos os lugares, dentro e fora das companhias. Muitas soluções de BI já conseguem capturar informações de boa parte delas — incluindo redes sociais, relatórios públicos do governo, e-mails e outros — mesmo quando eles são do tipo não-estruturado.

Uma característica fundamental de um processo de BI é que nele os dados são copiados da base de dados transacional e de outras fontes, para a base de dados analítica, o que permite que as informações sejam extraídas desta última sem que a performance do sistema transacional seja prejudicada e da forma mais eficiente, eficaz e visualmente agradável possível, sempre focada na tomada de decisão.

Os dados podem, na sua origem, estar estruturados ou não, ou seja, podem vir de um sistema de informação com dados organizados e corretamente normalizados em tabelas e colunas, e podem ainda serem extraídos de sites da internet como texto puro, áudio, vídeo ou outros formatos. Apesar disso, em ambas as situações eles devem ser tratados e gravados em um formato que facilite a extração de informações e apoie as decisões dos gestores.

A grande descentralização e heterogeneidade dos dados é o grande desafio do momento.

Vencer o desafio reside em integrá-los, interpretá-los e transformá-los, de alguma forma, em informação relevante ao negócio, possibilitando, com a devida análise, a criação de conhecimento.

Este conhecimento pode ser o diferencial de uma empresa, pois possibilita ressaltar os seus pontos fortes e mitigar os riscos envolvidos nos pontos fracos. Tendo um retrato fiel da realidade, uma empresa consegue, com base nos dados históricos, fazer uma previsão bastante assertiva do futuro e utilizá-la como embasamento para suas decisões.

Algumas fontes de dados:

Web: abordagens da web com conteúdo dinâmicos e assíncronos que possibilitam os usuários interagirem com a página e com outros usuários, ex: Redes Sociais.

OLTP: (*Online Transaction Processing*) A maioria dos sistemas de informação se enquadra na categoria dos Sistemas (OLTP), também chamada de processamento de transações em tempo real. Um sistema deste tipo exige uma base de dados modelada para otimizar a inclusão e alteração de dados, obedecendo a padrões rígidos de normalização, evitando redundâncias, permitindo a integridade referencial e outras consistências. São sistemas que registram as transações, como, por exemplo, os ERPs (Enterprise Resource Planning).

Cloud Computing: dados disponíveis na computação nas nuvens que tem como base a utilização da memória e das capacidades de armazenamento, compartilhados e interligados por meio da Internet.

Linking Web Data: Refere-se a dados ligados descrevendo um método de publicação de dados estruturados que podem ser interligados baseando-se em tecnologias da web estendendo-se para compartilhar informações que podem ser lidos automaticamente por computadores.

Integration of Web Data: Refere-se à integração de dados, essa integração combina dados residentes em diferentes fontes, tendo em vista fornecer aos usuários uma visão única dos dados.

BI as a Service: Refere-se à arquitetura SOA, orientado a serviços, o que facilita a partilha e acesso de informações em tempo real.

ETL: São as ferramentas usadas para transferir os dados, já prontos para uso, para o DW. Para a extração, especifica-se a fonte e quais dados se quer. Em seguida, eles vão para uma área em que podem ser transformados e tratados (para evitar, por exemplo, múltiplas instâncias do mesmo registro). Por fim, devidamente organizados, são enviados para o DW.

- **Extraction**(extração): Consiste na leitura dos dados a partir de uma ou mais fontes, como, na obtenção de todos os registros adicionados ou alterados depois do último processo de ETL executado, na leitura de páginas da web, de web services, xml, xls, csv, json, além de formatos Big Data, como Hadoop, Cassandra, MongoDB e HBase.
- **Transformation** (transformação): Consiste na alteração da forma ou do conteúdo do dado, para que se encaixe a estrutura da base destino. A transformação ocorre, por exemplo, ao se fazer um lookup pelo nome da cidade para armazenar o seu código, ao ajustar o nome do

bairro a partir do seu CEP, ao invalidar um registro por ter CPF ou e-mail inválidos, ao fazer decodificações, renomeações, cálculos, agregações, etc.

- **Load** (carga): Consiste na gravação dos dados em seu destino, independente do seu formato, que pode ser em um arquivo, em uma tabela da base transacional, em uma dimensão ou tabela fato de um cubo ou em formato Big Data.

Real Time ETL: Processo de extração, transformação e carga de dados em tempo real. Basicamente falando é uma forma de integrar os dados em tempo real, podendo ser feitos por intervalos de lotes em curto espaço de tempo ou dependendo do segmento feito apenas algumas vezes ao dia.

Data Warehouse: O DW permite acesso fácil a todos os dados da companhia, vindos das diferentes aplicações instaladas na empresa — e previamente tratados pelo ETL. São, normalmente, estruturados em tabelas relacionais que podem ser visualizadas, mas não alteradas (isso só é possível na aplicação-fonte).

Quando o DW é alimentado com dados compatíveis e sincronizados, passamos a ter todos os dados da empresa em um único local e em um formato consistente. Com isso, é possível ter uma visão padronizada, única e holística do que está acontecendo na empresa. O volume de dados, porém, continua sendo um desafio.

O Data Warehouse constitui-se em um grande banco de dados ou armazém de dados, responsável pelo armazenamento de altos volumes de dados, ou seja, os dados brutos de uma organização.

Data Mart: Repositório de dados, subdivisão de um DW, também responsável pelo armazenamento de dados, porém mais específicos e em menor escala.

OLAP: O OLAP ajuda a obter respostas mais rapidamente. Isso é possível porque o OLAP calcula previamente e guarda respostas às possíveis questões dos usuários. Com isso, é possível questionar os dados de forma abrangente e, assim, obter uma visão de alto nível enquanto os inspeciona até o nível do registro individual.

Com respostas mais rápidas, todo o método de análise muda: a informação é extraída mais facilmente dos dados e isso pode garantir vantagens competitivas à organização por meio de cubos decisórios, local onde os dados são analisados e processados gerando informações essenciais ao negócio.

Data Mining: É o processo de extração de informação dos dados corporativos, em busca de correlações, anomalias e tendências. Para que ofereça as respostas corretas, é preciso que sejam feitas as perguntas adequadas. Profissionais especializados podem ajudar a determinar quais resultados são significativos.

Visualização: O BI transforma dados, que no começo eram descontextualizados, em informação relevante para o usuário final, que pode ser ou não técnica. É preciso, portanto, que esses dados sejam didáticos, claros e intuitivos, que possam ser facilmente compreendidos.

Em geral, as ferramentas de visualização de dados utilizam gráficos e mapas para auxiliar esta compreensão. Os formatos tradicionais incluem gráficos de barra, de pizza, radar, funil, chegando até a relatórios e painéis interativos.

Dashboards: Trata-se das primeiras interfaces gráficas com as quais o usuário tem contato. Eles concentram os relatórios e as informações mais importantes, personalizados de acordo com as necessidades do usuário. É neles que as informações mais importantes são organizadas de forma simples e em um único local.

Os dashboards fazem parte da última etapa do ciclo de BI é a da saída dos dados, que pode acontecer em diversos formatos. Os relatórios, o formato mais usual, têm layout pré-definido, aceitam uma gama de filtros e geram documentos em pdf, xls, etc.

Os dashboards têm a finalidade principal de permitir o acompanhamento de indicadores em tela, de forma gráfica e interativa. Já as análises são feitas por meio de uma ferramenta de navegação OLAP, pelo próprio usuário final, que pode manipular os cubos criados, além de salvar e compartilhar suas consultas analíticas.

E-mails e Posts: Outras formas usuais de saída são os E-mails e Posts, que podem ser disparados para sinalizar a ocorrência de um evento ou para alertar que determinado indicador atingiu um nível crítico.

