



Cadeira de Tecnologias de Informação

Ano lectivo 2009/2010

Conceitos Fundamentais de Gestão de Dados

Tópicos

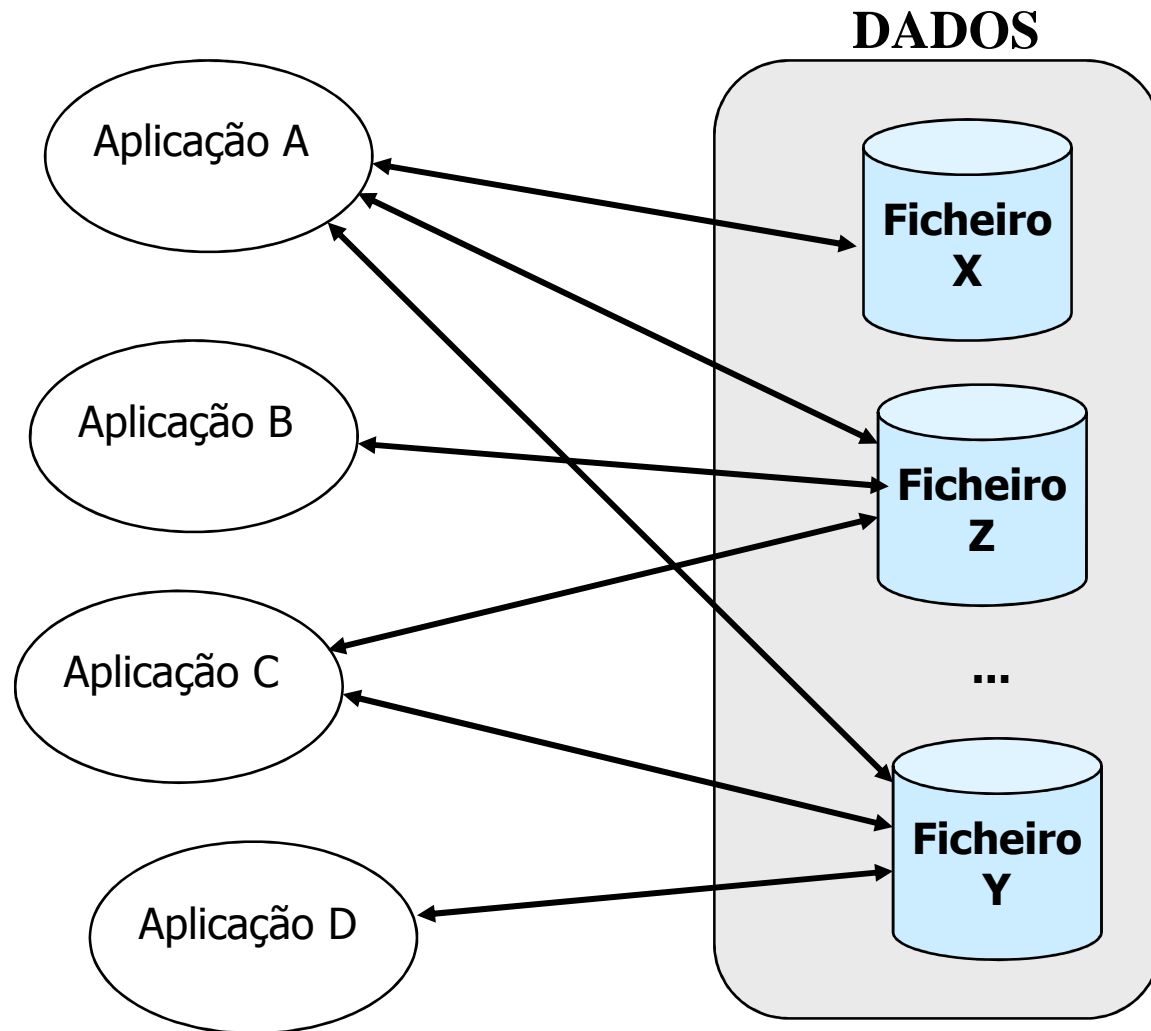
1. Conceitos de Base de Dados e de Sistema de Gestão de Bases de Dados
2. Modelo Relacional de Dados
 - Estrutura
 - Operações Básicas de Interrogação
3. Conceitos de *Data Warehouse* e *Data Mart*
4. Conceito de *Data Governance*

Conceitos Fundamentais

Uma **Base de Dados** é um conjunto interrelacionado de dados numa determinada área.

Um **Sistema de Gestão de Bases de Dados (ou Gestor de Bases de Dados)** é o software que gere o armazenamento, manipulação e pesquisa dos dados existentes na base de dados, funcionando como um interface entre as aplicações e os dados necessários para a execução dessas aplicações (exemplos: IMS, DB2, MySQL, Informix Dynamic Server, Oracle Server, Sybase SQL Server, Microsoft SQL Server, Access).

Sistemas Tradicionais - Gestão de Ficheiros



Limitações dos Sistemas de Gestão de Ficheiros

- **Redundância de Dados**

Os mesmos dados são armazenados em vários ficheiros

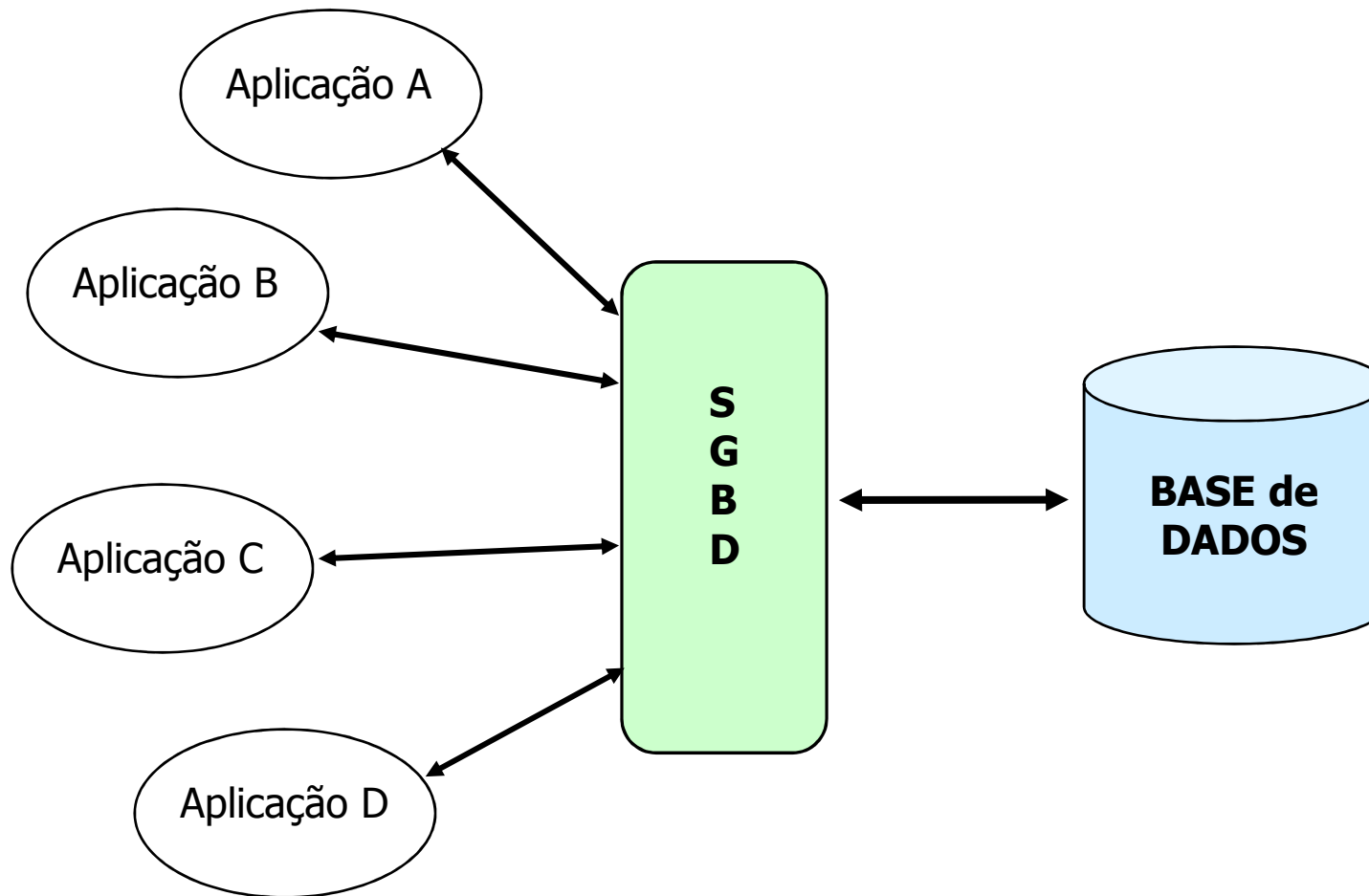
- **Isolamento dos Dados**

As aplicações dificilmente acedem aos ficheiros de outras aplicações

- **Inconsistência dos Dados**

Várias cópias diferentes dos mesmos dados

A Tecnologia de Bases de Dados



A Tecnologia de Bases de Dados

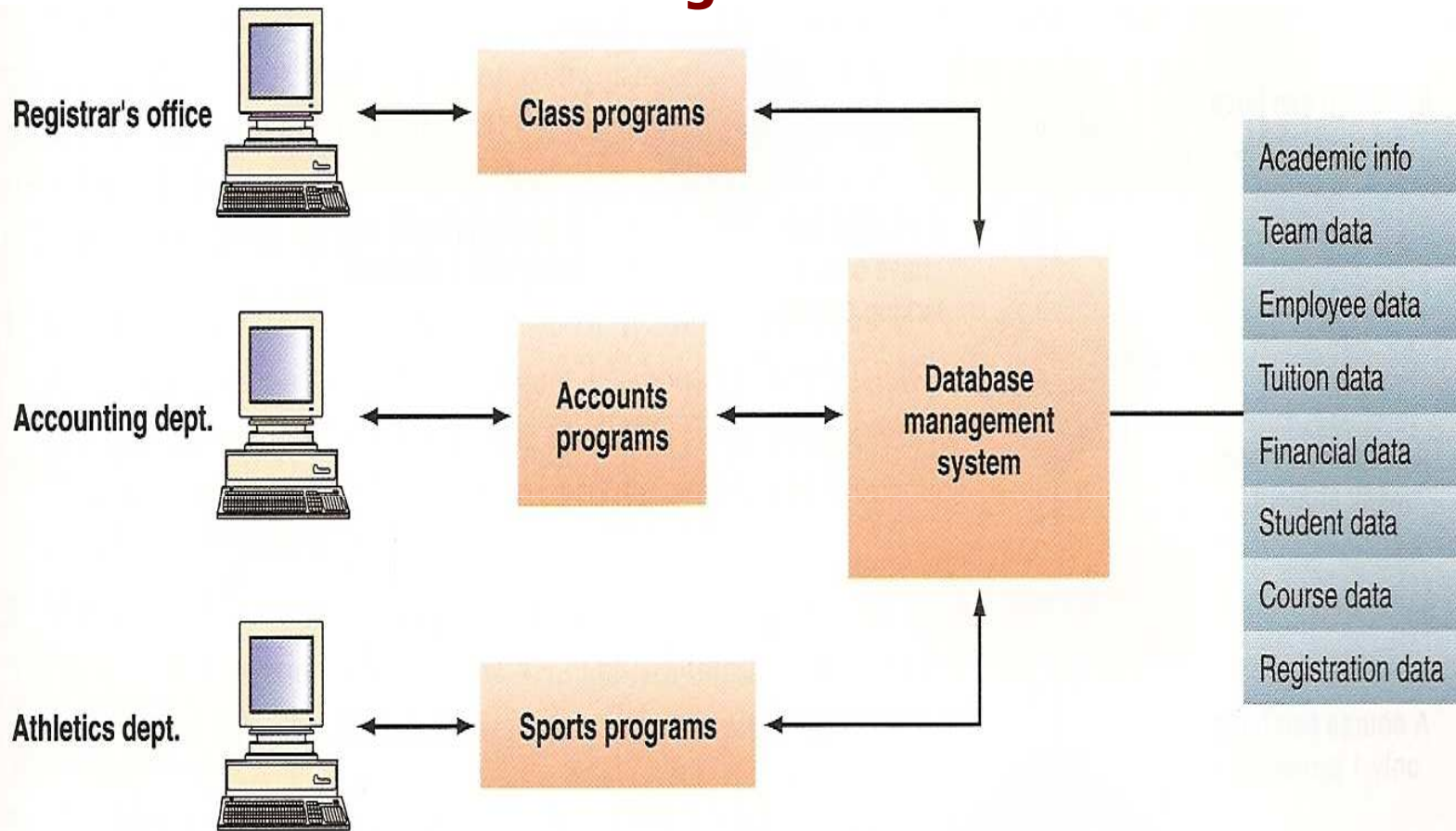


FIGURE 4.3 A database management system (DBMS) provides access to all data in the database.

Reiner, R.K.; Turban, E.; Potter, R.E. (2007). *Introduction to Information Systems – Supporting and Transforming Business*, John Wiley.

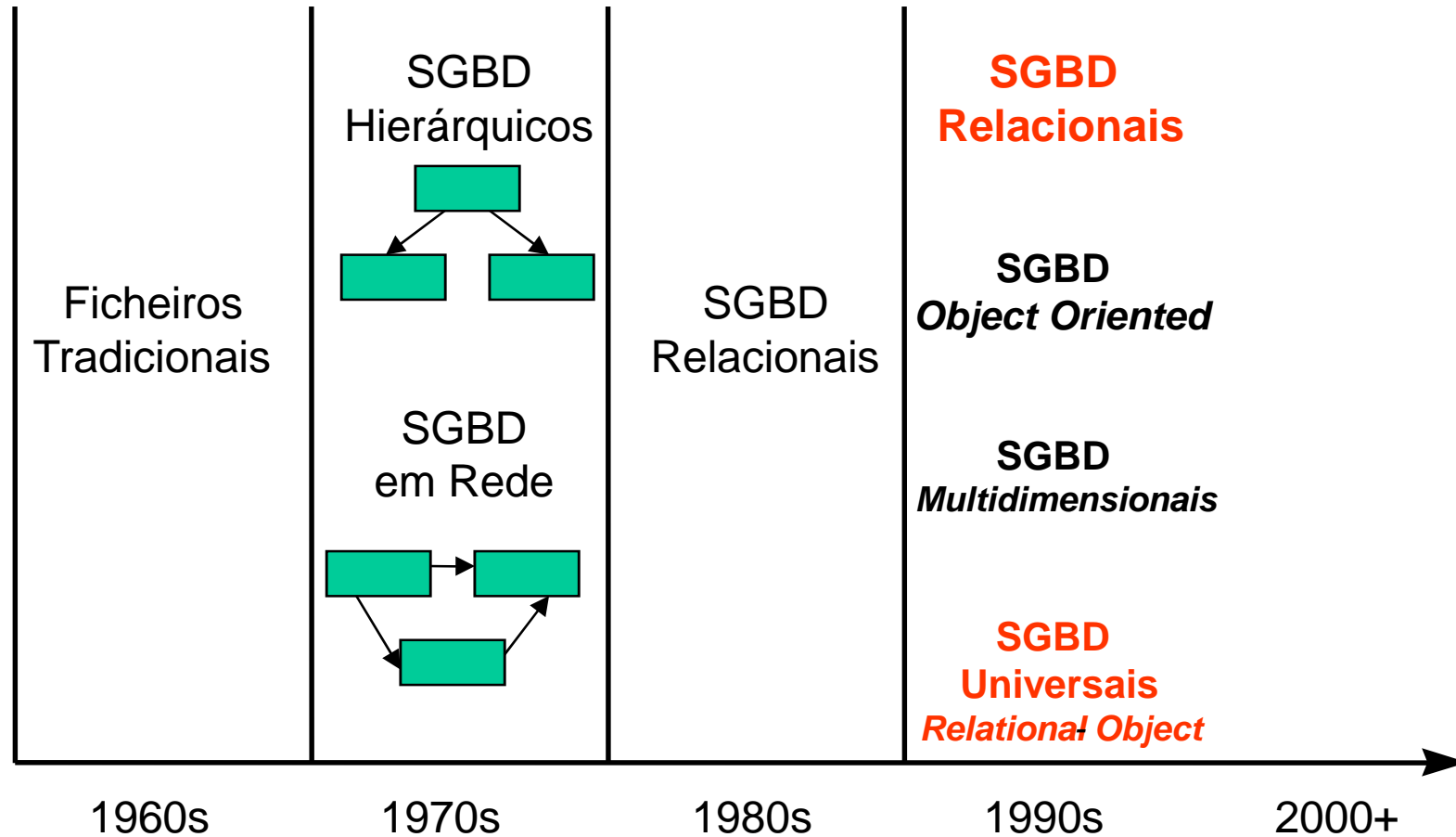
Algumas Vantagens da Tecnologia de Bases de Dados

- Integração de dados de suporte a múltiplas aplicações
 - Diminuição de redundâncias
 - Integridade dos dados (evitar inconsistências)
- Facilitar a pesquisa (permite questões *ad hoc*)
- Aumentar a flexibilidade das aplicações (independência dos dados relativamente aos programas)
- Desenvolvimento de mecanismos de segurança
- Controlo da concorrência

Requisitos Fundamentais de um SGBD

- **Segurança** - protecção da base de dados contra acessos não autorizados
- **Integridade** - validação de operações que coloquem em risco a consistência dos dados
- **Controlo de concorrência nos acessos** - coordenação da partilha dos dados pelos vários utilizadores (SGBDs multiutilizador)
- **Recuperação de falhas** - restaurar a integridade da base de dados depois da ocorrência de uma falha. Mecanismos de recuperação (fundamentalmente baseados em redundância): *backups, transaction logging* (ficheiro *transaction log*, dados para repor as últimas transacções)

Evolução dos SGBD



Modelo Relacional

O modelo relacional é o mais amplamente disponível no mercado.
A estrutura fundamental do modelo relacional é a tabela (formalmente designada por relação)

Ex: Tabela Empregado

NºEmp	Nome	Telef	Categ	Salario	Comissão	Função
10	Antunes	12554	B	1000	15	Analista
20	Bento	54321	A	2500	50	Director
30	Correia		E	960.5		Porteiro
40	Dias	23457	C	990.5		Programador
50	Edmundo		B	1200	12.5	Contabilista
...

Modelo Relacional

Exemplo de duas tabelas relacionadas

Empregado

n_emp	nome ...	cod_categ
01	João ...	c3
02	José ...	c1
03	Gil ...	c2
04	Júlio ...	c1
...

Categoria

cod_cat	design	venc_base
c1	Director	3000
c2	Técnico	2000
c3	Admin	1200
c4	Auxil	900
...		

Modelo Relacional

Linguagens de um SGBD Relacional

DDL - *Data Definition Language*;

DML - *Data Manipulation Language*.

- *SQL - Structured Query Language*

Linguagem para definição e manipulação de dados comum aos SGBD relacionais (inclui DDL e DML).

- *QBE - Query By Example*

Interface gráfico que facilita a elaboração de estruturas de interrogação da base de dados. Acaba por ser um gerador de SQL (a única linguagem que o SGBD entende)

No Access estão disponíveis as linguagens SQL e QBE

Modelo Relacional

Operações Básicas de Interrogação (Álgebra Relacional)

Projeção

Seleciona um subconjunto de colunas de uma tabela de dados (atributos).

Seleção ou Restrição

Seleciona um subconjunto de linhas de uma tabela de dados.

As condições de seleção podem envolver:

- Constantes (valores do domínio de um atributo, texto entre aspas)
- Nomes de um atributo
- Comparações aritméticas (=, <>, <, <=, >=)
- Operadores lógicos (*and*, *or*, *not*)

Junção

Combina linhas de tabelas diferentes com base na comparação de valores das colunas seleccionadas.

Modelo Relacional

Operação de Projecção

Tabela Cliente

N_Cli	Nome_Cli	Morada	Cidade	País
1	António Abreu	R. Alberto Antunes	Porto	Portugal
2	Bernardo Bento	Rue de La Paix	Bruxelas	Bélgica
3	Carlos Castro	R. Clara Campos	Lisboa	Portugal
...
20	Manuel Matos	R. Marco Moita	Maputo	Moçambique
21	Mário Martins	R. Maria Mendes	Luanda	Angola

Projecção da Tabela Cliente sobre Nome e Morada

Nome_Cli	Morada
António Abreu	R. Alberto Antunes
Bernardo Bento	Rue de La Paix
Carlos Castro	R. Clara Campos
...	...
Manuel Matos	R. Marco Moita
Mário Martins	R. Maria Mendes

Modelo Relacional

Operação de Selecção ou Restrição

Tabela Cliente

N_Cli	Nome_Cli	Morada	Cidade	País
1	António Abreu	R. Alberto Antunes	Porto	Portugal
2	Bernardo Bento	Rue de La Paix	Bruxelas	Bélgica
3	Carlos Castro	R. Clara Campos	Lisboa	Portugal
...
20	Manuel Matos	R. Marco Moita	Maputo	Moçambique
21	Mário Martins	R. Maria Mendes	Luanda	Angola

Selecção dos Clientes de Portugal

N_Cli	Nome_Cli	Morada	Cidade	País
1	António Abreu	R. Alberto Antunes	Porto	Portugal
3	Carlos Castro	R. Clara Campos	Lisboa	Portugal

Modelo Relacional

Operação de Junção

Junção das Tabelas Empregado e Categoria

Empregado

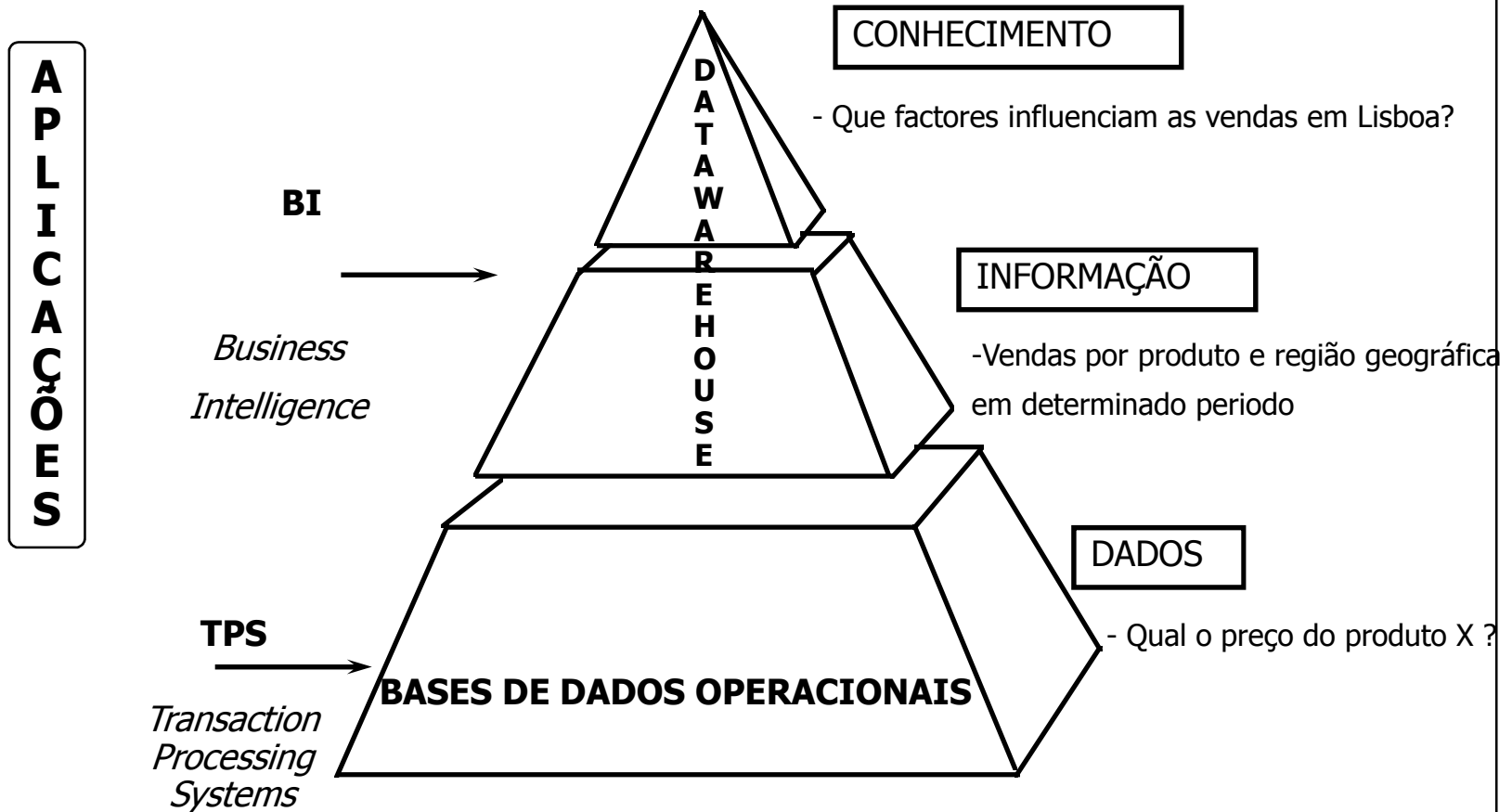
n_emp	nome	...	cod_categ
01	João	...	c3
02	José	...	c1
03	Gil	...	c2
04	Júlio	...	c1
...

Categoria

cod_cat	design	venc_base
c1	Director	3000
c2	Técnico	2000
c3	Admin	1200
c4	Auxil	900

n_emp	nome	...	cod_categ	designação	venc_base
01	João	...	c3	Admin	1200
02	José	...	c1	Director	3000
03	Gil	...	c2	Técnico	2000
04	Júlio	...	c1	Director	3000

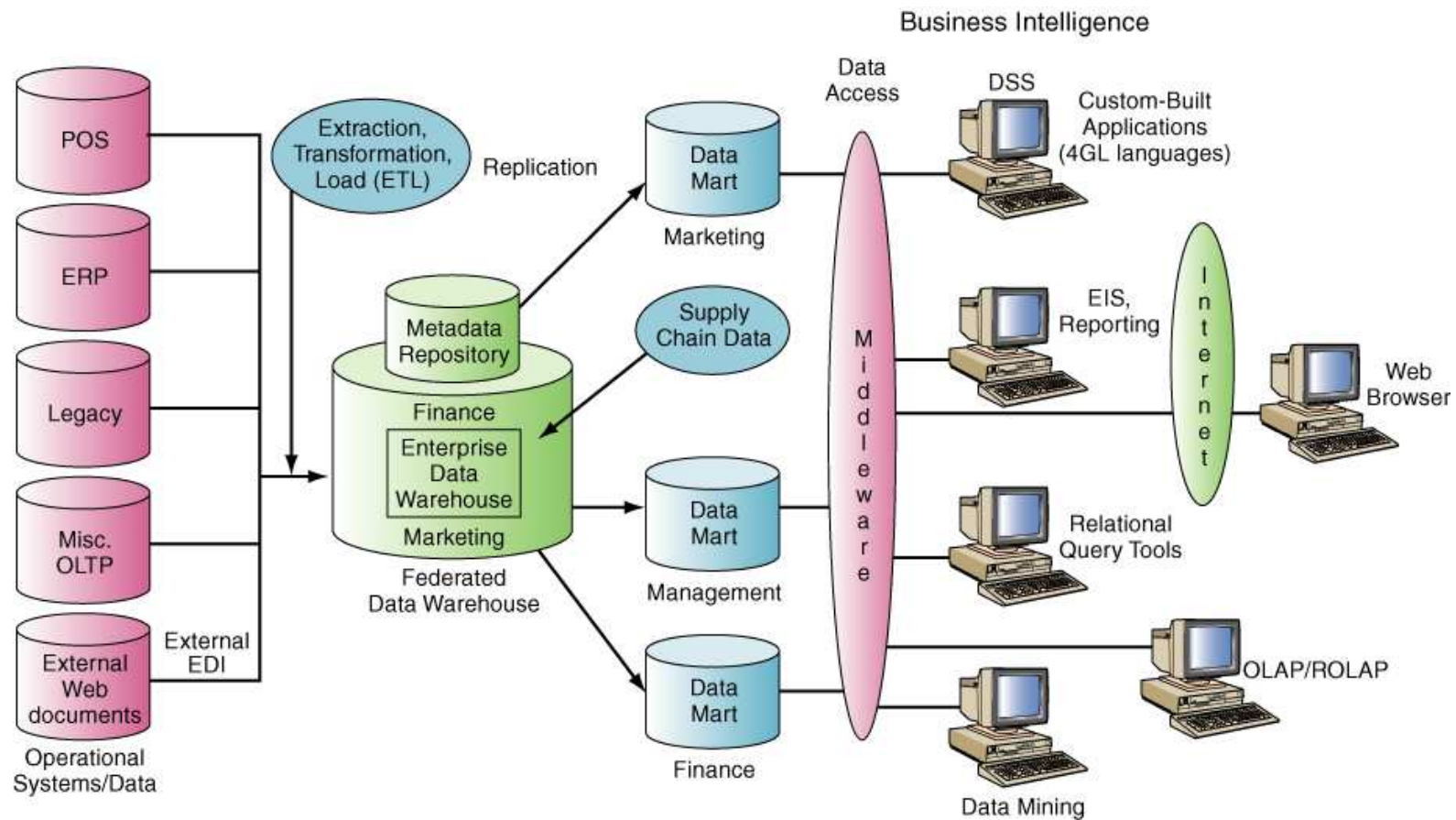
Tipificação das Bases de Dados



Data Warehouse

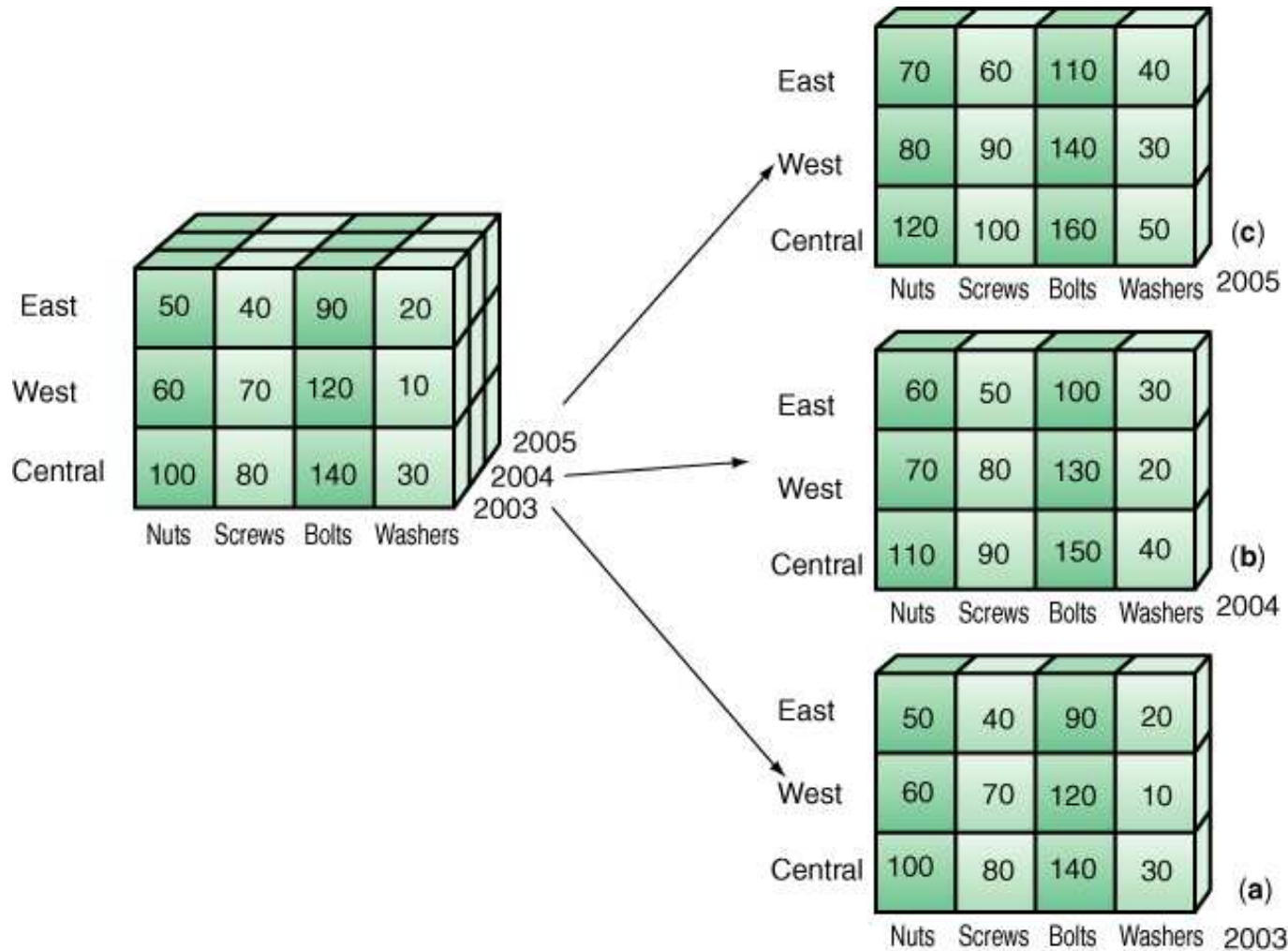
- **Um *Data Warehouse* (DW) é um repositório de dados históricos, organizados por assunto (clientes, produtos, etc) cujo objectivo é suportar a tomada de decisões**
- Os dados e informações existentes no DW são provenientes das bases de dados operacionais e são extraídos através de ferramentas ETL (*Extraction, Transformation and Loading*)
- Os dados do DW são normalmente apresentados aos utilizadores na forma multidimensional

Data Warehouse Framework & Views



Reiner, R.K.; Turban, E.; Potter, R.E. (2007). *Introduction to Information Systems – Supporting and Transforming Business*, John Wiley.

Base de Dados Multidimensional



Reiner, R.K.; Turban, E.; Potter, R.E. (2007). *Introduction to Information Systems – Supporting and Transforming Business*, John Wiley.

Vantagens do *Data Warehouse*

- Disponibilidade de uma visão consolidada dos dados organizacionais
- Acesso rápido e fácil via *Web browsers*

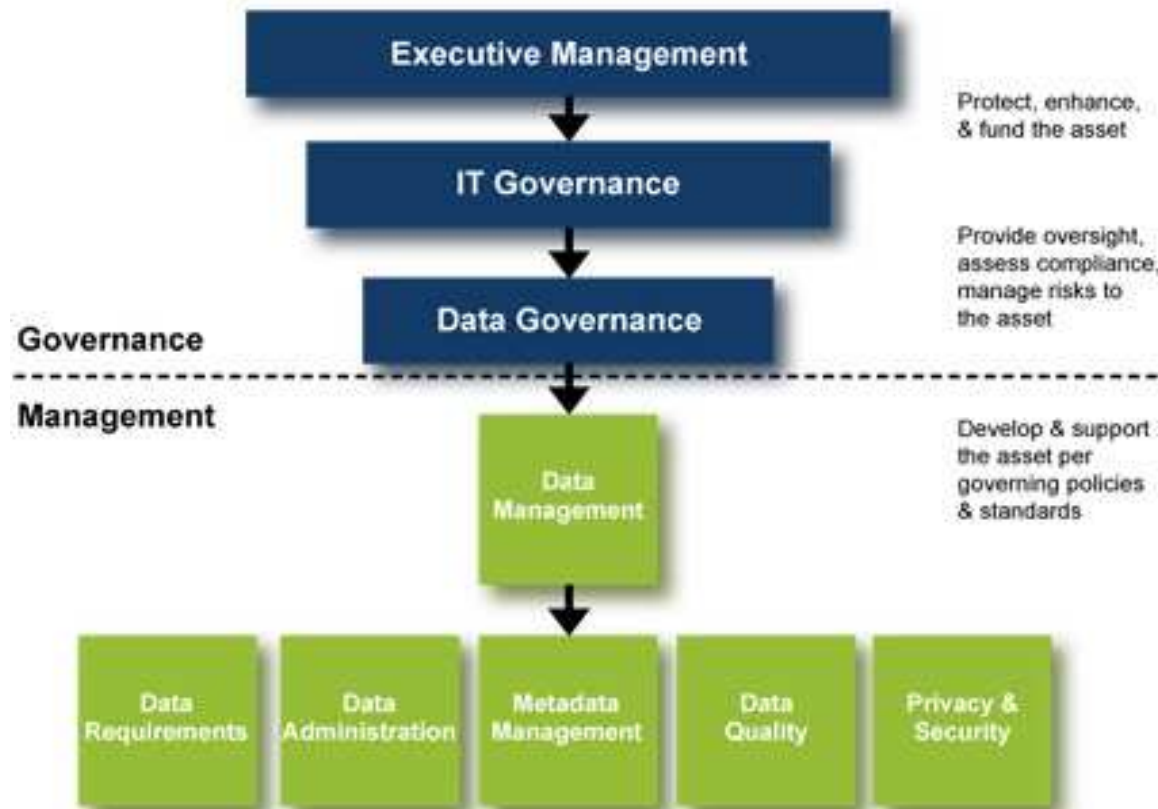
Data Marts

- Um *data mart* é um sub-cojunto de um *datawarehouse* desenhado para satisfazer as necessidades de uma área de negócio, de um departamento, de uma área geográfica, etc
- O *datawarehouse* pode ser construído *bottom up* a partir de um conjunto de *data marts* ou *top down* em que os *data marts* são cópias de subconjuntos do *datawarehouse*

Data Governance (Governança dos Dados)

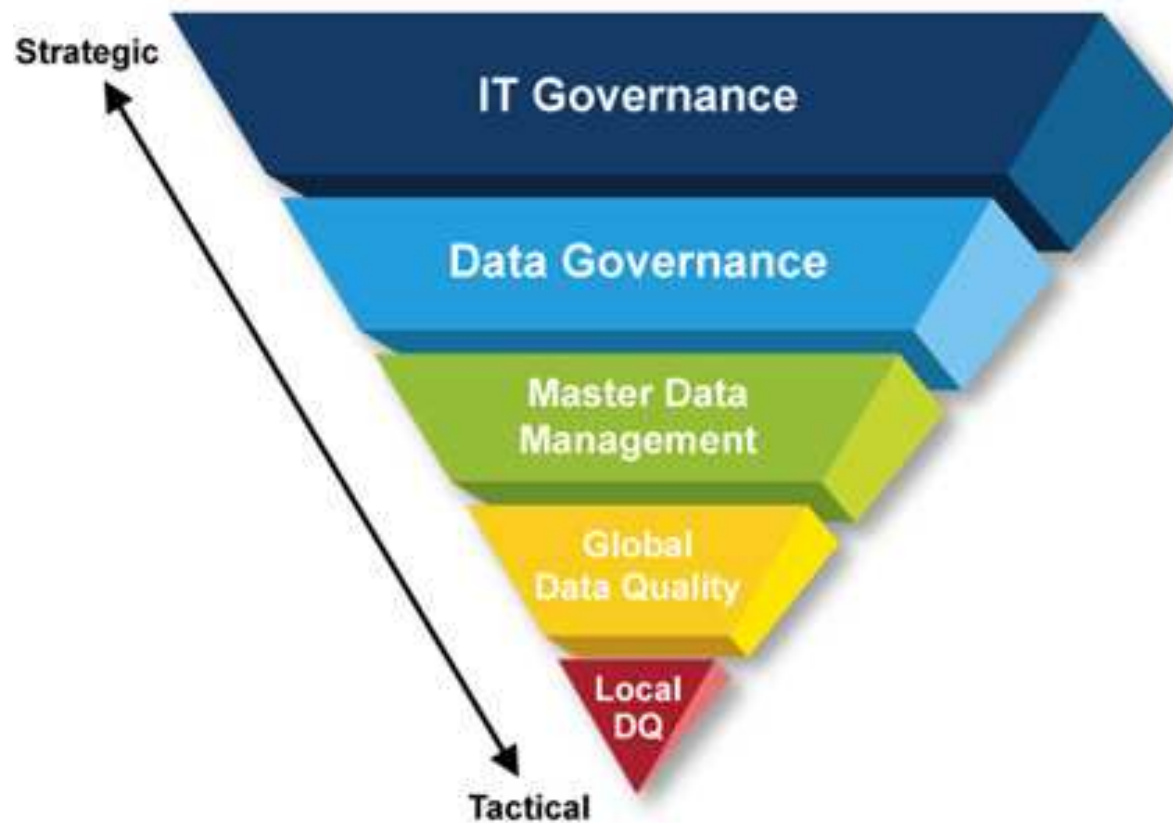
- *Data Governance* (DG) é um conjunto de processos que asseguram que os dados, que são importantes activos, são formalmente geridos em toda a organização
- DG é uma disciplina emergente com uma definição ainda em evolução
- **A disciplina contempla uma convergência da qualidade dos dados, gestão de dados, gestão de processos empresariais e gestão de riscos envolvendo a manipulação de dados de uma organização**
- Através da DG as organizações pretendem efectuar um controlo adequado sobre os processos e métodos utilizados pelos seus empregados que manipulam os dados

Data Governance



Reiner, R.K.; Turban, E.; Potter, R.E. (2007). *Introduction to Information Systems – Supporting and Transforming Business*, John Wiley.

Data Governance (Governança dos Dados)



Reiner, R.K.; Turban, E.; Potter, R.E. (2007). *Introduction to Information Systems – Supporting and Transforming Business*, John Wiley.