

Banco de Dados I

Aula 1

Introdução a Banco de Dados

Meta

Apresentar conceitos importantes e as principais características de um banco de dados.

Objetivos

1. Distinguir dados de informação;
2. Definir o que é Banco de Dados e Banco de Dados Relacional;
3. Descrever a Tecnologia de Banco de Dados;
4. Identificar as três fases de um projeto de Banco de Dados.

Nos tempos mais remotos...

Desde os primórdios, o homem gera e transmite informações. Primeiro, foram as pinturas rupestres, depois esculpiram a pedra e, mais tarde, chegaram ao **papiro**.



Figura 1.1 Pintura rupestre e papiro egípcio

Com a invenção do papel, a informação ficou mais leve e, com o passar dos anos, houve a necessidade de transmitir conhecimentos e descobertas que cresciam a cada dia. O número de informações aumentou, assim como aumentou o número de pessoas que queriam se informar. Assim, as informações viraram livros e os livros lotaram as bibliotecas.



Figura 1.2. Com o aumento da disponibilidade de informações, o número de pessoas que desejavam se informar também aumentou.

O conhecimento gerava cada vez mais conhecimento que, como consequência, gerava mais informação. Desta maneira, mais pessoas procuravam informações ao mesmo tempo. A invenção do computador veio para buscar solucionar este problema do armazenamento e acesso à grande quantidade de informações que não paravam de progredir. Com o tempo, o computador foi se tornando uma tecnologia acessível a um número cada vez maior de pessoas. Mais computadores em uso geravam mais informações. A partir da necessidade de trocar as informações de forma mais rápida e eficiente, nasceu a internet.

A Web aumentou ainda mais o volume e acesso às informações, crescendo exponencialmente – e sem cessar - até nossos dias, lotando os Bancos de Dados.

Hoje, os Bancos de Dados estão presentes em todos os lugares da indústria da tecnologia da informação (TI) e nos negócios em geral. Usamos os Bancos de Dados de maneira direta e indireta todos os dias – transações bancárias, reservas de viagem, relações de emprego, buscas em sites da Web, compras, vendas, etc. Estamos falando de um grande volume de dados, ou seja, informações que precisam ser armazenadas,

acessadas e atualizadas por um indefinido espaço e tempo, proporcionando aos seus usuários agilidade e qualidade de resposta no cruzamento das informações.

Dados X Informações

Até aqui, você já ouviu falar de dados e de informações. Mas você sabe o que eles significam? São nomes diferentes para a mesma coisa? Definir esses conceitos é um passo necessário para compreender o que pode ser armazenado nos Bancos de Dados.

Dado: é qualquer elemento identificado em sua forma bruta que por si só, não conduz a uma compreensão de determinado fato ou situação. O dado pode ser apresentado na forma de números, palavras, imagens ou sons.

Ex: 39°C.

Este dado isolado não diz nada. Ele relata apenas uma medida de temperatura em graus Celsius que pode estar associada a muitos eventos (temperatura corporal, temperatura do ambiente, etc.).

Informação: é o significado (semântica) do dado, isto é, o dado somado ao atributo, relevância e contexto. São dados coletados, organizados e ordenados.

Informação = dado + significado

Box de Curiosidade

O que cerveja tem a ver com fraldas?

Uma das maiores redes de varejo dos Estados Unidos, o Wal-Mart, descobriu em seu gigantesco armazém de dados (códigos de produtos vendidos), que a venda de fraldas descartáveis estava associada à de cervejas (informação). O que isso significa: que as crianças estão bebendo cerveja? Não! Uma investigação mais detalhada revelou que, em geral, eram os homens que saíam à noite para comprar fraldas e aproveitavam para levar algumas latinhas de cerveja para casa.

O setor responsável pela organização dos produtos não pensou duas vezes: os colocou lado a lado. Resultado: a venda de fraldas e cerveja disparou.

Então, temos que: dados = códigos de produtos vendidos e informação = associação dos produtos fraldas e cerveja.

Fim box de Curiosidade

Já pensou tomar uma decisão importante com pouca informação? O risco de algo dar errado é grande, não é mesmo? Na tomada de decisões um Banco de Dados é uma fonte da qual podemos extrair uma grande variedade de informações.

Pois bem, como as informações são extraídas do Banco de Dados, não é possível processar informações em computadores, por exemplo. O que se processa nos computadores são os dados que representam essas informações. Só o ser humano é capaz de elaborar informações por meio de associações de conceitos. Assim, é correto falar em Banco de Dados ou Base de Dados e não em Banco de Informações.

O que é Banco de Dados?

Um Banco de Dados é uma coleção de arquivos (conhecido também como tabelas), que armazenam dados e suas respectivas associações.

Em um computador, os dados são organizados logicamente numa hierarquia que envolve campos, registros, arquivos e Bancos de Dados, tal como a escrita pode ser organizada em textos, palavras, sentenças, parágrafos e documentos. A figura 1.3 ilustra essa hierarquia de dados:

Banco de Dados	Banco de dados vendas		
	Arquivo cliente	Arquivo pedidos	Arquivo produtos
Arquivo	Arquivo cliente		
	NOME	IDADE	ENDEREÇO
	Carla Soares	42	Rua Meir, 33
	Roberto L.	54	Rua Beira, 11
	João P.	35	Rua G, 67
Registro	NOME IDADE ENDEREÇO		
	Carla Soares	42	Rua Meir, 33
Campo	Carla Soares (Campo NOME)		

Figura 1.3: Hierarquia de Dados

- Um campo consiste num grupamento de caracteres. Exemplo: o nome. Pode ter também os campos idade ou endereço, por exemplo.

NOME
Carla Soares

- Um registro consiste num grupo de campos relacionados. Exemplo: o nome, a idade e o endereço da cliente Carla Soares.

Carla Soares 42 Rua Meir, 33

- Um arquivo consiste num grupo de registros do mesmo tipo. Exemplo: Arquivo cliente.

Arquivo cliente

NOME	IDADE	ENDEREÇO
Carla Soares	42	Rua Meir, 33
Roberto L.	54	Rua Beira, 11
João P.	35	Rua G, 67

- Um grupo de arquivos é chamado de Banco de Dados. Exemplo: Arquivo cliente, Arquivo pedidos e Arquivo produtos.

Banco de dados vendas

Arquivo cliente	Arquivo pedidos	Arquivo produtos
-----------------	-----------------	------------------

De igual forma, Arquivo de pedidos e Arquivo produtos contêm um grupo de registros e, em cada registro, um grupo de campos.

Agora, vamos mostrar a figura 1.3 como um conjunto de registros dispostos em estruturas regulares, ou seja, como tabelas. Observe a Figura 1.4:

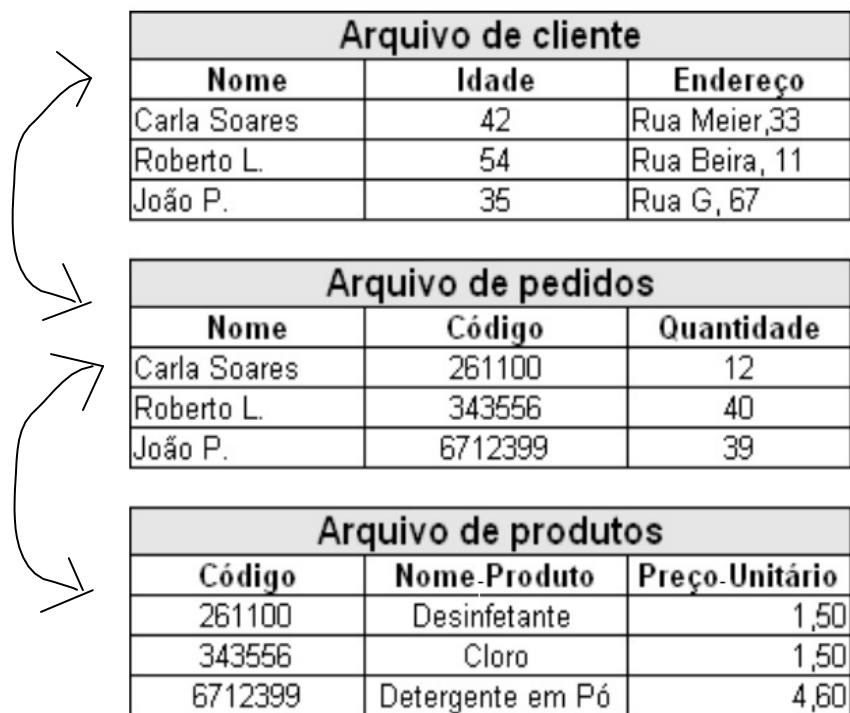


Figura 1.4. Arquivos dispostos como tabelas

Observa-se que os dados armazenados em linhas são registros e os dados armazenados em colunas são os campos. Os dados de uma tabela normalmente descrevem um assunto específico, pois a figura 1.4. retrata o caso de clientes que fazem compras de produtos através de pedidos.

E o que são essas duas setas ao lado das tabelas?

Você já vai entender: o conceito de tabelas + as setas = Banco de Dados Relacionais. Vejamos o porquê.

Um **Banco de Dados Relacional** organiza seus dados em relações. Cada relação pode ser vista como uma tabela pela qual cada coluna corresponde aos campos ou atributos da relação e as linhas correspondem aos registros ou tuplas ou elementos da relação.

Um conceito importante em um Banco de Dados Relacional é o **atributo chave**, que permite identificar e diferenciar uma tupla de outra. Através do uso de chaves é possível acelerar o acesso a elementos e estabelecer relacionamentos entre as múltiplas tabelas de um Sistema de Banco de Dados Relacional.

Essa visão de dados organizados em tabelas oferece um conceito simples e familiar para a estruturação dos dados, sendo um dos motivos do sucesso dos Sistemas de Banco de Dados Relacionais.

Atributo chave

Atributo através do qual é possível identificar determinado registro. Uma chave não pode ser repetida, ou seja, o conjunto de valores que constituem a chave deve ser único dentro de uma tabela.

Agora que você já conhece os conceitos acima descritos, que tal observar um exemplo?

Na figura 1.4. consideremos o atributo chave Nome do Arquivo cliente, pois podemos relacioná-lo com o Arquivo de pedidos. Se o atributo chave fosse Carla Soares no Arquivo de clientes, obteríamos os dados de Idade e Endereço dela. Assim, com o mesmo atributo chave poderíamos relacioná-la com o Arquivo de pedidos, e encontrarmos o atributo Código e o atributo Quantidade da cliente Carla Soares. Por outro lado, não seria possível acessar os dados do Arquivo de produtos pelo atributo chave Nome, pois tal atributo não existe.

Se acessarmos o atributo chave Carla Soares no Arquivo de pedidos, podemos obter os atributos Código e Quantidade pertencentes a ela podendo, através do atributo Código, relacioná-la com o atributo chave Código no Arquivo de produtos e, assim, obtermos os dados do Nome do produto e o Preço Unitário.

Em resumo, podemos relacionar os três arquivos através do atributo chave. No caso da cliente Carla Soares, teremos os seguintes dados:

- Idade: 42 anos;
- Endereço: Rua Meier 33;
- Um pedido de produto com o código 261100;
- Quantidades: 12;
- Nome do produto: desinfetante;
- Preço unitário: R\$1,50.

Agora, imagine o que poderia acontecer se esses Arquivos de Dados estivessem isolados? Suponhamos que o Departamento Pessoal possua o Arquivo cliente feito em

Excel, o Departamento de Contabilidade possua o Arquivo pedidos feito em Word e o Departamento Financeiro possua o Arquivo produto feito em Access. Se algum Departamento precisar de alguma informação terá que solicitá-la ao outro Departamento. Mas que confusão, não é mesmo? Além de confuso, muito tempo seria perdido até obter a informação completa sobre o cliente!

Claro que em uma empresa ou instituição real cada Arquivo acima citado possui mais atributos que o do nosso exemplo. Para evitar que os Arquivos de Dados estejam isolados e feitos em diferentes ambientes, é necessário que façamos uma **Integração de Dados**.

Veja, a seguir, como a Integração de Dados pode ser útil.

Integrado

Por integrado, queremos dizer que o Banco de Dados pode ser considerado como uma unificação de vários arquivos.

Temos mais uma definição que coloca o **banco de dado** como um conjunto de dados integrados que visa atender um conjunto de aplicações. Por conjunto de aplicações, entendemos as aplicações que são feitas para que qualquer usuário possa utilizar e acessar o **Banco de Dados**.

Exemplo: Um buscador de página *Web* é uma aplicação que recupera informação de um gigantesco Banco de Dados.

Vamos supor que temos certo tipo de produto e as informações referentes a ele, tais como vendas, compras e produção que estejam armazenadas em arquivos diferentes, como demonstrado na figura 1.5.



Figura 1.5. Dados não integrados

Nesta forma de armazenamento encontramos problemas de falta de Integração de Dados:

a) Redundância e Inconsistência dos dados:

- Arquivos de formatos diferentes;
- Programas desenvolvidos em diferentes linguagens;
- Informação repetida em diferentes lugares;
- Dados que não representam a realidade.

b) Dificuldade de extração de informação;

c) Redundância de rotinas;

d) Dificuldades no desenvolvimento de novas aplicações.

e) Problemas com acesso as informações;

Exemplo:

Item a: visto que os arquivos e programas aplicativos são criados por programadores diferentes durante um longo período de tempo, os arquivos provavelmente terão formatos diferentes e os programas serão escritos em diversas linguagens de programação.

Poderemos encontrar o mesmo elemento de informação duplicado em diversos arquivos. Suponha que o atributo código do produto e o atributo nome possam aparecer no Arquivo de Vendas e o no Arquivo de Produção. Essa redundância levará a altos custos de armazenamento e acesso. Se realizarmos a mudança do nome do produto só no Arquivo de Vendas e não no Arquivo de Produção, resultará em uma inconsistência de dados.

Item b: suponha que o gerente de produção necessite encontrar os nomes de todos os produtos produzidos em quantias maiores que 1000 unidades. Caso este tipo de lista não exista ou não tenha sido antecipada quando o sistema original foi projetado, não haverá nenhum programa aplicativo para gerar essa lista. Neste caso, o gerente tem duas saídas: ou ele pega a lista de todos os produtos produzidos e extrai a informação necessária manualmente ou pede ao departamento de processamento de dados que faça tal aplicativo. Ambas as alternativas são obviamente insatisfatórias.

Imagine que tal programa tenha sido de fato escrito e que alguns dias mais tarde, o mesmo diretor necessite destacar da mesma lista apenas os nomes de produtos produzidos com preço de custo maior que R\$ 5000,00. Novamente o gerente tem duas opções e nenhuma é satisfatória. Este tipo de ambiente convencional de processamento de arquivos não permite que os dados necessários sejam recuperados de uma maneira conveniente e eficiente.

Item c: como já mencionado, os programas aplicativos podem ser criados por programadores diferentes em diversas linguagens de programação. Então, diversos programas podem conter rotinas iguais, ou seja, ocasionando duplicidade de rotinas ou processos desnecessários.

Item d: as dificuldades no desenvolvimento de novas aplicações devem-se ao isolamento de dados, visto que os dados estão espalhados em diversos arquivos podendo ter formatos diferentes.

Item e: arquivos isolados não fornecem informação completa do cliente. Se o mesmo cliente fosse cadastrado em uma tabela por várias vezes e necessitássemos atualizá-lo ou excluí-lo, isto teria que ser feito em diversos lugares.

Então, qual a solução para evitar tais problemas? Integração de dados!

Em Bancos de Dados onde há integração dos dados, cada informação (vendas, compras e produção) é armazenada uma única vez, como se observa na figura 1.8.



Figura 1.8. Dados integrados

Por isso, o Banco de Dado é também um conjunto de dados integrados que visa atender um conjunto de aplicações.

O desenvolvimento da teoria de Banco de Dados ajudou, sobretudo, na independência de dados, ou seja, qualquer mudança na estrutura física ou na estratégia de acesso

não implica em alteração nos aplicativos que utilizam tal dado.

Vantagens dos Bancos de Dados Integrados

- Pode reduzir ou eliminar a redundância (repetição de dados);
- A inconsistência (ou seja, dados que não representam a realidade) pode ser evitada (até certo ponto);
- Pode aplicar restrições de segurança: nem todo usuário do Sistema de Banco de Dados deve ter acesso a todos os dados. Por exemplo: num sistema bancário, o departamento de pessoal necessita apenas de parte do Banco de Dados que tenha informações sobre os diversos empregados do banco. Eles não necessitam ter acesso à informação sobre as contas dos clientes do banco;
- Pode manter a integridade dos dados: garante a qualidade dos dados em um Banco de Dados. Por exemplo: se o código de um produto é introduzido com um valor de 0049 em uma tabela chamada produtos, o Banco de Dados não deve permitir que outro produto tenha um código com o mesmo valor.
- Redução da perda de espaço de armazenamento;

Um Banco de Dados com dados integrados precisa ser gerenciado e isto é feito através do **Sistema de Gerenciamento de Banco de Dados**.

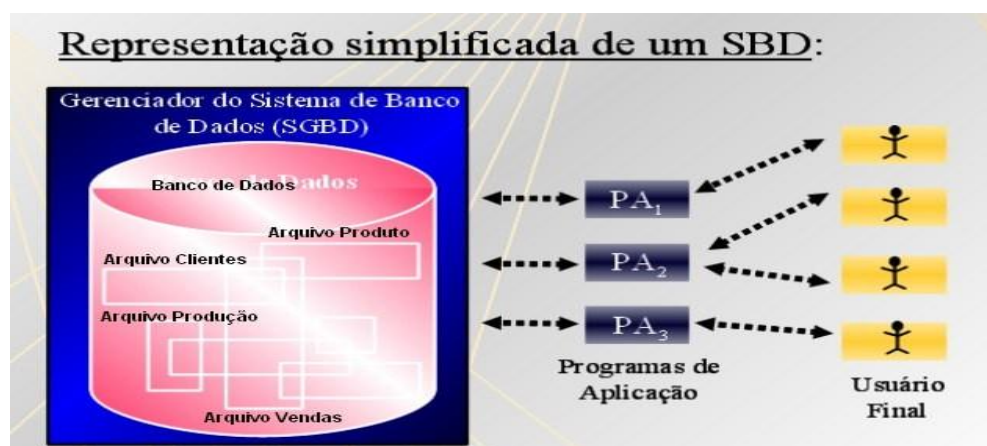


Figura 1.6. O Sistema de Banco de Dados

Sistema de Gerenciamento de Banco de Dados (SGBD)

É o software (coleção de programas) responsável pelo Banco de Dados. Operações como consultas, inserções, atualizações, exclusões, recuperação de dados, armazenamento dos dados, criação de usuários e tabelas, diretivas de segurança, etc. são controladas pelo SGBD. Além disso, o SGBD garante a integridade dos dados, cuida da segurança, retorna consultas, compartilha dados, controla a redundância, as transações, e tudo isso de forma transparente para o usuário.

Podemos citar como exemplos de SGBDs: SQL Server, Oracle, Firebird, MySQL, Interbase, entre outros. Estes programas, em geral, são chamados SGBDs relacionais.

A figura 1.6 mostra um Sistema de Banco de Dados que é composto de:

- SGBD,
- Conjunto de Programas de Aplicação,
- Conjunto de Usuário Final.

Mas você sabe quem é o Usuário Final?

Um usuário final pode ser o programador de aplicação (responsável pela escrita de programas de aplicações de Banco de Dados), o usuário que acessa o Banco de Dados a partir de uma aplicação on-line ou o administrador do Banco de Dados.

Banco de Dados e Tecnologia de Informação

Banco de Dados é o componente da **Tecnologia de Informação** voltado para o armazenamento da informação a ser utilizada em um processo de tomada de decisão. A estrutura e comportamento de um banco de dados devem propiciar esse armazenamento de forma persistente e consistente.

Tecnologia de Informação

É o conjunto de recursos não humanos dedicados à comunicação, processamento e armazenamento da informação e a maneira como esses recursos são organizados em um sistema capaz de executar um conjunto de tarefas.

Temos então que o Banco de Dados é o componente da Tecnologia de Informação, cuja função é dar suporte ao armazenamento dentro do ciclo de vida da informação. Veja a figura 1.7.

2. Transmissão da Informação;
3. Processamento da Informação;
4. Armazenamento da Informação;
5. Exibição da Informação.



Figura 1.7. Ciclo de vida da Informação

Em geral, o suporte tecnológico da informação é composto pelas seguintes tecnologias:

1. Tecnologia de Interface Homem-Máquina (Captura e Exibição da Informação);
2. Tecnologia de Transmissão (Redes de Computadores);
3. Tecnologia de Processamento (Centralizado, Distribuído, Paralelo e Cliente-Servidor);
4. Tecnologia de Armazenamento.

Portanto, a Tecnologia de Armazenamento é a Tecnologia de Banco de Dados, encarregada pelo armazenamento (e recuperação) da informação.

A estrutura e o comportamento do Banco de Dados devem propiciar o armazenamento de forma persistente e consistente. Mas isso veremos mais adiante.

Evolução: Tecnologia de Informação em Banco de Dados

- Até 1960: Sistema de Arquivos integrados. Ex.: ISAM, VSAM;
- Final de 1960: Modelo Hierárquico. Ex.: IMS (IBM);
- 1970 e início de 1980: Modelo de Redes. (CODASYL) Ex.: IDMS, DMS-II(Unisys);

- Meados 1980: Modelo Relacional. Ex.: DB-2, SQL-DS (IBM), Oracle, Ingres;
- Final de 1980: Modelo Orientado a Objetos e Relacional Estendido (Objeto-Relacional) Ex.: BDOO: Vbase, O2, Orion, Gemstone, Jasmine, ObjectStore. BDOR: Postgres, Informix, Oracle 9i, IBM DB2;
- 1990: BD Inteligentes e Cliente-Servidor, BD e Web, BD multimídia, BD Espacial, Ativo, Temporal, Dedutivo;
- 2000: Modelo Objeto Relacional, BD de WWW, BD Mobile, SQL/MM.

Agora que já vimos diversos conceitos importantes sobre Banco de Dados, vamos colocá-los em prática através das seguintes atividades.

Para se ter sucesso... é preciso planejar!

Muitos dos problemas enfrentados no dia-a-dia de qualquer *empresa informatizada* resultam de um projeto “mal feito” ou “mal implementado”. A modelagem de banco de dados é um tema que nunca sai de moda. Independente do sistema, um banco de dados sempre deve ser iniciado por um bom projeto e uma boa modelagem, pois é justamente a partir deste ponto que será garantida a confiabilidade, eficiência e eficácia do sistema.

O Projeto de um Banco de Dados

A modelagem de banco de dados é estudada há mais de três décadas e vem sendo, ao longo desse tempo, definida por diversos autores. Apresentamos as seguintes definições:

- É o processo que determina a organização de um banco de dados, incluindo a sua estrutura, conteúdo e aplicações.
- É o processo de projeto da estrutura lógica e física de um ou mais bancos de dados, visando proporcionar as informações necessárias de um determinado conjunto de aplicações aos usuários de uma organização.

Mas por que devemos usar modelagem de dados?

- Proporciona informação concisa dos dados necessários para o negócio;
- Modelo de Dados independe de hardware e software;
- Identifica redundâncias;
- Identifica dados não utilizados;

- Permite estabelecer e manter regras de integridade dos dados;
- É um auxiliar precioso na concepção do banco de dados;
- É fundamental para um bom desempenho do sistema;
- Os dados devem estar organizados de modo a serem flexíveis para responder às necessidades dos usuários.

Obter dados com estas características é o principal objetivo da modelagem de dados. A modelagem do banco de dados é extremamente importante para que o sistema tenha uma boa performance.

Estudos indicam que quanto maior o tempo despendido no Projeto de Banco de Dados, menor será o tempo utilizado em sua manutenção. Se a devida atenção não for dada ao desenho do banco de dados, todo o desenvolvimento do sistema poderá ser comprometido. Um banco de dados bem projetado fornece um acesso conveniente a todas informações desejadas. Com uma boa estrutura, assegura-se resultados mais rápidos e precisos.

O Projeto de um Banco de Dados é um processo complexo que visa atingir algumas metas:

- Satisfazer os requisitos de informações especificadas por usuários e aplicações;
- Proporcionar uma estruturação natural e fácil para entender a informação;
- Dar suporte a quaisquer requisitos de processo e objetivos de desempenho, como tempo de reposta, tempo de processamento e espaço de armazenamento.

Para o cumprimento das metas mencionadas, o processo é dividido em diferentes fases: Projeto Conceitual, Projeto Lógico e Projeto Físico. Esta complexidade é melhor gerenciada quando o problema é subdividido em vários subproblemas independentes. O fato central nas três fases do Projeto de Banco de Dados é a modelagem do dado e suas propriedades.

Projeto (Modelagem) de Banco de Dados

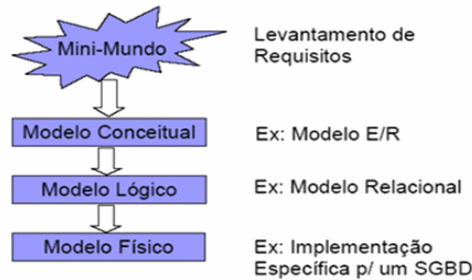


Figura 1.8. Projeto ou Modelagem de Banco de Dados

Modelo conceitual: É também uma descrição da estrutura do Banco de Dados, porém, de forma independente da implementação que será feita. Em outras palavras, ele independe de qual SGBD será utilizado na implementação. O Modelo Conceitual indica quais os dados que poderão aparecer no Banco de Dados, mas não informa de que forma estes mesmos dados serão armazenados no nível SGBD.

Modelo lógico: Este modelo descreve a estrutura do Banco de Dados no nível do usuário SGBD, neste caso, o administrador de dados. Este modelo já se aproxima mais da implementação que será feita, ou seja, torna-se dependente de qual SGBD será implementado.

Modelo físico: Este modelo é o último passo antes da geração de *scripts* de implementação. Ele é totalmente dependente do SGBD específico que será utilizado. Além das definições de chave (que já estão presentes no modelo lógico), o modelo físico contempla definições de armazenamento que não têm influência alguma nas etapas anteriores, mas que são essenciais no tocante à performance geral do Banco de Dados.

Conclusão

Os Bancos de Dados Relacionais são, atualmente, o tipo Banco de Dados mais utilizado no mundo. Um Banco de Dados Relacional organiza seus dados em *relações*. Cada relação pode ser vista como uma tabela pela qual cada coluna corresponde aos campos ou *atributos* da relação e as linhas correspondem aos registros, às *tuplas* ou elementos da relação.

Um conceito importante em um Banco de Dados Relacional é o conceito de atributo

chave, que permite identificar e diferenciar uma tupla de outra. Através do uso de chaves é possível acelerar o acesso a elementos e estabelecer relacionamentos entre as múltiplas tabelas de um Sistema de Banco de Dados Relacional. O desafio mais significativo ao implementar um Banco de Dados é desenvolver sua estrutura e comportamento corretamente.

Resumo

Vamos rever os principais conceitos vistos nesta aula:

- Dado: é qualquer elemento identificado em sua forma bruta que, por si só, não conduz a uma compreensão de determinado fato ou situação;
- Informação: é o componente que, efetivamente, transita entre os processos organizacionais e, dessa forma, estabelece relacionamentos entre os mesmos;
- Podemos definir o termo Banco de Dados como: "Um conjunto de arquivos relacionados". O termo arquivo é relacionado como uma tabela pela qual as colunas são campos e as linhas são os registros;
- Um Banco de Dados Relacional é aquele que: 1. Organiza seus dados como uma tabela pela qual cada coluna corresponde a atributos da relação e cada linha corresponde às tuplas ou elementos da relação e 2. Possui um atributo chave que serve para acessar elementos e estabelecer relacionamentos entre as múltiplas tabelas;
- O papel da Tecnologia da Informação é dar suporte à Informação em todo o seu ciclo de vida compreendendo: Captura, Transmissão, Processamento, Armazenamento e Exibição da Informação:
Dentro desse contexto, Banco de Dados é o componente da Tecnologia da Informação voltado para o Armazenamento da Informação;
- O projeto de banco de dados é constituído de 3 fases: modelagem conceitual, lógica e física.

Referências Bibliográficas

Date C.J. , 20034. Introdução a Sistemas de Bancos de Dados. 8ed. Americana. Rio de Janeiro. Elsevier.

Setzer V.W. & Corrêa da Silva F.S. 2005. Bancos de Dados. 1ed. São Paulo. Edgard Blucher.