

INTRODUÇÃO À TECNOLOGIA DE BANCOS DE DADOS

Conceitos básicos em bancos de dados

Desde que surgiu o conceito de banco de dados na ciência da computação, bem como toda a tecnologia que vem sendo desenvolvida associada a ele, essa área de dados tornou-se muito relevante para o crescimento do uso de computadores nas organizações, pois é uma das mais aplicadas dessa ciência. Por ser bastante aplicada, é uma área que fica em contato direto com os usuários finais dos sistemas e, portanto, é popular, diferentemente de outras áreas menos acessíveis da computação. Praticamente, todos os sistemas desenvolvidos para as mais diferentes aplicações utilizam algum tipo de banco de dados. Por ser popular, muitas vezes as pessoas empregam o termo “banco de dados” de maneira equivocada, pois ele é comumente utilizado para definir três diferentes conceitos relacionados a essa tecnologia. A seguir, esses conceitos serão apresentados com a intenção de esclarecer alguns pontos e evitar que sejam confundidos:

1. **Banco de Dados (BD):** este primeiro conceito é o mais simples de todos e representa um conjunto de dados relacionados e acessíveis. Dados são fatos do mundo real, que podem ser registrados e organizados. Considere, por exemplo, nomes, telefones e e-mails de pessoas que você conhece e que estão armazenados em sua agenda do celular. Esses dados estão relacionados entre si, afinal você só gravará dados de pessoas que conhece ou que de alguma forma lhe interessam ou possuem algum significado para você. Esses dados de pessoas estão armazenados em algum dispositivo de onde você pode acessar a qualquer momento que desejar. A agenda é um banco de dados.

2. **Sistema Gerenciador de BD (SGBD ou Database Management System – DBMS):** este segundo conceito representa um tipo de sistema composto por uma série de ferramentas criadas especificamente para gerenciar os bancos de dados, e que apresentam linguagens utilizadas para manter os bancos de dados.



3. **Sistema de BD:** já este terceiro conceito representa todo sistema desenvolvido com funções ou aplicações específicas, que utilizam bancos de dados e, por sua vez, foram desenvolvidos em algum SGBD.

É sinônimo de sistema de informação.

Segundo Elmasri e Navathe (2018), um banco de dados possui algumas propriedades implícitas:

- Representa aspectos de uma parte do mundo real, chamada de Mini Mundo (ou Universo de Discurso);
- é uma coleção lógica e coerente com algum significado inerente;
- é projetado, construído e povoado por dados, atendendo a uma finalidade específica, para ser usado por pessoas específicas e por aplicações específicas.

A Figura representa o processo de modelagem do Mini Mundo, que é algo irregular, num banco de dados que deve ser algo bastante estruturado. A tarefa de modelar um banco de dados a partir de uma parte do mundo real é apenas a primeira de muitas tarefas que compõem o processo de construção de um banco de dados.

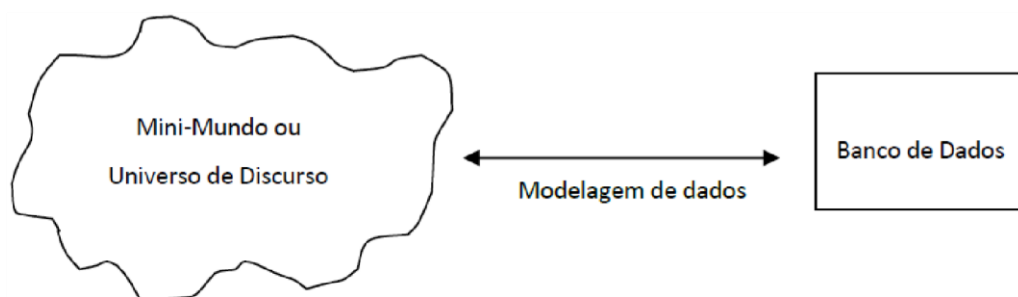


Figura: Ilustração da modelagem de dados entre o mundo real e o BD.



De forma bastante simplificada, o processo de criação de um BD poderia ser explicado da seguinte maneira: a partir da descrição detalhada dos requisitos de dados do sistema, feito pelo analista de sistemas, é construído o primeiro modelo conceitual de dados. Posteriormente, o projetista de dados faz a construção do projeto físico de dados, para somente depois o banco de dados ser de fato implementado pela equipe de programadores num SGBD. Todo esse processo será devidamente explicado ao longo da disciplina.

A importância dos bancos de dados nas organizações

Os dados estão presentes na vida dos seres humanos desde que surgiu a linguagem e começamos a nos comunicar e registrar os fatos. Desde que começamos a nomear objetos e diferenciar as pessoas por seus nomes, começamos a atribuir diferentes dados ou diferentes valores a esses objetos e pessoas. Essa diferenciação feita aos objetos e pessoas do mundo precisa de alguma forma estar organizada para ser compreendida.

Por muitos anos, a organização dos dados foi feita pelos seres humanos e registrada de forma manual. Inicialmente por desenhos nas paredes das cavernas, posteriormente em papel, até finalmente chegar ao registro automatizado dos dados nos computadores, como é feito hoje. Mas desde que surgiu algum tipo de registro e organização dos dados, e que estes estiveram ao acesso de outras pessoas, podemos dizer que surgiu um arquivo ou um banco de dados.

O fato é que o imenso volume de dados com o qual atualmente as pessoas e as empresas precisam lidar é imensamente maior do que há alguns anos.

Segundo Elmasri e Navathe (2018), os bancos de dados são um componente essencial na sociedade moderna, pois a maioria das pessoas encontra diariamente diversas atividades que envolvam alguma interação com algum banco de dados, por exemplo, quando vamos sacar dinheiro no caixa eletrônico, ou fazer uma compra no supermercado, ou até quando vamos fazer uma consulta online sobre um determinado assunto.

Já nas organizações, de uma maneira geral, praticamente todos os processos passam por algum tipo de sistema de informação computacional, que por sua vez interage com um banco de dados. Segundo Medeiros (2013), os sistemas de



informação estão na atualidade profundamente arraigados nas empresas. Como estamos na era da informação, portanto, não poderia ser diferente.

É muito difícil nos dias atuais encontrarmos uma empresa que ainda contenha algum tipo de processo ou procedimento que possa ser realizado de maneira independente de um sistema computacional, mesmo porque há a necessidade de um controle de todas as ações dos funcionários e isso se dá através do monitoramento dos sistemas. Isso faz com que uma imensa quantidade de dados seja gerada instantaneamente e diariamente no mundo e esse movimento ininterrupto vai acumulando um volume gigantesco de dados.

Além disso, todas as informações geradas nas últimas décadas das organizações ficaram armazenadas nos sistemas das empresas para fins de dados históricos, para construção de Data Warehouses, dentre outras aplicações. Isso ampliou ainda mais o enorme volume de dados armazenados. Diante desse cenário, podemos então perceber que atualmente temos uma quantidade de dados armazenados como nunca houve na história da humanidade e muito provavelmente essa quantidade tende a continuar aumentando nos próximos anos.

A tecnologia que apoia e mantém essa gigantesca quantidade de dados é a tecnologia de bancos de dados, que evoluiu desde seu surgimento, mas que se manteve em sua estrutura inicial de forma a simplificar ao máximo o trabalho das pessoas envolvidas na área de gerenciar e manipular os dados.

Diante desse cenário, onde percebemos nitidamente a importância que as informações (e consequentemente os dados) têm nas nossas vidas, faz-se necessário conceituar claramente o que é um banco de dados. Nas próximas seções, entenderemos os principais conceitos dessa tecnologia e quem são as pessoas diretamente envolvidas nessa área.

O ambiente de um sistema de BD

Para melhor compreendermos os conceitos básicos associados à tecnologia de Banco de Dados (BD), vale observarmos a Figura, que representa de forma resumida o ambiente de um sistema de banco de dados.



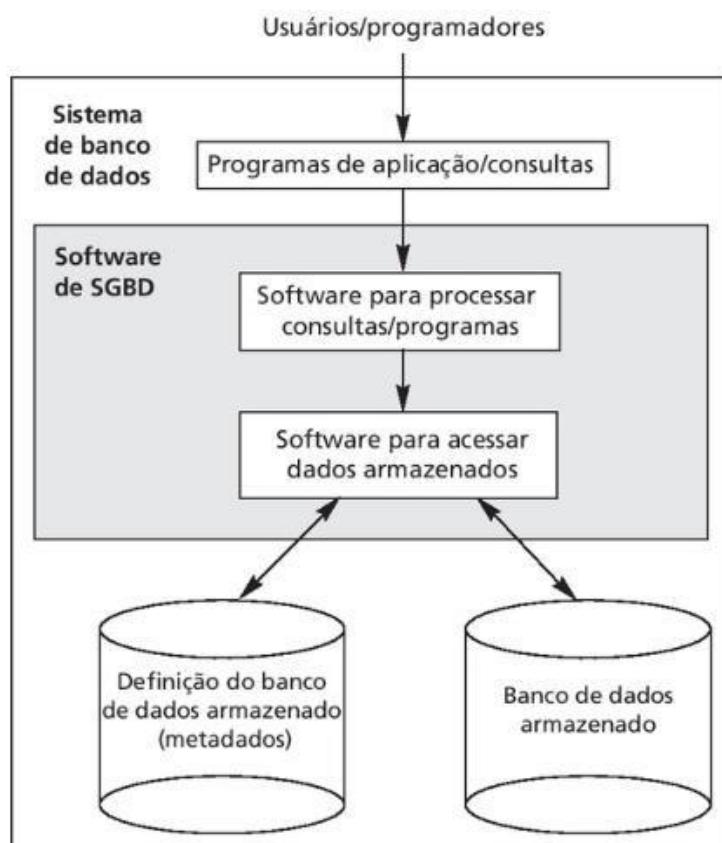


Figura: Diagrama simplificado de um ambiente de sistema de banco de dados.

Fonte: Elmasri e Navathe (2018).

Interpretando a Figura, podemos observar que as pessoas (programadores e usuários) interagem com os sistemas de bancos de dados através de programas de aplicação. Esses programas de aplicação podem ser construídos pelos programadores em diferentes tipos de linguagem de programação, para serem usados pelos usuários em diferentes tipos de dispositivos, para diferentes ambientes da empresa. Por exemplo, pode acessar o mesmo sistema de banco de dados uma aplicação voltada para os funcionários que é acessada somente na intranet da empresa, outra aplicação voltada para os clientes pode ser acessada através da internet pelo site da empresa, dentre outras.

Os programas de aplicação realizam consultas (do inglês, queries) no BD. Uma query executada no BD irá realizar alguma ação de manipulação nos dados,



por exemplo: ler (recuperar) os dados que estão cadastrados no banco, ou alterar (inserir, remover ou modificar) os dados.

Podemos também observar que há uma camada intermediária entre as aplicações e os BD propriamente ditos, que é o Sistema Gerenciador de BD (SGBD). O SGBD recebe as consultas feitas pelos usuários, enviadas pelas aplicações e as processa, repassando a programas específicos a função de buscar os dados desejados nos bancos de dados.

De maneira simplificada, podemos separar os arquivos que armazenam os dados em dois grandes grupos: num primeiro estão os dados armazenados, ou os dados propriamente ditos, ou seja, os valores relativos aos fatos do mundo real que foram cadastrados; num segundo grupo ficam contidas as definições dos dados ou os metadados, que são informações sobre a estrutura do BD, tais como qual é o tipo de valor que aquele dado pode conter (numérico ou alfabético). Essa separação é feita porque a estrutura dos dados, uma vez definida na fase de projeto do BD, pouco irá mudar ao longo da utilização desse sistema de banco de dados, enquanto os dados armazenados são alterados constantemente no dia a dia das organizações. Por esse motivo, o SGBD gerencia e mantém esses dois grupos de arquivo de maneira diferente.

O mais importante de fixar dessa imagem que representa o ambiente de um sistema de banco de dados é a importância que os SGBD têm para facilitar o desenvolvimento de aplicações que acessam os BDs, uma vez que fazem todo o processamento intermediário entre as consultas enviadas nas diversas linguagens de programação e os BDs armazenados.

Características básicas dos SGBD

Antes de serem criados os primeiros SGBD, pelo menos até a metade da década de 1970, os sistemas computacionais existentes utilizavam, de forma muito restrita, as estruturas de arquivos disponíveis nos Sistemas Operacionais (SO) da época para gerenciar os seus dados. Essa era uma tarefa muito complexa e restrita, pois se tinha que conhecer muito bem o sistema de gerenciamento de arquivos e de memória do SO, e para cada nova aplicação que surgisse, toda a lógica de manipulação dos dados tinha que ser refeita. Então, quando os primeiros SGBD



foram implementados, eles tinham de cumprir o objetivo de solucionar os principais problemas enfrentados pelos programadores da época.

Assim, segundo Elmasri e Navathe (2018), os SGBD foram criados contendo as seguintes principais características, que são mantidas até os dias atuais:

- **Natureza autodescritiva de um sistema de BD:** visava solucionar o problema da dependência dos sistemas de BD desenvolvidos em linguagens de programação da época de outros sistemas, tais como o Sistema Operacional. Desejava-se um sistema que fosse independente dos demais e, portanto, ele precisava conter nele mesmo todas as informações necessárias para seu funcionamento. A tecnologia de BD para solucionar esse problema implementou o chamado **catálogo** do banco de dados, que armazena o esquema do banco e todas as informações necessárias para funcionamento do SGBD, sem dependência de outros tipos de programas. A informação armazenada no catálogo do BD é chamada de **metadados** (informações sobre os dados). Com todas as informações necessárias para seu funcionamento contidas no próprio SGBD, ressaltando a sua natureza autodescritiva, esse sistema tornou-se independente dos demais.

- **Isolamento entre programas e dados, e abstração de dados:** antes dos SGBD, no processamento tradicional que utilizava os arquivos, a estrutura do arquivo, sua localização física e outros detalhes de armazenamento estavam embutidos nos programas de aplicação. Assim, qualquer mudança na estrutura dos dados refletia mudanças em todos os programas de aplicação. Essa era uma tarefa muito complexa e exigia muito esforço da equipe de desenvolvimento de software. Já os SGBD oferecem uma **abstração dos dados**, armazenando as estruturas (ou definição) dos dados no seu catálogo, enquanto os dados propriamente ditos ficam armazenados separadamente. Quando a estrutura dos dados muda, apenas uma parte do SGBD é acionada para ser modificada. Já todos os programas de aplicação que utilizam os bancos de dados não são alterados. Isso facilita muito o trabalho da equipe de desenvolvimento de software e cria um isolamento entre os programas e os dados propriamente ditos.

- **Conceito de abstração dos dados:** um modelo de dados é usado para esconder detalhes de armazenamento, com uma visão conceitual do BD (ELMASRI e NAVATHE, 2018).



- **Suporte a múltiplas visões dos dados:** um banco de dados típico tem muitos usuários e cada qual possui interesses diferentes, que podem solicitar diferentes perspectivas ou **visões** dos dados no BD. O SGBD é capaz de suportar diferentes visões dos dados, a depender do usuário, somente as que interessam. Isso é importante tanto para simplificar para o usuário quanto por motivos de segurança, uma vez que certos dados não serão disponibilizados a usuários não autorizados. Uma visão (ou do inglês, view) pode ser um subconjunto dos dados do BD ou conter dados virtuais, temporários e derivados de arquivos físicos do BD, mas que não estão armazenados explicitamente.

- **Compartilhamento de dados e processamento de transação multiusuário:** um SGBD multiusuário é aquele que permite o acesso ao BD, feito por diversos usuários diferentes, ao mesmo tempo (concorrentemente), ou seja, os usuários irão concorrer pelos recursos daquele BD. Para que esse compartilhamento de dados entre os usuários funcione sem problemas, o SGBD deverá realizar um controle dessa concorrência, enquanto realiza o processamento das várias transações que são enviadas ao mesmo tempo. Então, todo SGBD multiusuário possui um módulo de controle de concorrência para garantir que as transações sejam executadas da maneira correta e eficiente.

Pessoas envolvidas com BD

Como já discutimos, os BDs estão em toda parte, sendo utilizados por praticamente todas as pessoas de alguma forma. Nesta seção, entretanto, estudaremos as pessoas que estão mais diretamente envolvidas com o processo de criação, manipulação e gerenciamento dos bancos de dados nas organizações.

Para um pequeno BD individual, como no exemplo da agenda de contatos do seu celular, em geral apenas uma pessoa constrói e manipula os dados, não havendo o seu compartilhamento com outras pessoas. Porém nas organizações, onde centenas ou milhares de usuários irão acessar o BD, geralmente existe um grupo grande de pessoas envolvidas com essa tecnologia e que trabalham em seu dia a dia com os bancos de dados.



Segundo Elmasri e Navathe (2018), as pessoas mais diretamente ligadas ao trabalho de projetar, criar, implementar, gerenciar e manipular os dados são apresentadas a seguir.

Os **usuários finais** são pessoas cujas profissões requerem o acesso constante a um sistema de BD para consulta, atualização e emissão de relatórios. Esse acesso se dá através do uso de sistemas de aplicação, que são construídos sobre os SGBD, ou seja, os usuários finais não irão acessar os bancos de dados diretamente pelo SGBD, mas sim por sistemas construídos para eles. A aplicação e o BD existem prioritariamente para os seus usuários, por isso eles são muito importantes e considerados os verdadeiros “donos” do BD. Quando os projetistas e analistas de BD estão realizando o seu trabalho, eles interagem com esses usuários finais para conhecê-los e devem levar em consideração o fato de que, muitas vezes, os usuários são pessoas com pouca familiaridade com a tecnologia de BD. O diálogo entre o analista de sistemas e os usuários finais deve ser feito da maneira mais clara e facilitada possível, para que as necessidades informacionais dos usuários sejam traduzidas corretamente como os requisitos de dados do sistema.

Os **analistas de sistemas** e os **programadores de aplicações** (engenheiros de software) são as pessoas que determinam e implementam as especificações das transações customizadas para atender às necessidades dos usuários. O trabalho de análise do sistema tem início no diálogo entre os usuários e o analista, que resulta numa lista de requisitos de dados do sistema. Esses requisitos serão utilizados para a construção do primeiro modelo abstrato dos dados, o modelo conceitual, que é a primeira etapa no processo de criação de um BD. Em paralelo a esse trabalho, também são projetadas as transações que serão programadas e, posteriormente, testadas e documentadas, até serem colocadas em ação no sistema, quando se inicia a fase de utilização e manutenção. Uma vez que o sistema esteja sendo utilizado, enquanto ele existir, sempre terá que ser mantido pela equipe de engenharia de software.

Os **projetistas do banco de dados** são responsáveis pela identificação dos dados que serão armazenados no BD e também por escolher as estruturas apropriadas para representar esses dados (modelo físico). Eles recebem o modelo conceitual criado pelos analistas de sistemas que o transformam num modelo físico, geralmente utilizando tabelas, já pensando em como esse BD será implementado



num SGBD. Os projetistas também interagem com os potenciais usuários dos sistemas de BD e criam as visões que atendem aos requisitos desses usuários. O projeto final do BD, que é o resultado final do trabalho dos projetistas, deve ser capaz de suportar todos os requisitos de todos os grupos de usuários dos sistemas. Até esse momento, o BD ainda não foi de fato criado no SGBD, ele ainda está em fase de projeto.

O **administrador de bancos de dados**, ou como é mais conhecido por sua sigla DBA (do inglês, DataBase Administrator), é o administrador-chefe de todos os recursos que envolvem o banco de dados, os SGBD e outros sistemas relacionados. O DBA é responsável pela autorização para o acesso ao BD, por questões de segurança e desempenho do SGBD, pela coordenação e monitoração do uso do BD, por adquirir recursos de software e hardware conforme necessidade, e por mais uma longa série de tarefas. Por ser a pessoa mais importante na área de BD nas organizações, o DBA será estudado com mais detalhes a seguir. Nós vimos aqui uma lista de pessoas que interagem diretamente com a tecnologia de bancos de dados e que realizam tarefas distintas, porém é muito comum encontrar em empresas, principalmente de pequeno e médio porte, uma única pessoa desempenhando os papéis de todas essas diferentes funções. Nesse caso, é muito importante que essa pessoa saiba diferenciar cada uma das tarefas que estiver desempenhando no momento adequado, seguindo as etapas do processo de criação do BD corretamente, sem pular ou eliminar nenhuma etapa, para garantir a qualidade final do BD criado.

Resumidamente, a função de administrar um BD comporta as tarefas de: instalar, configurar, monitorar e solucionar problemas que possam surgir com a utilização de um SGBD específico. Geralmente, para ser considerado um DBA de um SGBD específico, o mercado de trabalho exige uma certificação específica concedida pela empresa fabricante do SGBD.

Detalhando um pouco mais as funções desempenhadas no cotidiano de um DBA, temos: a criação, ou pelo menos, o conhecimento do esquema lógico do BD; a definição de checagem de segurança e integridade dos dados; a tomada de decisões de como os dados são representados na base de dados armazenados; a definição, ou pelo menos, o conhecimento do projeto físico do BD; a definição de procedimentos de recuperação dos dados em caso da ocorrência de falhas; o



monitoramento diário das ações dos usuários e o monitoramento diário do desempenho do SGBD; o contato com usuários para averiguação de disponibilidade dos dados por eles requisitados e ajuda na determinação e resolução de problemas; e os ajustes apropriados à medida que ocorram mudanças de requisitos.

