

SISTEMAS DE INFORMAÇÃO II

CAPÍTULO 1.1

SISTEMAS DE DATA WAREHOUSING: INFRA-ESTRUTURA INFORMACIONAL E FUNCIONAL

FILIPPE SÁ
FILIPPE.SA@ISEC.PT

1. Sistemas de Data Warehousing: Infra-estrutura Informacional e Funcional

Tópicos Programáticos:

1.2. Ambiente e estrutura funcional de um data warehouse

1.2.1. Arquitectura generica um Data Warehouse

1.2.1.1 Arquitectura de um Data Warehouse – Kimball

1.2.1.2 Arquitectura de um Data Warehouse - Inmon

1.2.2 Seleção de Arquitectura (Kimball Vs Inmon

1.2.3 Arquitetura de Chaudhuri

1.2.4 Arquitetura de Valente

1.2.5. Classificação de Data Warehouses

1.3. Ciclo de vida e desenvolvimento incremental de um data warehouse

1.3.1. Evolução dum DW

1.3.2. Ciclo de Desenvolvimento

1.3.3. Monitorização e Optimização

1.4.1. Business Intelligence

1.4.1.1 O que é o BI

1.4.1.2 Business Intelligence Loop

1.4.1.3 Infraestruturas tecnologicas de Business Intelligence

1.4.1.4 Exemplos de Interrogações típicas BI

1.4.1.5 BI no contexto das organizações

1.4.1.6 Funcionalidades de BI

1.4.1.7 Principais softwares de BI

1.4.1.4 Principais softwares de BI

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1. Arquitectura de um Data Warehouse

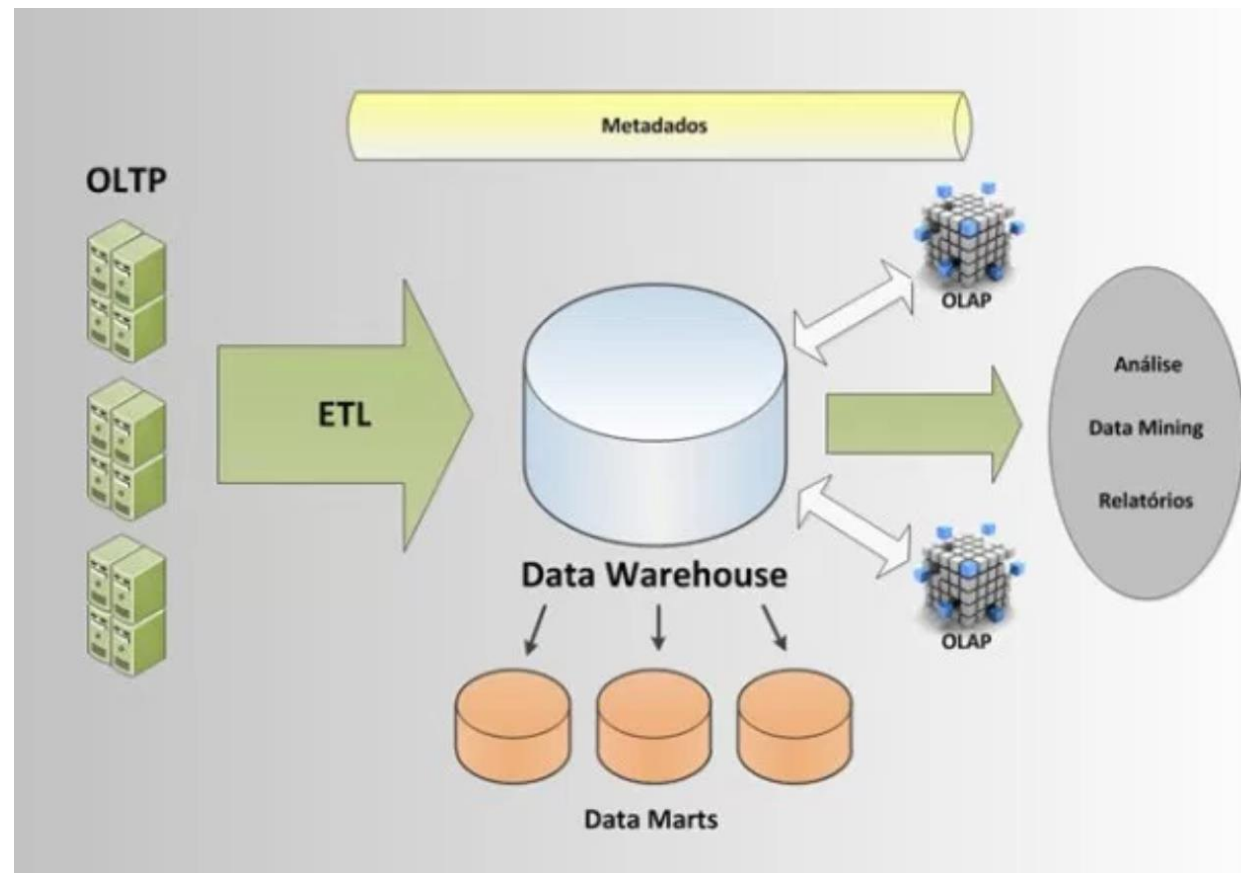
Uma estrutura de uma DW pode variar de acordo com aquilo que o integra;

Podemos falar de:

- DW com ODS (Operational Data Store), que não é mais que uma base de dados desenhada para integrar dados de múltiplas fontes em que estes são tratados antes de serem enviados ao DW para armazenamento;
- DW com múltiplos data marts, que são um subconjunto de dados de DW referentes a determinado assunto ou departamento.

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1. Arquitectura Generica um Data Warehouse



Fonte Magalhães (2017)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1. Arquitectura Generica de um Data Warehouse

No entanto, todos os DW são compostos por vários níveis, cujos principais são os seguintes:

1. Nível de fonte de dados (Data Source)
2. Nível de extração de dados (Data Extraction)
3. Nível de estagio de dados (Data Staging)
4. Nível de armazenamento (Data Storage)
5. Nível de apresentação de dados (Data Presentation)

Fonte Magalhães (2017)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1. Arquitectura Generica de um Data Warehouse

1. Nível de fonte de dados (*Data Source*) – Todas as fontes de dados originais que vão alimentar o DW, como BD's, ficheiros de texto, excel, etc;
2. Nível de extração de dados (*Data Extraction*) – Os dados são importados a partir das fontes de dados para o DW, podendo existir filtragem nos dados, mas nada significativo;
3. Nível de estagio de dados (*Data Staging*) – Primeiro armazenamento dos dados e conjunto de processos que vão iniciar o tratamento dos dados;
4. Nível de armazenamento (*Data Storage*) – Área onde os dados são armazenados após a execução de todos os processos de filtragem e transformação;
5. Nível de Apresentação de dados (*Data Presentation*) – Informação que é apresentada, de várias formas, aos utilizadores. Neste nível, são geralmente usadas ferramentas *Online Analytical Processing* (OLAP) e de geração de relatórios.

Fonte Magalhães (2017)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitectura de um Data Warehouse - Kimball

Arquitectura Kimball

Apresentada por Ralph Kimball;

Composta por quatro componentes separados e distintos necessários à criação de um DW

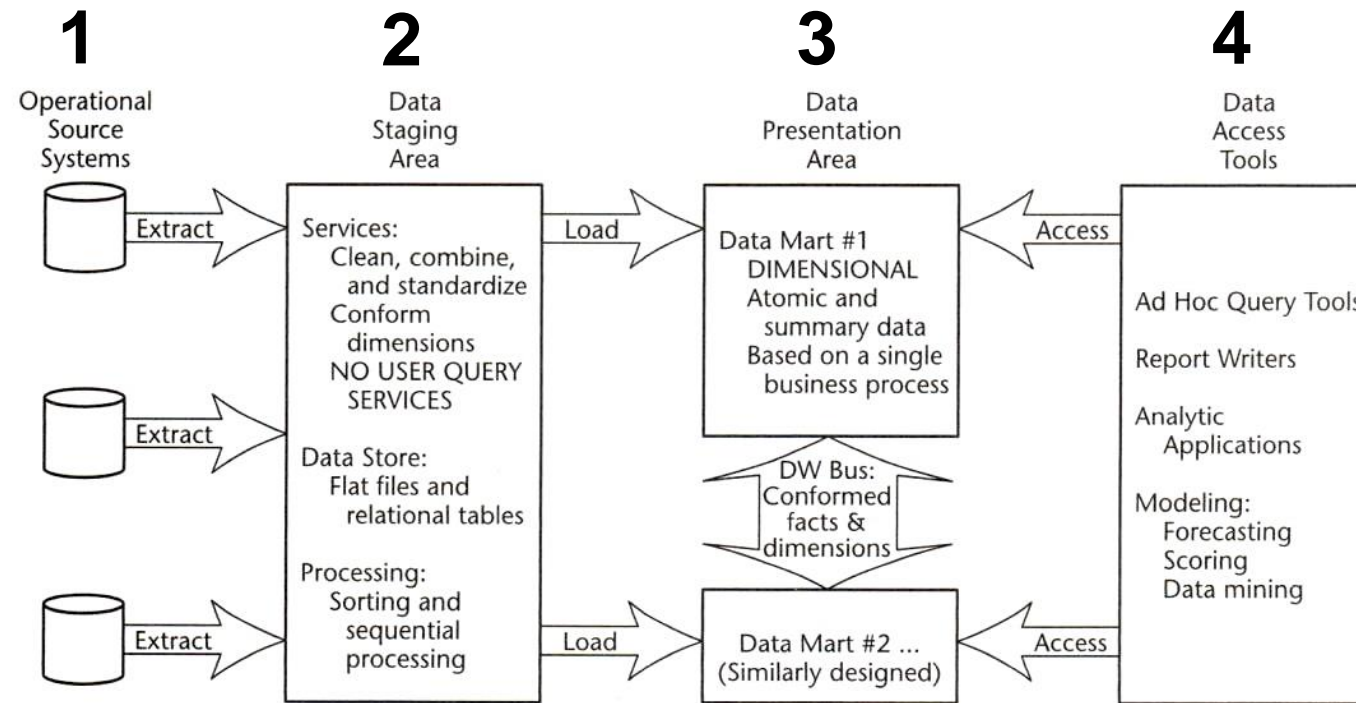
1. Sistemas Operacionais de origem
2. Data Staging (área de estágio de dados)
3. Data Presentation (área de apresentação de dados)
4. Ferramentas de aplicações de dados

Fonte Magalhães 2017

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitetura de um Data Warehouse - Kimball

Elementos Básicos dum Data Warehouse - Kimball



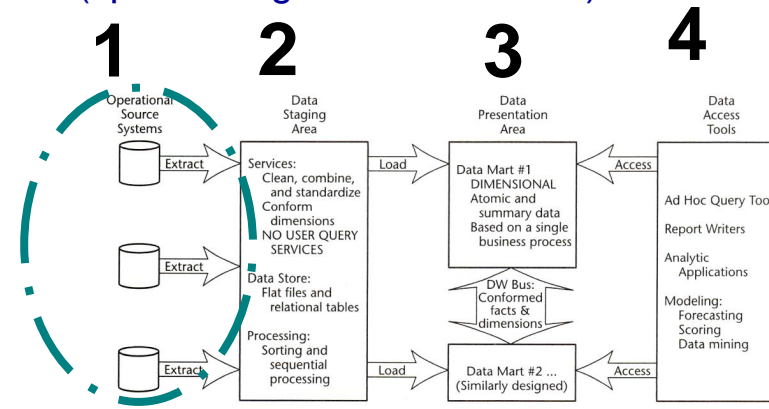
Retirado de: *The data Warehouse Toolkit*, Ralph Kimball

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitectura de um Data Warehouse - Kimball

Elementos Básicos Arquitetura DW Kimball: 1-Sist. Operacionais Origem

- São estes sistemas que alimentam o DW com dados
- O utilizador, em princípio, não tem qualquer controlo sobre eles, pelo que poderão ser considerados externos ao DW;
- Registam as transacções do negócio. Podem ser pensados como exteriores ao DW, já que, muitas vezes, temos pouco controlo sobre o conteúdo ou formato dos dados;
- A prioridade dos sistemas operacionais é o desempenho transaccional e disponibilidade;
- As interrogações típicas nestes sistemas são estreitas (tipo um registo de cada vez) e são parte do normal fluxo transaccional;
- Os sistemas operacionais mantêm poucos dados históricos e, havendo um DW, não são, de todo, responsáveis pelo registo do passado;
- Normalmente estes sistemas são não integrados, com todos os problemas falados atrás.



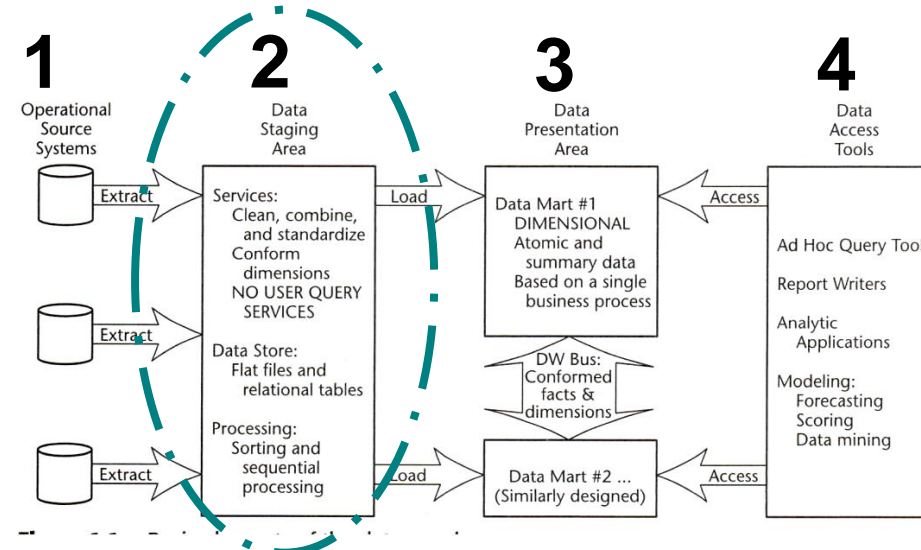
1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitectura de um Data Warehouse - Kimball

Elementos Básicos Arquitetura DW Kimball:

2-Área de Estágio dos Dados

- É simultaneamente uma área de armazenamento e um conjunto de processos normalmente referidos como extracção-transformação-carga (Extract-Transform-Load:ETL);
- É tudo o que existe entre os sistemas fonte e a área de apresentação;
- É território proibido dos utilizadores finais e não proporciona serviços de interrogação e apresentação. Ou seja não é acessível ao utilizador e não permite consultas ou visualizações;
- A extracção é o primeiro passo do processo de pesquisa de dados para o DW;
- Segue-se a transformação (purificação, combinação e atribuição de chaves);
- Área dominada por actividades simples como processamento sequencial e ordenações;
- O passo final passa por estruturar fisicamente os dados e carregá-los na área de apresentação.

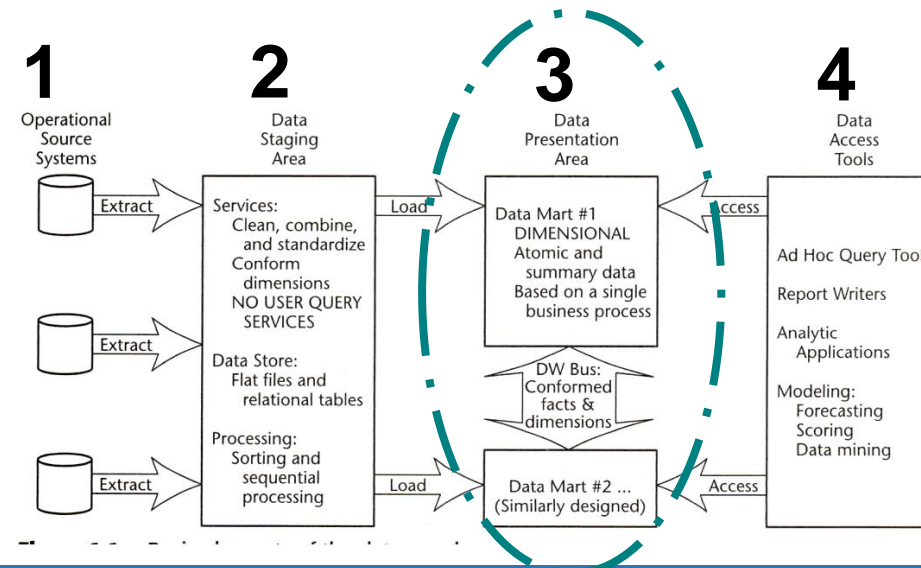


1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitectura de um Data Warehouse - Kimball

Elementos Básicos Arquitectura DW Kimball: 3-Área de Apresentação

- É onde os dados são organizados, armazenados e disponibilizados para interrogações directas ou outras aplicações analíticas, por parte dos utilizadores;
- É o DW que os utilizadores conhecem, sendo tudo o que eles vêem e tocam através de ferramentas de acesso;
- O epíteto “colocar os dados na saída” seria adequado para esta área;
- Modelada dimensionalmente;
- Podemos pensar esta área como uma série de data marts integrados;
- Um data mart é uma fatia do total da área de apresentação, representando os dados de um processo de negócio;
- A modelação dimensional resulta da forma intuitiva de pensar o negócio como um cubo de dados;
- Foco: simplicidade -> rapidez.

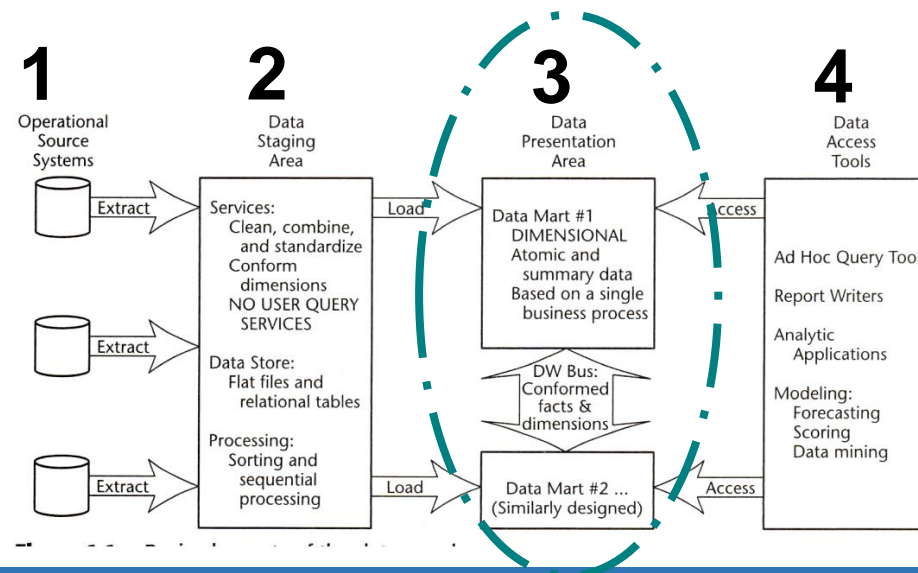


1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitectura de um Data Warehouse - Kimball

Elementos Básicos Arquitetura DW Kimball: 3-Área de Apresentação

- Pode conter dados detalhe (atômicos) e agregados;
- Os dados de detalhe são necessários para responder a interrogações não previsíveis (ad hoc);
- Enquanto os Data Marts podem conter dados agregados que permitem melhorar o desempenho, os dados de detalhe devem ser armazenados, usando o mesmo modelo dimensional;
- A granularidade determina o nível de detalhe máximo que é permitido nas consultas;
- Em resumo, Kimball refere que: “Na área de apresentação, os dados disponíveis para interrogações devem ser modelados dimensionalmente, atômicos, e conformes com a arquitetura de bus do Data Warehouse (utilização de dimensões e factos - esquemas em estrela relacionais ou cubos em BD multidimensionais)”

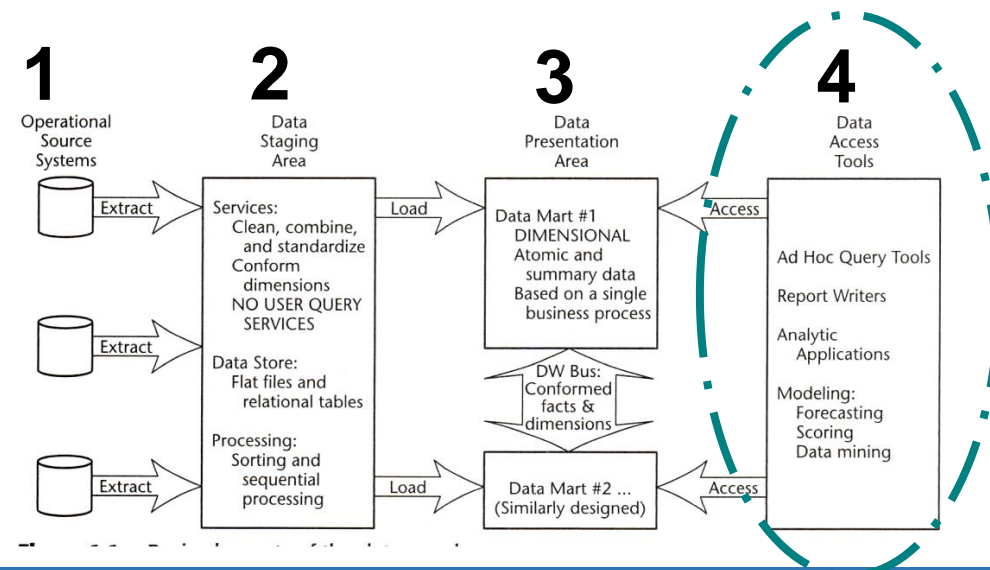


1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.1 Arquitetura de um Data Warehouse - Kimball

Elementos Básicos Arquitetura DW Kimball: 4-Ferramentas de Acesso

- Por definição, executam as interrogações à área de apresentação do Data Warehouse;
- Proporcionam um conjunto de capacidades que permitem explorar a área de apresentação para a tomada de decisões;
- Podem ser tão simples como ferramentas para colocar interrogações had hoc;
- Podem ser complexas como aplicações de data mining ou modelação;
- Aproximadamente 80-90% dos utilizadores acedem ao DW através de aplicações analíticas pré-construídas;
- 10-20% dos utilizadores farão uso de ferramentas de interrogação ad-hoc



1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.1.2 Arquitectura de um Data Warehouse - Inmon

Arquitectura de Inmon

Bill Inmon desenvolveu a arquitetura Corporate Information Factory (CIF)

Conhecido como o 'pai do DW'

Os dados são extraídos da fonte de dados e processados num sistema ETL por aquisição de dados;

Deste processo resultam dados atómicos que são armazenados num DW empresarial;

Este constitui o elemento central da arquitetura de Inmon, um repositório de dados atómicos que são capturados a um nível muito baixo de detalhe.

Os dados são guardados numa base de dados relacional com normalização à terceira forma normal.

Um particularidade desta arquitetura é o facto de além de serem utilizados na criação dos data marts, os dados podem ser utilizados na pesquisa dos utilizadores;

Os dados Data Mart são centrados no departamento, o que fará ter um data mart por sector.

1.2. Ambiente e estrutura funcional de um Data Warehouse

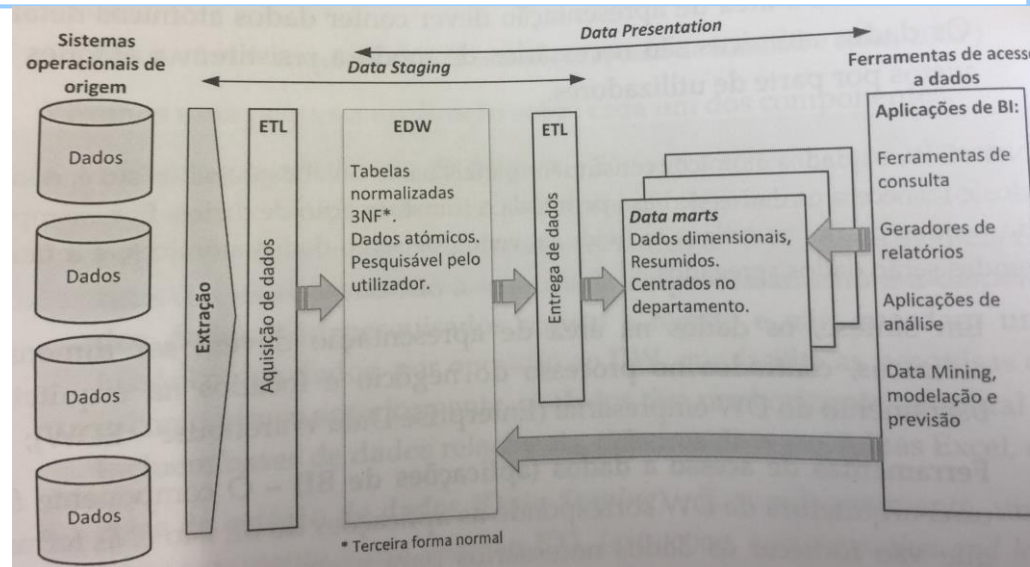
1.2.1.2 Arquitectura de um Data Warehouse - Inmon

Arquitectura de Inmon

As ferramentas de acesso aos dados são as já referidas na arquitetura de Kimball

Uma particularidade de Inmon é sobreposição da Data Presentation com a Data Staging

Permite pesquisas no EDW – Enterprise Data Warehouse



Fonte Magalhães (2017)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.2. Selecção de Arquitetura

Kimball Vs Inmon

As duas arquiteturas partilham um aspecto: Têm um armazenamento integrado e único de dados atómicos

- Kimball denominou-a Dimensional Data Warehouse
- Inmon denominou-a Enterprise Data Warehouse
- Em fases iniciais, enquanto Inmon aposta na criação de um DW que suporta toda as áreas de uma organização, Kimball aponta para a criação de Data Marts (DM)

Kimball utiliza:

- modelos dimensionais (esquema em estrela, floco de neve) para organizar os dados numa DW dimensional.
- os sistemas analíticos podem aceder diretamente aos dados.
- Aproximação designada por bottom-up, no qual, os data marts são criados de modo a servirem as necessidades analíticas dos departamentos (podem ser combinados num grande DW)
- O modelo dimensional é utilizado para suprir as necessidades dos departamento de uma empresa

Inmon utiliza:

- Abordagem top-down
- Utiliza o modelo ER no Enterprise Data Warehouse.
- Ao contrario de Kinball, Inmon recorre ao modelo dimensional somente para os data marts
- Utiliza os data marts dedicados aos departamentos
- Começa por criar um EDW e só depois os data marts de acordo com as necessidades analíticas dos departamentos

Fonte Magalhães (2017)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.2. Seleção de Arquitetura

Kimball Vs Inmon

Tabela 1 – Diferenças nas Metodologias Inmon e Kimball. Fonte: George, 2012 (Adaptado)

	Inmon	Kimball
Ferramentas dimensionais		✓
Orientada aos processos		✓
Facilidade de <i>design</i>		✓
Ferramentas relacionais	✓	
Dados normalizados	✓	
Análise temporal contínua e discreta	✓	

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.2. Selecção de Arquitetura

Kimball Vs Inmon

Pros e Contras das Metodologias. Fonte: George, 2012 (Adaptado)

	Inmon	Kimball
Construção do DW	Dispendiosa em tempo	Simples e rápida
Manutenção	Fácil	Difícil e por vezes redundante
Custo	Alto inicialmente, com custos decrescentes	Baixo, com custos que se mantêm
Duração	Longo tempo de montagem	Curto tempo de montagem
Conhecimento Exigido	Equipa especialista	Equipa generalista

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitectura Chaudhuri

Arquitectura de Chaudhuri

Além de conhecer os componentes envolvidos na construção do DW é necessário compreender os fluxos de dados que ocorrem no sistema

Chaudhuri propõe uma arquitetura, com base no fluxo dos dados

Esta arquitetura pode ser dividida em:

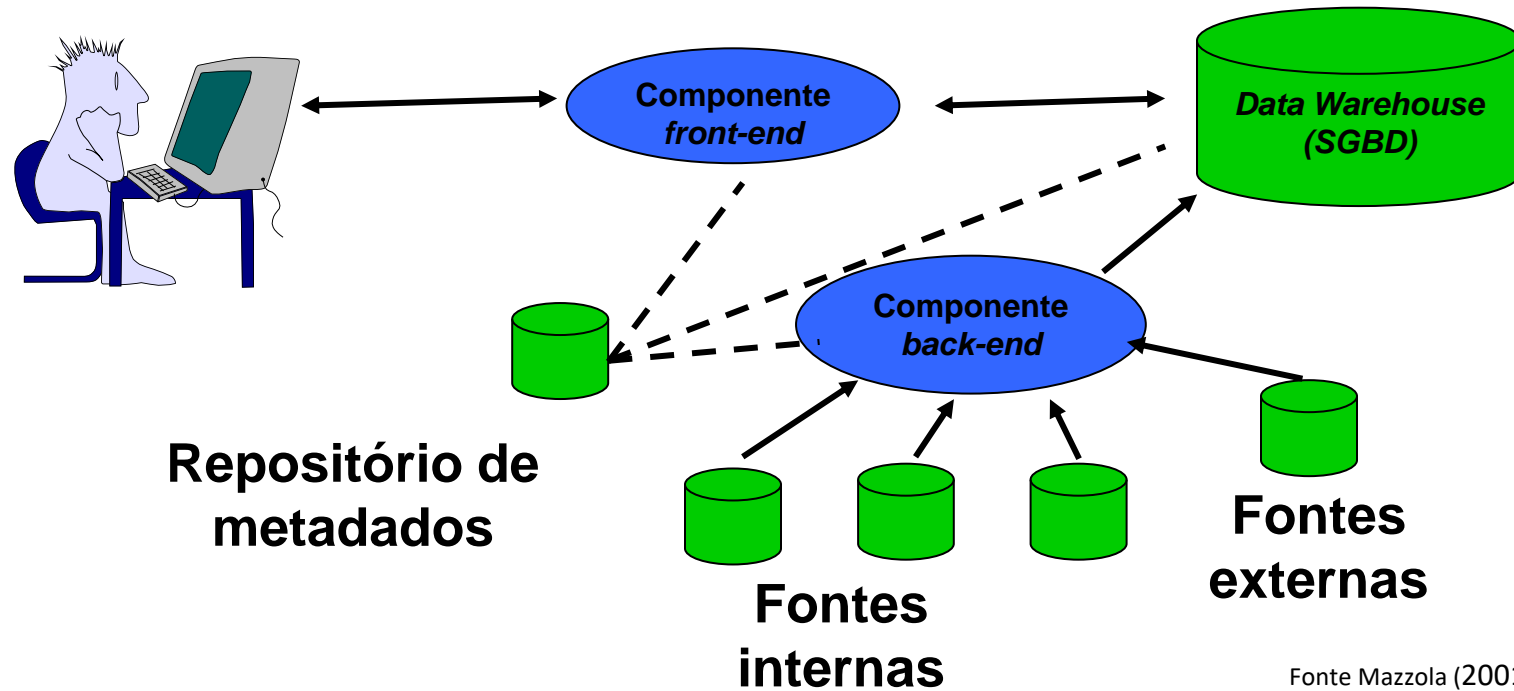
1. Fontes de dados de onde o DW irá retirar os seus dados de origem;
2. Um conjunto de estruturas de dados armazenados analiticamente: o DW do sistema;
3. Um Sistema Gestão de Base de Dados (SGBD) otimizado para atender os requisitos analíticos dos sistemas de DW;
4. Um componente back end: conjunto de aplicações responsáveis por extrair, filtrar, transformar, integrar e carregar os dados de diferentes origens no DW;
5. Um componente frontend: conjunto de aplicações responsáveis por disponibilizar aos utilizadores finais acesso à DW;
6. Um repositório para armazenar e gerir os metadados do sistema.

Fonte Mazzola (2001)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitetura de um Data Warehouse - Chaudhuri

Arquitetura de Chaudhuri



1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitetura de um Data Warehouse - Chaudhuri

Arquitetura de Chaudhuri

Cinco principais fluxos fazem parte do sistema:

1. fluxo de entrada (inflow')
2. fluxo de saída (outflow)
3. fluxo de subida (upflow)
4. fluxo de descida (downflow)
5. metafluxo (metaflow).

Fonte Mazzola (2001)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitectura de um Data Warehouse - Chaudhuri

Arquitectura de Chaudhuri

1 - O primeiro fluxo é o de entrada dos dados no sistema (inflow), que envolve extrair, filtrar, transformar, integrar e carregar os dados de várias fontes no DW.

Deve-se considerar as fontes de dados que pertencem à empresa e as fontes externas.

O fluxo de entrada é geralmente implementado com ajuda de ferramentas especialmente desenvolvidas para este fim.

2 - O segundo fluxo é o de descida dos dados (dowflow), ou seja, em tempos pré determinados, de dois a cinco anos dependendo da empresa, os dados armazenados no DW passam para o estado de dados antigos. Este é o fluxo que remove do DW aqueles dados considerados velhos, que não são utilizados com regularidade.

Fonte Mazzola (2001)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitectura de um Data Warehouse - Chaudhuri

Arquitectura de Chaudhuri

3 – O terceiro fluxo é o de subida dos dados (upflow), onde é dado o foco à necessidade de colocar os dados em formatos mais acessíveis para os utilizadores finais.

Este processo sumariza e agrupa os dados dentro de “vistas” mais adequadas aos utilizadores finais e às aplicações de frontend (tabelas sumarizadas, gráficos, páginas HTML, base de dados pessoais, entre outros formatos).

Também é função do fluxo de subida a distribuição dos dados para os diferentes níveis do sistema como, por exemplo, Data Marts e BD's pessoais.

4 – O quarto fluxo é o de saída dos dados (outflow), cuja função é disponibilizar acesso aos utilizadores finais. Este processo é implementado através de uma variedade de ferramentas de frontend como, por exemplo, ferramentas com características OLAP, estatísticas, Data Mining, Executive Information System, Decision Support Systems (DSS), entre outras.

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitectura de um Data Warehouse - Chaudhuri

Arquitectura de Chaudhuri

5 - O quinto e último fluxo pode ser designado de metafluxo (metaflow)

Ao contrário dos quatro fluxos de dados anteriores, que descrevem como os dados percorrem o DW o metafluxo move metadados, ou seja, dados sobre os outros fluxos.

O repositório de metadados é responsável pela gestão do sistema como um todo, indicando:

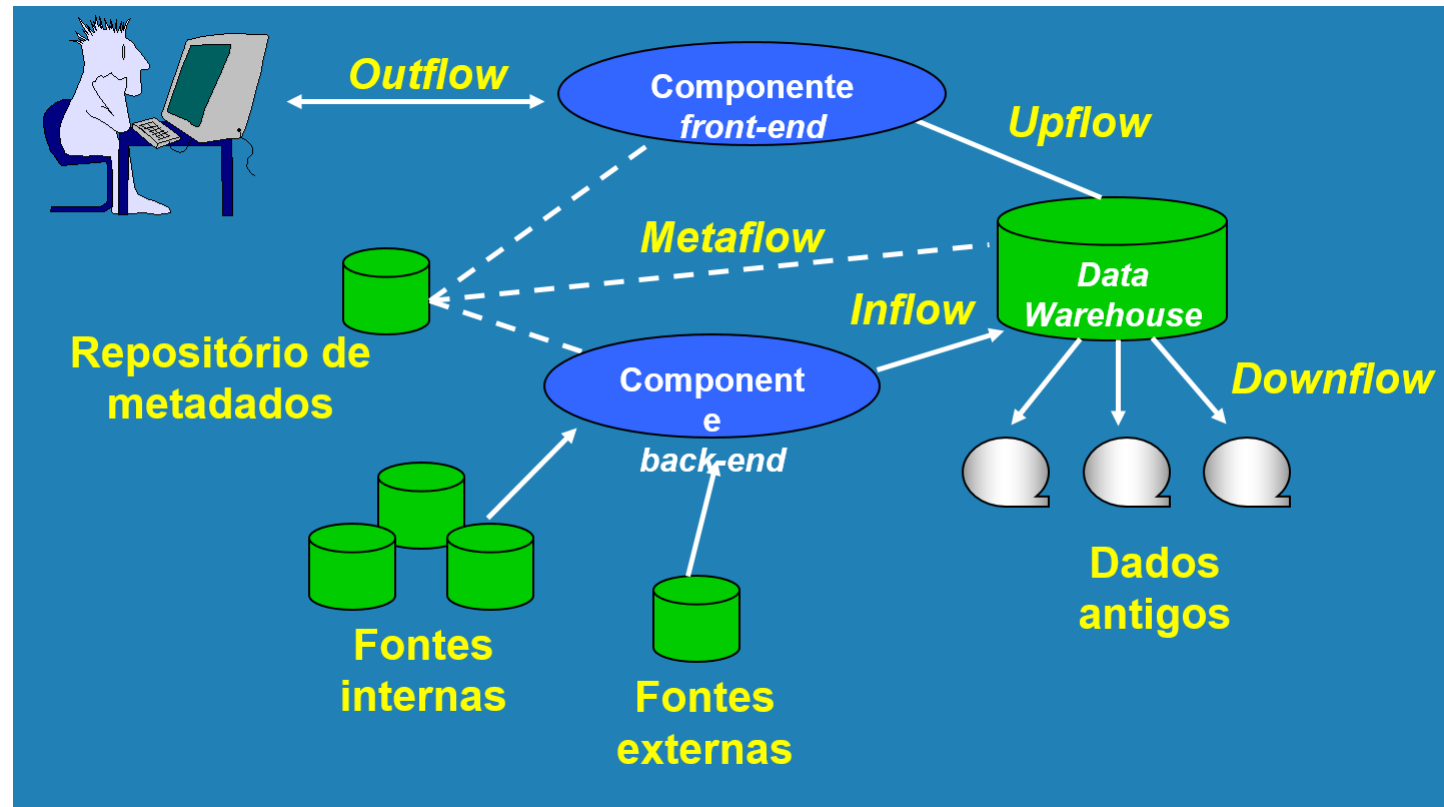
- a origem dos dados
- como são transformados
- quando são atualizados
- o seu significado
- como são acedidos
- quem os visualiza.

Fonte Mazzola (2001)

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.3 Arquitectura de um Data Warehouse - Chaudhuri

Arquitectura de Chaudhuri



1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.4 Arquitectura de um Data Warehouse - Valente

Arquitectura de Valente

Valente refere que é possível definir uma arquitetura mais simples,

Nesta arquitetura pode-se visualizar as bases de dados que compõe o DW. Existe um extrator que é responsável pela deteção automática das mudanças nas bases de dados originais. Sempre que existe uma mudança no conteúdo da BD, informação nova ou atualizada, e que é relevante para o DW é propagada para o integrador.

O integrador é responsável em integrar o conteúdo fornecido pelos extratores e fornecê-lo ao DW.

Para integrar os dados de diferentes tipos de bases de dados é necessário que estes passem por alguns processos.

1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.4 Arquitectura de um Data Warehouse - Valente

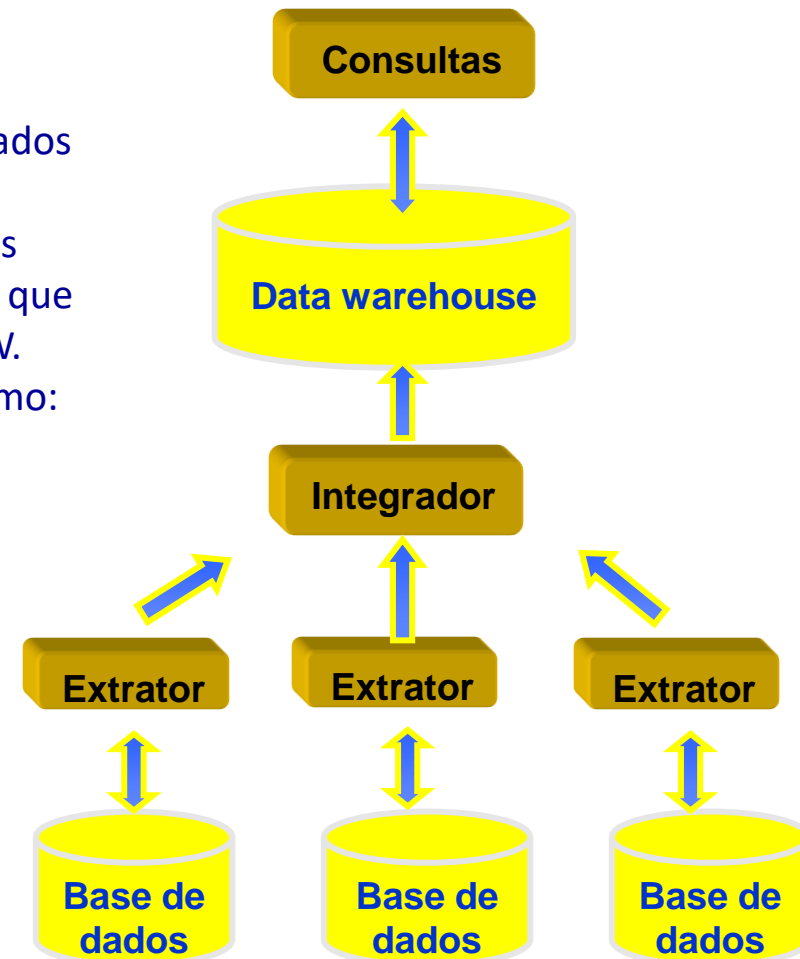
Arquitectura de Valente

Primeiro, os dados devem ser ajustados ao modelo de dados utilizado pelo DW

Posteriormente os novos dados devem ser unidos com os dados já existentes, resolvendo possíveis inconsistências que possam existir entre as bases de dados e os dados do DW.

Uma lista das tarefas do integrador pode conter itens como: especificar as diferenças das bases de dados, definir relacionamentos entre dados de múltiplas bd's, resolver duplicações/inconsistências e determinar como as informações serão integradas dentro do D W.

Fonte Mazzola (2001)



1.2. Ambiente e estrutura funcional de um Data Warehouse

1.2.5. Classificação de Data Warehouses

3 Modelos

Enterprise-Wide Data Warehouses

- Recolhe toda a informação acerca de assuntos que abrangem toda a organização;
- Trata-se de grandes projectos que requerem investimentos em larga escala em tempo e recursos.

Data Mart

- Um subconjunto dos dados de toda a organização com valor para um grupo específico de utilizadores. O seu âmbito é confinado a um grupo específico ou grupos seleccionados, tal como um data mart para marketing.
 - O Data Mart pode ser independente ou dependente (directo do Data Warehouse)
- São geralmente vocacionados a suprir as necessidades informacionais de uma parte da organização, tal como um departamento, sendo mais focados.

Warehouse Virtual

- Um conjunto de vistas sobre as bases de dados operacionais;
- Só algumas das vistas sumarizadas podem ser materializadas;
- Proporcionam vistas das bases de dados operacionais que são materializadas para permitir um acesso eficiente.

1.3. Ciclo de vida e desenvolvimento incremental de um DW

1.3.1. Evolução dum DW

1.3.1.1. Perfil de utilização e implicações

O utilizador do DW é uma pessoa a que poderemos chamar de “Analista DSS”. Trata-se de uma pessoa que é, antes de mais, um gestor e, em segundo lugar, um técnico.

- A perspectiva do Analista DSS é do tipo “dá-me o que eu penso que quero e, então, dir-te-ei o que verdadeiramente me interessa”. Trata-se do processo heurístico. Esta atitude face ao DW tem implicações imediatas no ciclo de desenvolvimento das aplicações centradas em DW;
- O ciclo de desenvolvimento clássico (SDLC - System Development Life Cycle) assume que os requisitos são conhecidos antes da concepção e assim conduzem todo o processo, podendo classificar-se como “requirements driven”;
- Também, se o DW (ou um DMart inicial) tem sucesso, rapidamente o âmbito é alargado (novos temas ou novos DMarts).

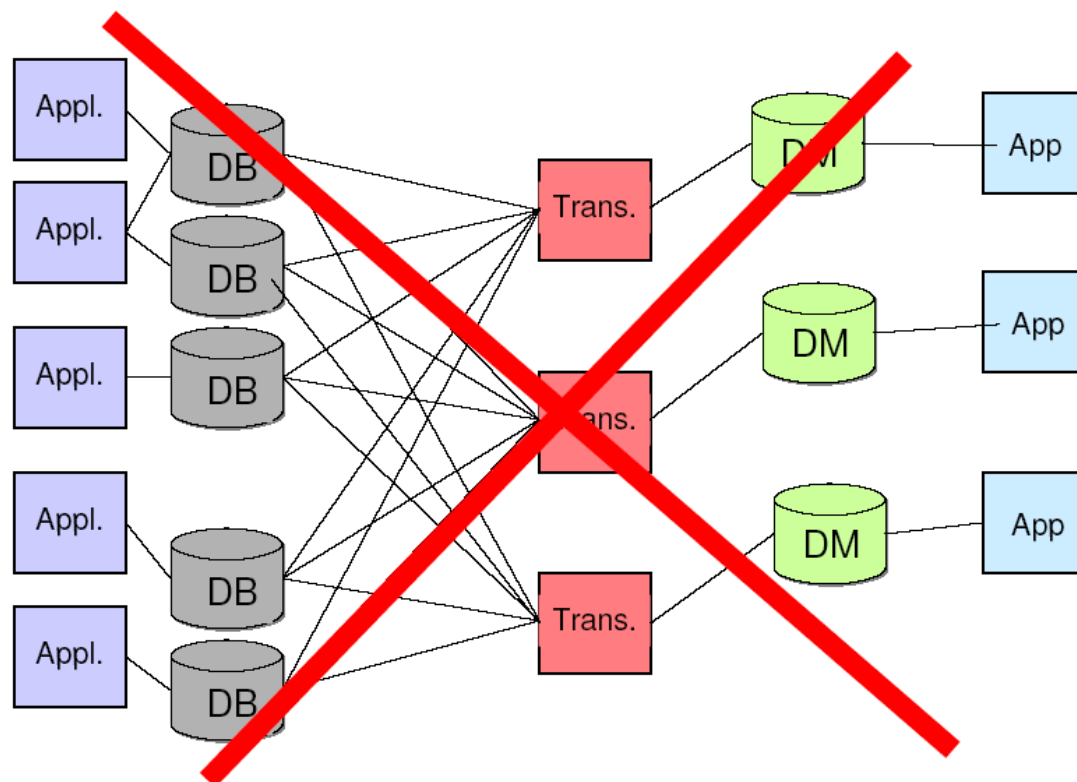
1.3. Ciclo de vida e desenvolvimento incremental de um DW

1.3.1. Evolução dum DW

1.3.1.2. Perigos: Muitas forma de criar um DW não integrado (1)

Muitas vezes, a concepção e instalação de um DM e o seu sucesso, leva a que outros departamentos na empresa empreendam iguais iniciativas.

Mas, se forem iniciativas isoladas, há o perigo de voltar-se às velhas “ilhas de informação”, sem integração e com desperdício de recursos.



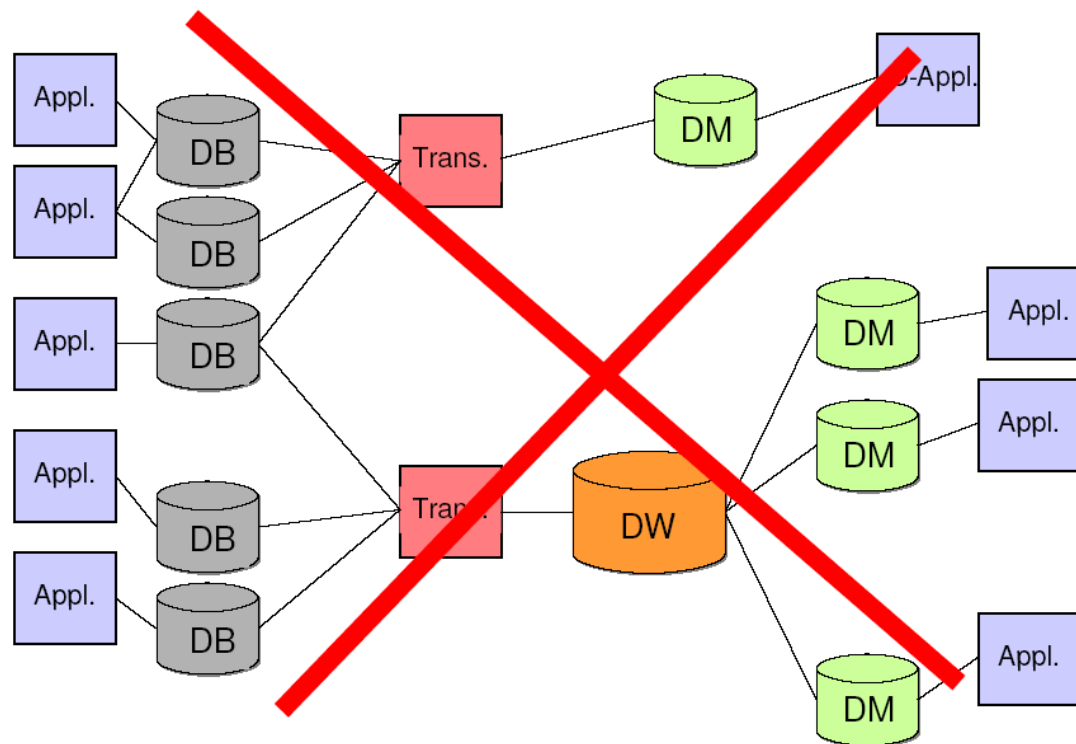
1.3. Ciclo de vida e desenvolvimento incremental de um DW

1.3.1. Evolução dum DW

1.3.1.2. Perigos: Muitas forma de criar um DW não integrado (2)

Neste caso é visível **uma integração parcial**. Se vários DMs têm como fonte o próprio DW e, como tal, um referencial informacional comum, um deles, usa fontes de dados “proprietárias” e como tal, dispõe de um ETL próprio, estando desintegrado do conjunto.

Está-se **mais perto da integração total**: provavelmente uma fase intermédia da solução do caso anterior. Mas, pode tratar-se de uma **regressão**...



1.3. Ciclo de vida e desenvolvimento incremental de um DW

1.3.2. Desenvolvimento do DW

1.3.2.1. Ciclo de Desenvolvimento

O ciclo de desenvolvimento clássico (SDLC - System Development Life Cycle) assume que os requisitos são conhecidos antes do desenho e assim conduzem todo o processo, podendo classificar-se como “requirements driven”;

No caso das **aplicações do mundo DSS (Sistemas de suporte à decisão)**, os requisitos são, **normalmente, a última coisa a ser descoberta**, tratando-se de um processo inverso do anterior (identificado, em alguma literatura, como CLDS - inverso de SDLC). Aqui, o processo de desenvolvimento é conduzido pelos dados, já que é o seu conhecimento que inicia o processo, podendo classificar-se como “data driven”.

1.3. Ciclo de vida e desenvolvimento incremental de um DW

1.3.2. Desenvolvimento do DW

1.3.2.1. Ciclo de Desenvolvimento

Como se depreende do ciclo de desenvolvimento do DW, é **importante o “feedback” do utilizador**, por forma a que sejam feitos os ajustes necessários. Grande parte deste feedback pode ser conseguido indirectamente, através do processo de monitorização [INMON96], de modo a:

- identificar os dados que são utilizados e com que frequência;
- calcular o tempo de resposta do sistema aos pedidos dos utilizadores;
- determinar quem usa efectivamente o DW;
- especificar quanto do DW é utilizado;
- estabelecer um padrão temporal da utilização do DW.

1.3. Ciclo de vida e desenvolvimento incremental de um DW

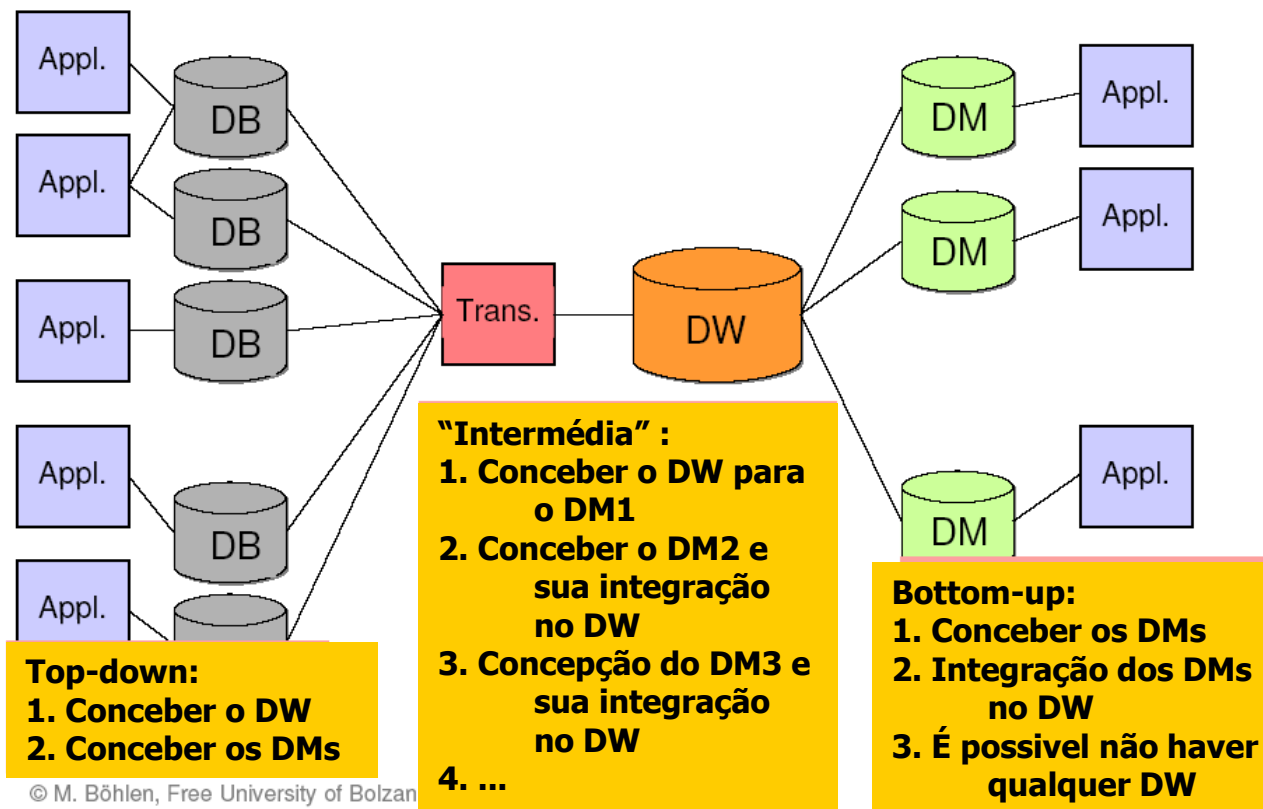
1.3.2. Desenvolvimento do DW

1.3.2.2. Bottom-Up vs. Top-Down

O tipo de **abordagem no desenvolvimento** do DW resulta, em regra, **do processo evolutivo histórico**.

Assim, pode ser:

- 1) **Top-down**, tradicional, mais pensado e arquitectado;
- 2) **Bottom-up**, quando o tempo é factor determinante e uma iniciativa isolada foi absorvida pela criação de um referencial informacional;
- 3) **Intermédia**, mais evolutiva.



1. BI – Da Informação ao Conhecimento



1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.1 O que é a Business Intelligence

“BI é um termo abrangente que inclui aplicações, infraestruturas, ferramentas e as melhores práticas que permitem o acesso à análise da informações e assim melhorar e otimizar decisões e desempenho” (Gartner It Glossary)

“Bi é um processo que aproveita a tecnologia existente para transformar dados em informação utilizável por decisores, gestores e outros profissionais na tomada de decisões empresariais estratégicas e de gestão” (Magalhães 2017)

“BI é considerado um instrumento de análise descritiva, uma vez que descreve um passado e um estado presente” (Pratt, 2017)

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.1 O que é a Business Intelligence

“Os sistemas de Business Intelligence surgiram com o objetivo de ajudar as empresas na recolha, compreensão e exploração dos dados, de forma a extrair informação, capaz de dar suporte à tomada de decisão (Watson & Wixom 2007)”

“Estes sistemas combinam a recolha de dados, o armazenamento dos mesmos, a gestão do conhecimento com diversas ferramentas de análise, as quais possibilitam extrair informação relevante a partir dos dados armazenados” (Santos & Ramos 2006). “

De salientar que o BI permite reconhecer por exemplo tendências de mercado e obter informações essenciais à tomada de decisão

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.1 O que é a Business Intelligence

BI é uma combinação de tecnologias

- Data Warehousing (DW)
- On-Line Analytical Processing (OLAP)
- Data Mining (DM)
- Visualização de Dados (VIS)
- Análises de Cenários (what-if)
- Customer Relationship Management (CRM)

O “oposto” da Inteligência Artificial (AI)

- Os sistemas AI tomam decisões para os utilizadores
- Os sistemas BI ajudam os utilizadores a tomar as decisões certas, baseadas nos dados disponíveis
- Contudo, muitas técnicas BI têm raízes na AI.

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.1 O que é a Business Intelligence

- **Ou simplesmente:**
- **Ajuda na Inteligência empresarial.**
- **Objectivos:**
- **Possibilitar de uma maneira simples e eficaz aos utilizadores a possibilidade de recolher e obter informações corporativas.**
- **Permitem às empresas acederem e recolherem informações contidas, por exemplo, em DWs, com a finalidade de gerir os recursos para a tomadas de decisão e definições de estratégias.**

1. BI – Da Informação ao Conhecimento

➤ **Business Intelligence** significa:

- Utilizar as potencialidades das actuais bases de dados para construir **data warehouses**
 - Uma data warehouse é uma estrutura baseada em tabelas, mas concebida e desenhada
 - Para responder a queries OLAP (*on-line analytical processing*) que envolvem agregação de dados (factos) por diversas **dimensões**
 - Para proporcionar a aplicação de algoritmos de *data-mining* às suas **tabelas de factos**
- Tomar decisões com base nos resultados anteriores, no conhecimento pessoal da organização, no conhecimento criado e gerido pela organização (*know-how*), na sua cultura e no impacto social e humano de possíveis modificações

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

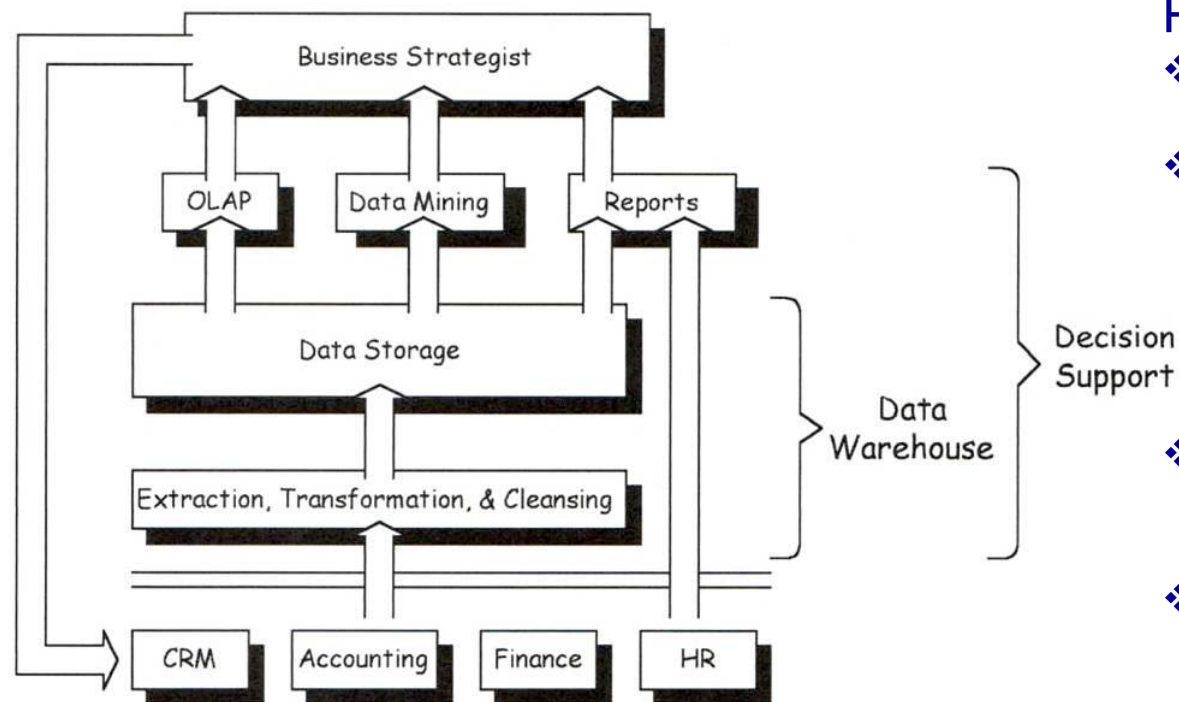
1.4.1.1 O que é a Business Intelligence

- Extrair e integrar dados de múltiplas fontes.
- Fazer uso da experiência já existente.
- Actualizar dados contextualizados.
- Trabalhar com cenários hipotéticos.
- Procurar relações de causa e efeito.
- Transformar os registos obtidos em informações úteis para o conhecimento empresarial.

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.2 Business Intelligence Loop

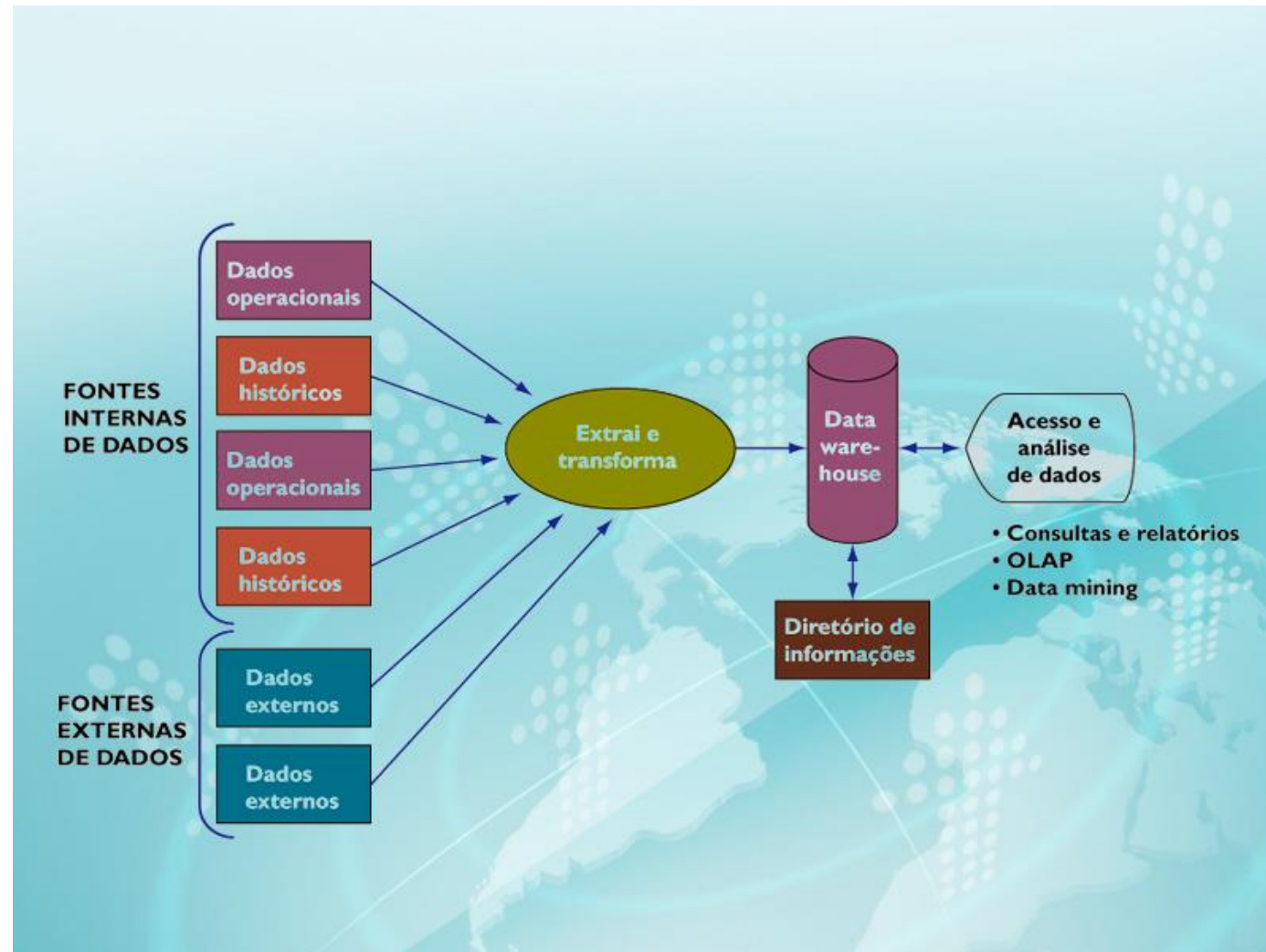


Retirado de: *Object-Oriented Data Warehouse Design*, William Giovinazzo

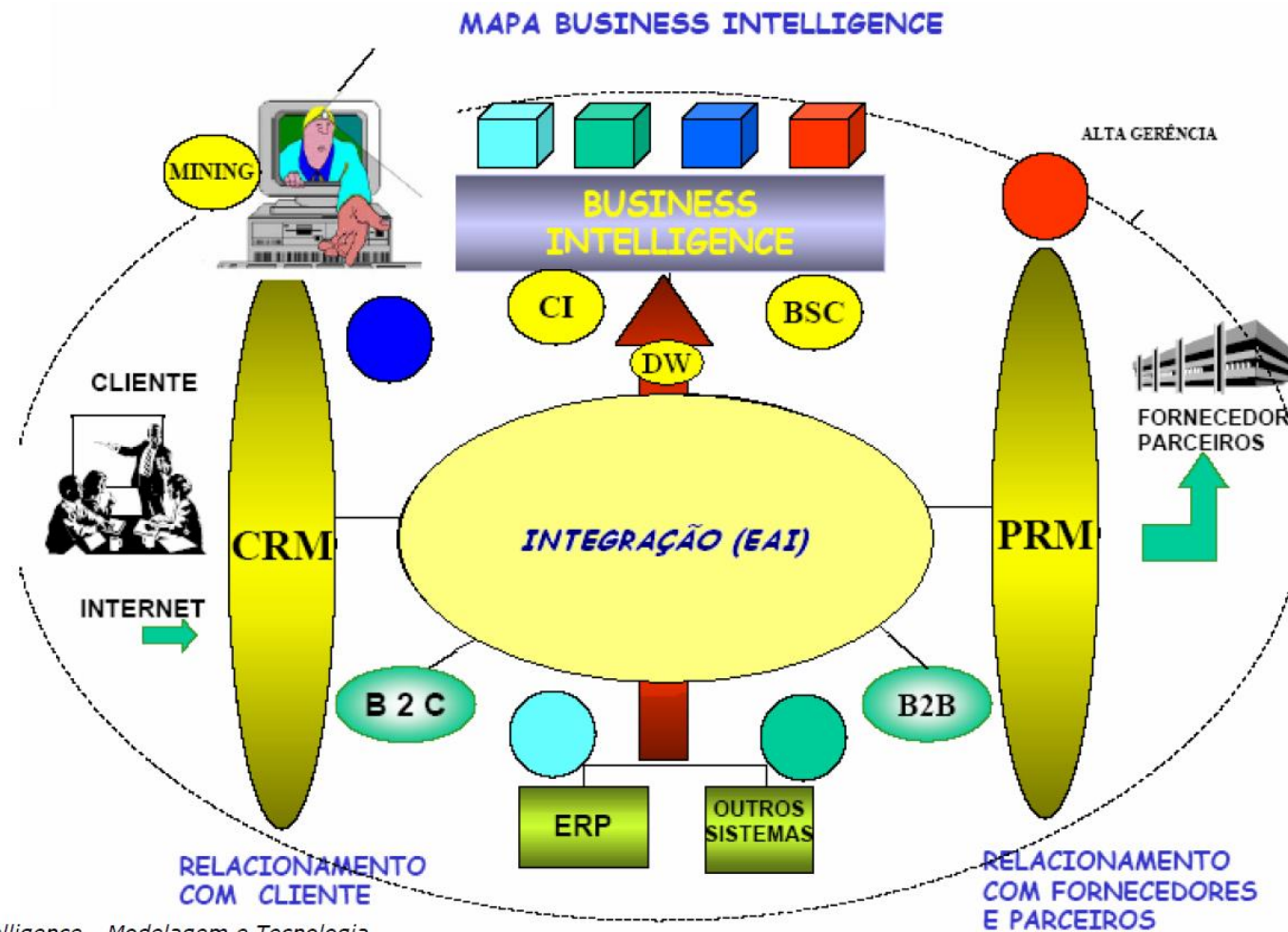
Porquê a BI?

- ❖ A organização deve desenvolver uma estratégia
- ❖ Para pensar uma estratégia ganhadora, deve poder antecipar-se o futuro. Compreender o passado é a melhor forma de prever o futuro.
- ❖ O estratega de negócio procura comportamentos no ambiente da empresa
- ❖ O Data Warehouse é o componente central do Business Intelligent Loop

ONDE FICAM OS DADOS DO BI?



MAPA DO BI



1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.3 Infraestruturas tecnologicas de Business Intelligence

- Os sistemas de Business Intelligence integram a atividade de exploração da Data Warehouse, denominada como “repositório de dados” construído para consolidar a informação de uma organização num formato consistente, permitindo posteriormente a análise de dados.

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.4 Exemplos de Interrogações típicas BI

- Q1:** Em 11 de Outubro de 2008, mostrar os 5 produtos mais vendidos para cada subcategoria de produtos que tenha contribuído com mais de 20% das vendas da sua categoria.
- Q2:** Relativamente a 23 de Maio de 2009, determinar a prioridade de envio e lucro bruto esperado para as encomendas que tenham os 10 maiores lucros brutos de entre as encomendas que não tenham sido ainda expedidas. Considerar só as encomendas relativas ao segmento de mercado de livros.
- **Os modelos de bases de dados típicos e respetivas bases de dados não são adequadas para responder a este tipo de interrogações.**

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.2.1.5 Bi no Contexto das Organizações

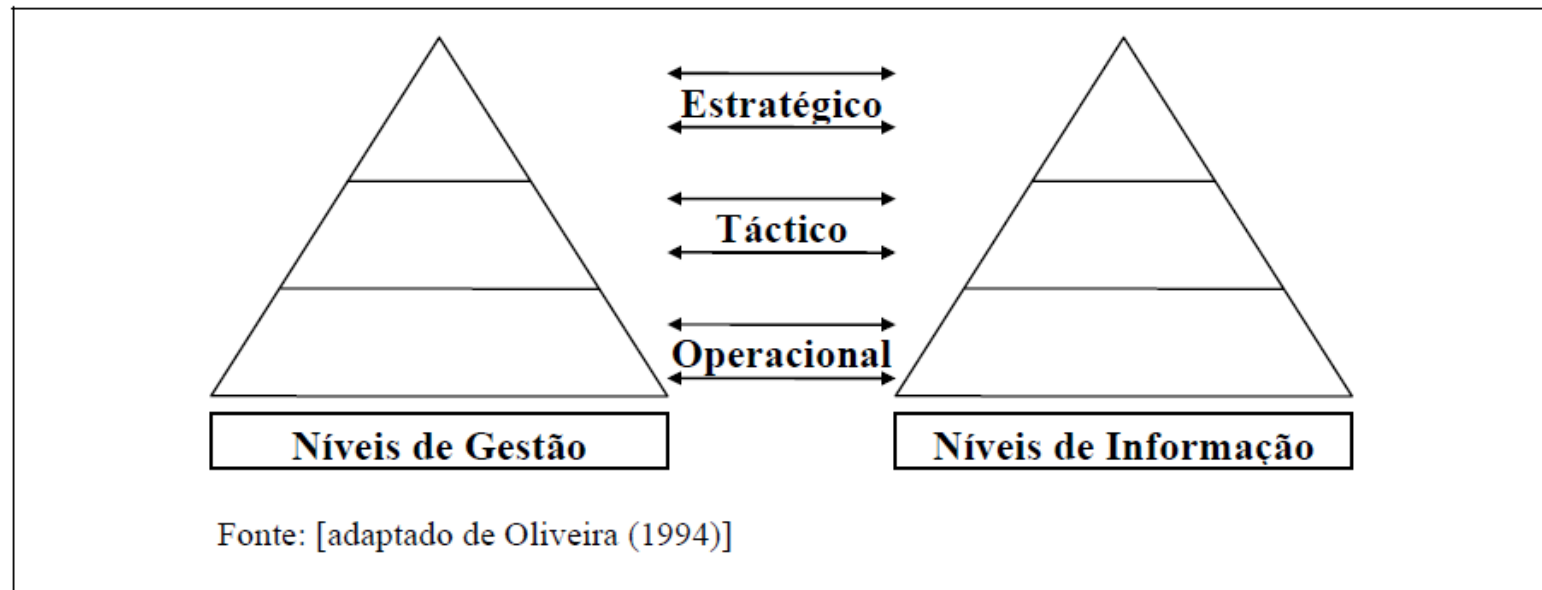
Figure 6.3 Organizations are composed of levels, with each using information technology to automate activities or assist in decision making.



1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4. 1 Business Intelligence

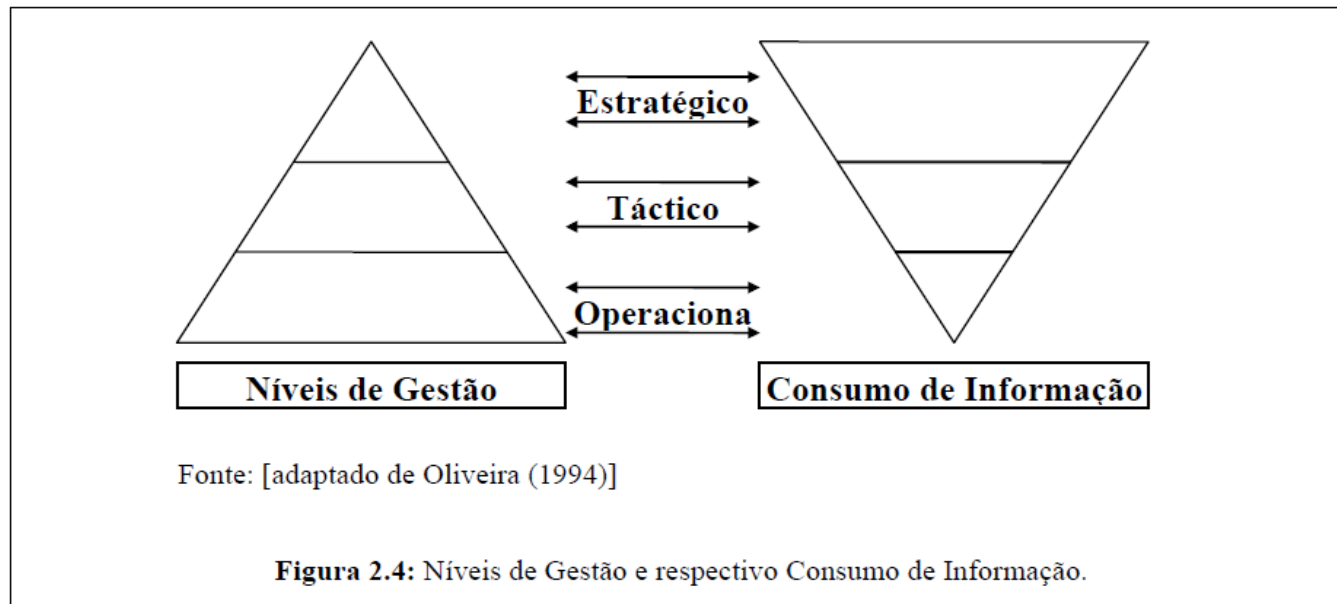
1.4.1.5 Bi no Contexto das Organizações



1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.2.1 Business Intelligence

1.2.1.5 Bi no Contexto das Organizações



1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.5 Bi no Contexto das Organizações

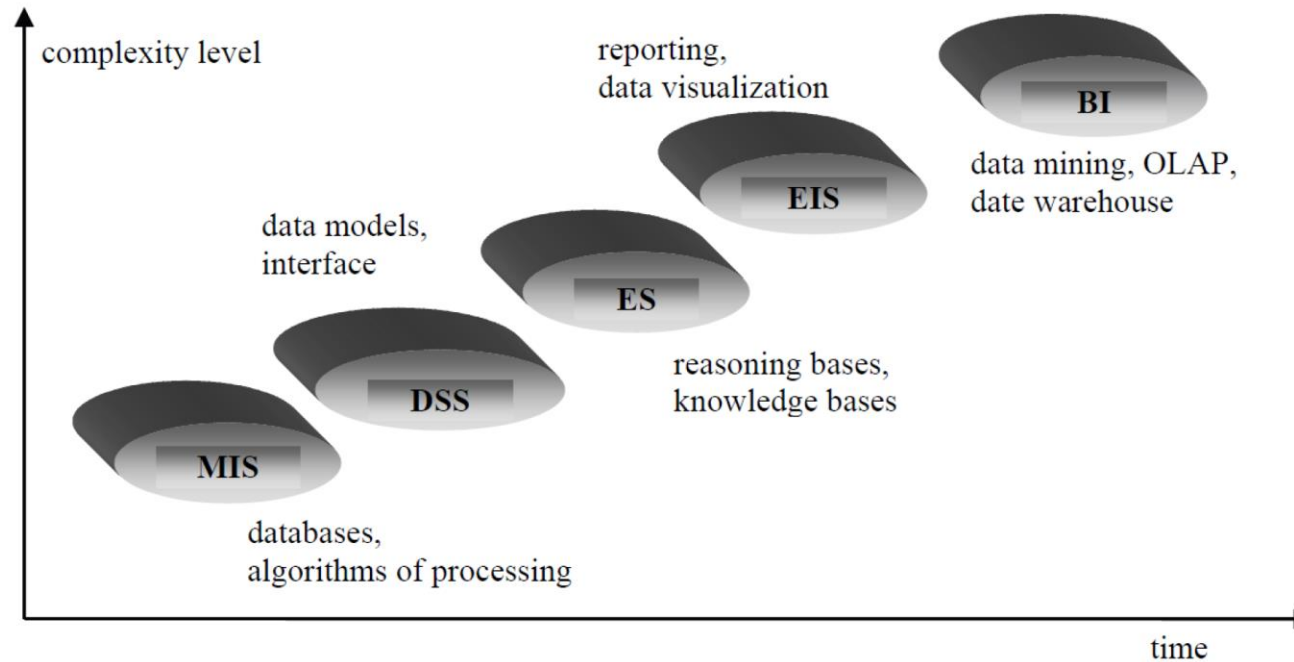


Figura 7 – Vantagens dos Sistemas de BI. Fonte: Olszak e Ziemba, 2007

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4.1 Business Intelligence

1.4.1.6 Funcionalidades de BI

- Uma ferramenta de BI deve apresentar funcionalidades que permitam:
 - Obter e modelar dados provenientes de sistemas de gestão de base de dados (SGBDs);
 - Extrair, transformar, efetuar limpeza e carregamento de dados;
 - Visualizar e efetuar análises de dados;
 - Mostrar dados em forma de dashboards;
 - Implementação de processos de Data Mining;
 - Criação de relatórios de dados.

1. Sistemas de Data Warehousing: Infraestrutura Informacional e Funcional

1.4. Business Intelligence

1.4.1.1 Principais softwares de BI

- JasperSoft
- OpenI
- Palo
- Pentaho
- SpagoBI
- Vanilla
- Manta Flow
- ManageEngine Analytics Plus
- SAP BusinessObjects
- ...

QUESTÕES

Filipe Sá

Filipe.sa@isec.pt