Estratégia Concursos

Aula 00

Banco de Dados p/ PG-DF (Analista - Suporte e Infraestrutura) Com Videoaulas- Pós-Edital

Autor:

Equipe Informática e TI, Thiago Rodrigues Cavalcanti

27 de Dezembro de 2019

Banco de Dados. 3.1 Arquitetura. 3.2 Estrutura. 3.3 Administração de Banco de Dados 2
Apresentação do Professor 2
Motivação para o curso 3
Objetivos da aula 5
Fundamentos de Gestão do Conhecimento 5
Introdução a gestão do conhecimento 5
Dado, informação, conhecimento e inteligência. 7
Conhecimento tácito e explícito 13
Dados estruturados e não estruturados. 14
Conceitos de Banco de dados 21
Conceitos básicos 21
Características da abordagem de BD 26
Personagem do ecossistema de BD 30
Evolução histórica dos SGBDs 33
Modelo de dados e Arquitetura em três esquemas 42
Modelo de Dados, instâncias, esquemas 42
Arquitetura três esquemas 44
Ambiente do Sistema de banco de dados 49
Questões comentadas 53
Exercícios 60
Gabarito 68
Resumo 69
Mapa Mental 71
Considerações finais 72
Referências 72

THIAGO CAVALCANTI
PROFESSOR

BANCO DE DADOS. 3.1 ARQUITETURA. 3.2 ESTRUTURA. 3.3 ADMINISTRAÇÃO DE BANCO DE DADOS

APRESENTAÇÃO DO PROFESSOR

Olá senhoras e senhores,

Sejam bem-vindos a mais um curso de Tecnologia da Informação! Hoje damos início ao mais completo curso para o concurso do **Procuradoria Geral do Distrito Federal - PGDF -** no que se refere Banco de dados e Business Intelligence. Gosto sempre de dizer que é um prazer imenso fazer parte desta equipe de professores do Estratégia Concursos e ter a oportunidade de apresentar um pouco do meu conhecimento e experiência em concursos públicos!

Antes de começar de fato o conteúdo teórico desta aula, vou apresentar de forma rápida. Meu nome é Thiago, sou casado, pernambucano, tenho dois filhos, Vinícius (8 anos) e Lucas (*in memoriam*). Torço pelo Sport Clube do Recife. Sou cristão. Frequento a IPN – Igreja Presbiteriana Nacional. Formei-me em Ciência da Computação pela UFPE. Tenho mestrado em engenharia de software na mesma instituição. Atualmente faço doutorado em economia na UnB.

Frequento academia para manter a forma, mas meu hobby mesmo é pedalar! Decidi vender o carro e viver num desafio intermodal de transporte. Ia para o trabalho de *bike* sempre que possível! Ultimamente tenho usado mais Uber do que a magrela, mais isso é um detalhe! A pergunta é: onde eu trabalho? No Banco Central do Brasil!

Fruto de uma trajetória de dois anos de estudos diários. Aposentei as canetas em 2010. Hoje estou de licença do Banco Central para fazer doutorado que começou em março de 2017. Antes de me licenciar eu trabalhava com análise e modelagem de dados.

Minha mais recente experiência com gestão de dados é parte de uma estratégia profissional de alinhar meu trabalho diário como servidor público com minha carreira paralela de professor e consultor de Banco de Dados (BD) e *Business Intelligence* (BI). A ideia é conseguir me especializar cada vez mais no tema desta nova carreira dentro da TI, que o mercado está denominando de **cientista dos dados (*Data scientist*)**.

Entrei neste universo de professor de concurso há alguns anos. Desde 2012, tenho me dedicado especificamente ao conteúdo de BD e BI. Minhas experiências em cursos presencias aqui em Brasília e em diversas partes do Brasil, bem como na gravação sistemática de aulas on-line me ajudaram a desenvolver um conteúdo exclusivo para os alunos do Estratégia Concursos.

A ideia é desenvolver um material completo, recheado de questões e com diversas dicas para ajudar você no seu objetivo: **ser aprovado e nomeado**!



Para finalizar, não deixe de seguir minha página no Facebook® (profthiagocavalcanti), onde eu publico, sistematicamente, questões comentadas e dicas semanais. Tenho também uma conta no Instagram, lá eu posto motivações e dicas rápidas a respeito do conteúdo de banco de dados e **análise de informações**. Agora que você já me conhece! Vamos seguir em frente com o nosso curso!

Para facilitar usa vida você pode usar os QR codes ao lado para acessar meu perfil nas redes sociais. Se precisar falar comigo por e-mail, mande mensagem para:

rcthiago@gmail.com

MOTIVAÇÃO PARA O CURSO

Preparar esse curso é um desafio! Consolidar de forma amigável o conhecimento de banco de dados, análise de informações ou business Intelligence para concursos não é uma tarefa fácil! Calibrar o nível do teórico associado a uma didática eficiente tem sido minha meta nos últimos anos. Separamos o conteúdo de forma a segmentar e impulsionar seu aprendizado. Para que você entre na primeira aula com um pouco mais de segurança, vou aproveitar para fazer uma rápida apresentação sobre o assunto.

Você já ouviu falar sobre **Data Science ou ciência dos dados**? É um conceito relativamente recente que agrupa diversas atividades executadas sobre um conjunto de dados. Em especial sobre grandes conjuntos de dados. Para analisar os dados eles precisam estar **armazenados e organizados** de maneira **convenientes** para os cientistas dos dados. Essa base de dados facilita o trabalho e o entendimento do conteúdo armazenado.

Cientistas de dados são uma nova geração de especialistas em análise que têm habilidades técnicas para resolver problemas complexos e a curiosidade de explorar quais são os problemas que precisam ser resolvidos. A solução desses problemas passa por analisar os dados presentes em um banco de dados. Neste curso veremos o passo-a-passo para construção de um banco de dados.

Nossa primeira aula deve inserir você no universo dos bancos de dados. Um banco pode ser visto como uma estrutura que armazena algo, por exemplo, um banco de leite guarda leite guarda leite materno para que possa ser reutilizado de forma adequada em momentos posteriores. Um banco de dados guarda dados. Esses devem ser controlados de forma adequada. É nesse momento que surge um sistema para "cuidar" do acesso consistente aos dados.

Os sistemas de gerenciamento de banco de dados (SGBDs) contribuem para a disponibilidade de um conjunto de informações para diferentes usuários simultaneamente. É preciso decidir quais dados armazenar, estruturar e manter na base de dados. Para controlar esse sistema e todo o

desenvolvimento do projeto e da infraestrutura associada ao sistema de banco de dados várias tarefas têm que ser feitas.

Veremos que existem profissionais dedicados a tarefas específicas. Veremos ainda que a construção de um banco de dados, em especial um banco de dados relacional, passa por algumas etapas bem definidas. Essas etapas criam modelos de dados ou esquemas que permitem um melhor entendimento da estrutura de dados da organização.

Todos esses conceitos serão vistos em detalhes nas próximas páginas. Ao final, teremos nossa tradicional lista de exercícios. Espero conseguir contribuir para a sua aprovação. Vamos em frente?!

Teremos muito trabalho! Por isso, montamos um **curso teórico em PDF,** baseado nas mais diversas bancas, em especial do **CESPE**, apresentando o conteúdo observando as variadas formas de cobrança do mesmo pelas bancas examinadoras.

Teremos ainda videoaulas que apresentam o conteúdo teórico de forma detalhada para algumas partes da matéria. Existe uma força tarefa para gravação de todo o assunto, mas não temos como garantir o término deste trabalho até a data de publicação das aulas. Mas não se preocupe, nosso objetivo é garantir que você tenha capacidade e conhecimento para ser aprovado. Logo, todo conteúdo necessário para a prova estará presente nos PDFs. Vamos seguir em frente para a nossa primeira aula que tratará dos conceitos básicos de banco de dados.

OBJETIVOS DA AULA

Nossa aula começa com os objetivos. Ao final dela, ao ler o resumo, você deve ter certeza de que fixou os principais conceitos associados aos seguintes objetivos:

- Conhecer os fundamentos de gestão do conhecimento e fazer a diferença entre dados, informação, conhecimento e inteligência.
- Entender o que é um banco de dados e por que sistemas de banco de dados são uteis.
- Conhecer uma arquitetura genérica para sistemas de banco de dados denominada ANSI/SPARC.
- Descrever o ambiente de banco de dados em função dos diversos componentes.

Falaremos ainda sobre os personagens que se envolvem no funcionamento diário de um banco de dados e traçaremos uma rápida história dos Sistemas de Gerenciamento de Banco de dados.

FUNDAMENTOS DE GESTÃO DO CONHECIMENTO

INTRODUÇÃO A GESTÃO DO CONHECIMENTO

Nas próximas páginas, vamos fazer uma introdução ao assunto de gestão do conhecimento. Esse assunto faz parte da motivação para a utilização dos sistemas de inteligência de negócio que veremos em seguida (nas próximas aulas do nosso curso). A gestão do conhecimento tem por **objetivo fornecer ou aperfeiçoar a capacidade intelectual** das organizações.

Não se trata apenas da criação de uma base de dados central que contenha a descrição das experiências e informações da empresa. Podemos encontrar nas diversas definições de gestão do conhecimento aspectos diferentes mais complementares sobre o tema.

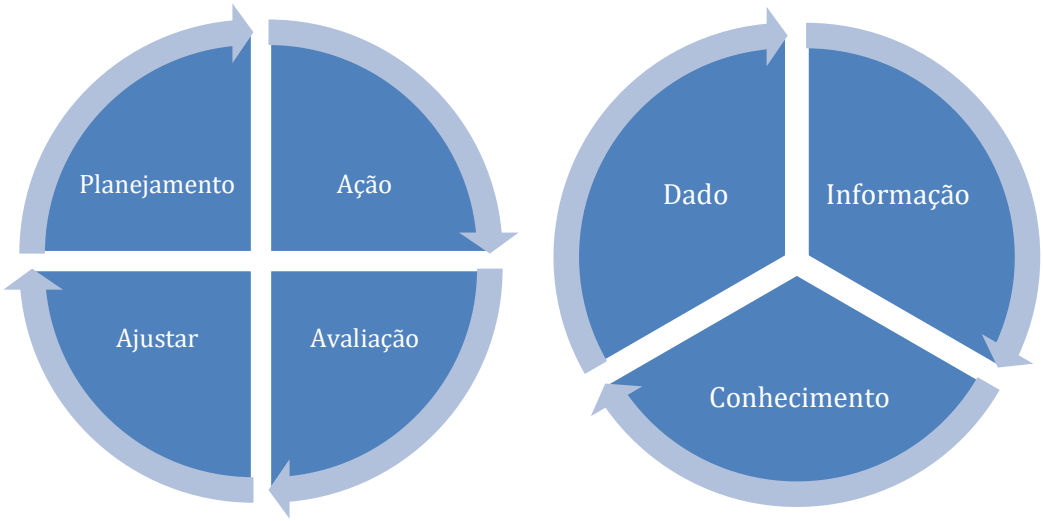
Hamilton afirma que é um processo de criação, transferência e uso de conhecimento com o objetivo de **melhorar o desempenho da organização**. Davenport descreve como um processo composto das seguintes atividades: **geração, codificação e transferência de conhecimento**.

O termo pode ser entendido ainda como um processo de criação, enriquecimento, capitalização e validação de conhecimento e *know-how* envolvendo todos os atores da organização, cujo objetivo é o **desempenho coletivo e a sobrevivência a longo prazo da empresa**. Trata-se, portanto, de um processo organizacional que facilita a aquisição, estruturação, integração e disseminação do conhecimento dos atores em toda a organização, com o objetivo de fornecer uma ajuda ao trabalho e melhorar a eficácia.

O ponto importante é **ampliar o uso e a criação de conhecimento** dentro de uma organização, com dois objetivos complementares e fortemente interligados: um relacionado ao patrimônio da organização e outro à inovação sustentável. Esses objetivos que são sustentados pelas suas dimensões econômicas, humanas, sociais e culturais.

Observa-se que a gestão do conhecimento **pressupõe a adoção de um conjunto de tecnologias e processos** cujo objetivo é apoiar a criação, transferência e aplicação do conhecimento nas organizações. A utilização de bancos de dados internos e externos, de informações dos parceiros e de métricas para quantificar a eficiência da gestão são partes relevantes neste contexto.

Em outras palavras, a gestão do conhecimento é um conjunto de atividades relacionadas a identificação e gestão do aprendizado organizacional. Ela permite as empresas ou órgãos públicos entenderem e organizarem suas estruturas de **informações** e **conhecimento**. Como todo processo de **gestão** devemos ter etapas de planejamento, organização, avaliação e controle (PDCA). A figura abaixo apresenta essa definição de forma organizada.



Vamos agora fazer duas rápidas questões rápidas para fixarmos um pouco mais deste conteúdo:

1. **Ano: 2017 Banca: FCC Órgão: TRT - 11ª Região (AM e RR) Prova: Analista Judiciário - Área Administrativa**

   Na Era da Informação, a administração de recursos humanos cedeu lugar a uma nova abordagem, a Gestão de Pessoas, onde as pessoas são vistas como

   a) recursos da organização.

   b) fornecedoras de conhecimento e competências.

   c) fornecedoras de mão de obra.

   d) previsíveis e flexíveis.

   e) dependentes da organização.

Comentário: Essa questão complementa o que acabamos de aprender sobre gestão do conhecimento. Esse modelo faz parte de uma evolução na gestão organizacional na qual as pessoas são vistas como parceiras, inteligentes e proativas, além de fornecedoras de conhecimento e competências. Assim, nossa resposta pode ser observada na alternativa B.

Gabarito: B

---

2. Ano: 2014 Banca: FCC Órgão: TCE-GO Prova: Analista de Controle Externo - Gestão de Conhecimento

Analise as afirmações a seguir:

I. O mapeamento de processos é um alicerce para todo e qualquer trabalho de Gestão do Conhecimento, pois uma empresa precisa conhecer como ela funciona, ou seja, como são realizadas suas operações, seus negócios e atividades.

II. O ser humano é o principal e imprescindível agente para uma eficiente implementação da Gestão do Conhecimento.

Em relação a elas, é correto afirmar que

a) I e II estão corretas.

b) I e II estão incorretas.

c) I está correta; II exagera a importância do ser humano na implementação da Gestão do Conhecimento.

d) I supervaloriza o papel do mapeamento de processos na Gestão do Conhecimento; II está incorreta, pois o principal agente para implementação da Gestão do Conhecimento é a motivação dos colaboradores.

e) I e II estão incompletas: I, por não mencionar o papel da cultura organizacional na Gestão do Conhecimento; II, por não considerar a importância da Tecnologia da Informação na Gestão do Conhecimento.

Comentário: Veja que ambas as alternativas estão corretas. O fator humano é importante, bem como o aspecto tecnológico associado a adoção da gestão do conhecimento. Logo, temos nosso gabarito na alternativa A.

Gabarito: A

Para entendermos o fluxo ou processo associado a gestão do conhecimento de forma mais consistente precisamos começar pela formalização de alguns conceitos básicos: **dados, informação, conhecimento e inteligência**. Esse será nosso objetivo a partir de agora. Faltaremos também de dados estruturados e não estruturados. Vem comigo!

DADO, INFORMAÇÃO, CONHECIMENTO E INTELIGÊNCIA.

Nosso conteúdo tem início em uma citação: “São múltiplos os desafios de um mundo globalizado, sendo difícil até mesmo obter consenso sobre quais seriam os prioritários. Todavia, um deles – **tornar o conhecimento acessível a todos** – destaca-se como um dos mais importantes devido aos seus efeitos estruturantes”.

É com esse texto que se inicia a apresentação do livro Inteligência, Informação e Conhecimento, uma coleção de artigos sobre gestão do conhecimento organizada em 2006 pela UNESCO. Para entender este assunto é preciso conhecer os conceitos básicos que vão subsidiar nosso entendimento, são eles: dados, informação, conhecimento e inteligência.

Dados podem ser definidos como sucessões de fatos brutos, que não foram organizados, processados, relacionados, avaliados ou interpretados, representando apenas partes isoladas de eventos, situações ou ocorrências. Constituem as unidades básicas a partir das quais informações poderão ser elaboradas ou obtidas. Os dados podem ser vistos como um conjunto de símbolos usados para armazenar e comunicar.

O dado é a forma primitiva que compõe os sistemas de informação. Podemos considerar os dados como a menor partícula estruturada que compõe uma informação. Quando os dados passam por algum tipo relacionamento, avaliação, interpretação ou organização tem-se a geração de **informação**. A partir do momento em que dados são transformados em informações, decisões podem ser tomadas.

A **informação** pode ser definida como um dado acrescido de **contexto, relevância e propósito**. Requer um mínimo de análise para sua obtenção, através de uma avaliação humana. Podemos entender melhor essa ideia usando uma concatenação de dados. “De”, “35”, “Thiago”, “=” e “idade” são dados quando observados individualmente. Porém, o agrupamento dos termos acima de forma lógica nos leva a uma informação: “Idade de Thiago = 35”.

A informação é gerada a partir de uma interpretação sobre os dados, estes podem ser contextualizados, categorizados, calculados ou condensados. São fatos sobre uma situação, pessoa ou evento. Podemos ainda transformar os dados em informação com significado filtrando, ordenando, estruturando. Percebam que não existe conclusão, mas a organização dos dados neste momento permite que algum tipo de análise possa ser executado.



RESUMINDO

Para transforma **dados** em **informações** precisamos que os mesmos sejam:

(1) precisos e oportunos

(2) específicos e organizados para um propósito

(3) apresentados dentro de um contexto que lhe dê significado e relevância, e

(4) que podem levar a um aumento na compreensão e diminuição da incerteza.

A informação é valiosa porque pode afetar o comportamento, uma decisão ou um resultado. Agora vamos passar as informações por um processo de validação, com foco naquilo que queremos obter. Ao receberem o tratamento adequado a partir de critérios inicialmente definidos, sob ponto de vista estratégico e com um caráter informativo obtemos o **conhecimento**.

Conhecimento é uma informação contextual, relevante e acionável. Segundo **Turban**, conhecimento é a informação em ação. Capital intelectual (ou recursos intelectuais) é outro termo frequentemente usado como sinônimo de conhecimento. Conhecimento é uma informação valiosa da mente, inclui **reflexão, síntese e contexto**. É difícil de estruturar, difícil de capturar em computadores, normalmente é tácito (não explícito) e sua transparência é complexa.

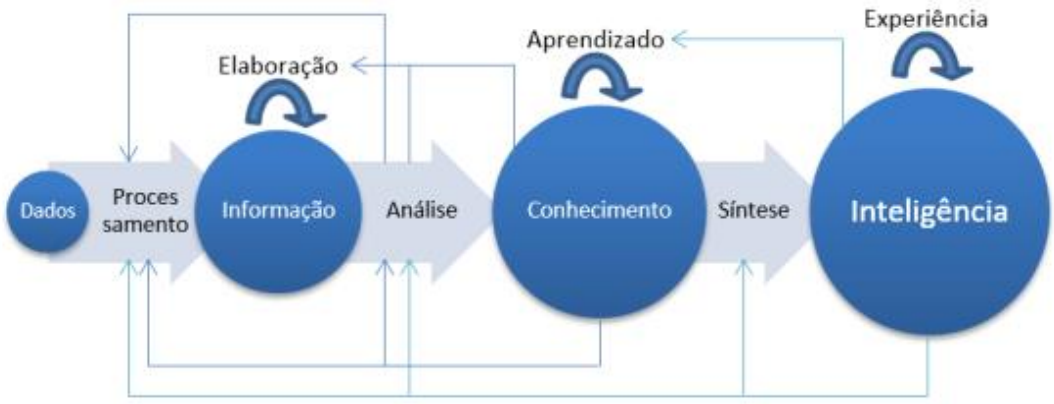
Veja abaixo um quadro comparativo entre conhecimento, informação e dados:



Dados	Informação	Conhecimento
Simples observações sobre o estado do mundo. • Facilmente estruturado • Facilmente obtido por máquinas • Frequentemente quantificado • Facilmente transferido	Dados dotados de relevância e propósito • Requer unidade de análise • Exige consenso em relação ao significado • Exige mediação humana	Informação valiosa da mente humana. Inclui reflexão, síntese e contexto • De difícil estruturação • De difícil captura em máquinas • Frequentemente tácito • De difícil transferência.

Faltou apresentar a inteligência! Para exibi-la de forma contextualizada e aproveitar para expor algumas características do processo quem tem início na aglomeração de dados brutos vamos apresentar uma figura abaixo e em seguida tecer alguns comentários sobre a mesma.





Dados compreendem a classe mais baixa de informação e incluem os itens que representam fatos, textos, gráficos, imagens estáticas, sons, segmentos de vídeo analógicos ou digitais etc. Os dados são coletados, por meio de processos organizacionais, nos ambientes interno e externo. Em suma, dados são sinais que não foram processados, correlacionados, integrados, avaliados ou interpretados de qualquer forma. Esta classe representa a matéria-prima a ser utilizada na produção de informações.

A próxima classe é a da informação propriamente dita. Nesta, os dados passam por algum tipo de processamento para serem exibidos em uma forma inteligível às pessoas que irão utilizá-los. Processar dados inclui a revelação de fotografias de um filme, as transmissões de rádio transformadas em um formato de relatório padronizado, a exibição de arquivos de computador como texto ou gráfico em uma tela, a grade de coordenadas em um mapa etc. O processo de transformação envolve a aplicação de procedimentos que incluem formatação, tradução, fusão, impressão e assim por diante. A maior parte deste processo **pode ser executada automaticamente**.

Uma vez que dados tenham sido transformados em informações, pelo menos em uma interpretação inicial, é possível refinar as informações mediante um processo de **elaboração**. As informações resultantes deste processo incluem características adicionais do problema, geram hipóteses, consequências das hipóteses, sugerem soluções para problemas, explanação e justificativas de sugestões, crítica de argumentos, entre outras saídas. Portanto, a transformação de dados em informações deve ser vista simplificadamente, como um tipo de pré-processamento de um processo de elaboração.

O próximo nível é o do **conhecimento**, que pode ser definido como sendo informações que foram analisadas e avaliadas sobre a sua **confiabilidade**, sua **relevância** e sua **importância**. Neste caso, o conhecimento é obtido pela interpretação e integração de vários dados e informações para iniciar a construção de um quadro de situação.

O processo de transformação é realizado por meio de avaliação de dados e informações. Os insumos provenientes das diversas fontes são analisados e combinados na síntese de um produto final, o conhecimento. É por meio do conhecimento que **aqueles que assessoram as decisões buscam uma compreensão mais efetiva da situação problema**.

O conhecimento **não é estático**, modificando-se mediante **a interação com o ambiente**, sendo este processo denominado **aprendizado**. Uma visão mais ampla é que o aprendizado é a integração de novas informações em estruturas de conhecimento, de modo a torná-las potencialmente utilizáveis em processos futuros de processamento e de elaboração. Além disto, conhecimentos novos podem resultar de um processo de inferência na própria estrutura do conhecimento.

O nível mais alto desta hierarquia é **a inteligência**, que pode ser entendida como sendo a informação como oportunidade, ou seja, **o conhecimento contextualmente relevante que permite atuar com vantagem no ambiente considerado**. Também pode ser vista como o conhecimento que foi sintetizado e aplicado a uma determinada situação, para ganhar maior profundidade de consciência da mesma.

Portanto, a inteligência resulta da síntese de corpos de conhecimentos, são usados julgamento e intuição daquele que toma decisões e uma visualização completa da situação é obtida. Nesta etapa os tomadores de decisão aplicam ao conhecimento gerado, suas habilidades, suas competências de

negócio e vivência na organização, para identificar direções estratégicas, tais como: novos projetos de pesquisa, acordos de cooperação, transferência de tecnologia e ações e reações da concorrência.

Idealmente, o entendimento da situação apoia a tomada de decisão a partir da visualização do cenário e cria as condições para que o planejamento possa ser realizado e as ações efetivadas. Além disso, podem ser revelados fatores críticos, possibilitando a antecipação a eventos, mediante o reconhecimento das consequências de novos ou iminentes efeitos de uma decisão. Por tudo isso, **a inteligência** deve ser a base do processo decisório, mesmo considerando que raramente é possível alcançar a compreensão total.

A transformação de conhecimento em inteligência é realizada por meio de síntese, sendo uma habilidade puramente humana baseada na experiência e intuição, que vai muito além da capacidade de qualquer sistema especialista ou de inteligência artificial. Síntese simplesmente não pode ser reduzida a procedimentos ou regras, por não considerarem o complexo.

Por fim, a **experiência** pode ser definida como **a efetividade da inteligência de uma organização**, que é aperfeiçoada pelas decisões tomadas e considerada geradora de algum tipo de vantagem. Sendo assim, a experiência agrega valor ao processo decisório de uma organização, por refletir toda a capacidade em atuar no ambiente competitivo.

Vejamos uma questão de uma prova anterior para fixarmos um pouco mais o conteúdo:

3. Ano: 2015 Banca: FCC Órgão: CNMP Prova: Analista do CNMP - Gestão Pública

Os Sistemas de Informação (SI) são construídos com Dados, Informação, Conhecimento e Inteligência. Sobre o tema, considere:

I. Informação é coletada nos ambientes interno e externo e representa, por exemplo: fatos, textos, gráficos.

II. A inteligência é realizada por meio de síntese, baseada em experiência e intuição, sendo uma habilidade humana.

III. Conhecimento demanda análise e avaliação sobre a confiabilidade, relevância e importância de dados e informações para a construção de um quadro de situação.

Está correto o que consta APENAS em:

a) I e III.

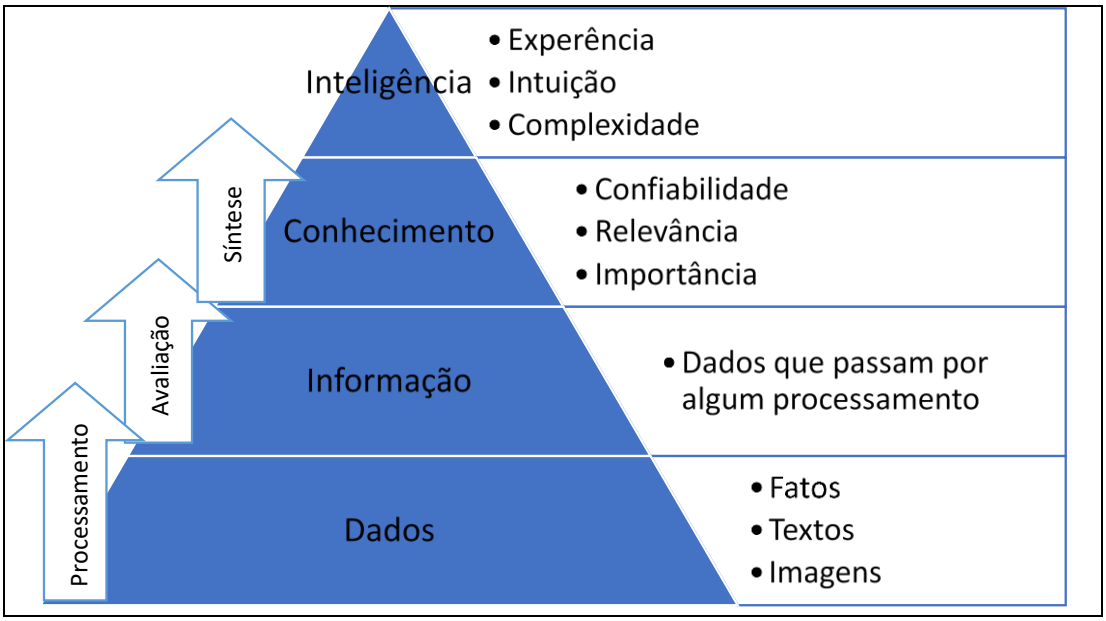
b) II.

c) II e III.

d) I.

e) III.

Comentário: Começamos o comentário desta questão solicitando que você observe com cuidado a figura abaixo. Em seguida, vamos avaliar cada uma das afirmações.



A afirmação I está incorreta, veja que os **dados** que são coletados em forma de fatos, textos e imagens. A segunda afirmação está correta e consistente com a figura. A inteligência pode ser vista como o conhecimento que foi **sintetizado e aplicado** a determinada situação para ganhar maior profundidade e consciência. Por fim, temos a última afirmação, que também está correta, pois o conhecimento, pode ser definido como informações cuja **relevância, confiabilidade e importância** foram avaliadas. Neste caso, o conhecimento é obtido pela interpretação e integração de vários dados e informações.

Desta forma, podemos marcar nossa resposta na alternativa C.

Gabarito: C.

CONHECIMENTO TÁCITO E EXPLÍCITO

Observando os conceitos acima apresentados é possível prever que uma empresa pode gerir de forma eficiente e informatizada o conhecimento. Devemos, portanto, viabilizar a conversão de conhecimento para em seguida geri-lo de maneira adequada.

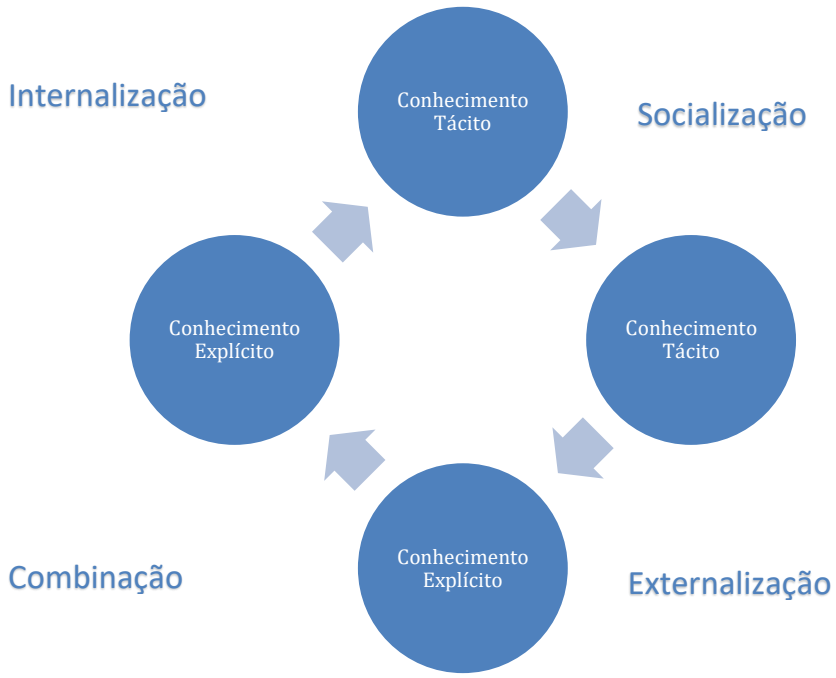
O conhecimento pode ser dividido em **tácito**, armazenado nas pessoas e adquirido através de experiências de vida, e **explícito** ou codificado, encontrado na forma de texto, formalizado e pode ser facilmente transmitido através do um meio físico. Desta forma, há quatro formas básicas pelas quais o conhecimento pode ser transmitido:

1º **Socialização** que é a interação do conhecimento através da troca de ideias e do compartilhamento de experiências, por meio de observação ou/e da prática.

2º **Externalização** neste caso o conhecimento pessoal passa para o domínio público por meio de uma documentação. Este pode ser reutilizado por outro, portanto deve ser classificado.

3º **Internalização** adquirido a partir de leitura. Vem de alguma publicação escrita que pode estar disponível em livro, revistas, na Internet ...

4º **Combinação** neste caso o conhecimento explícito de diferentes fontes é agrupado ou misturado visando agregar valor para gerar um novo conhecimento.



Vejam a figura acima as formas de conversão do conhecimento. Agora que já temos o conhecimento a respeito do assunto, vamos fazer uma rápida questão do **CESPE** para fixarmos o conteúdo.



4. Ano: 2016 Banca: CESPE Órgão: TCE-SC Cargo: Auditor de TI

Julgue os itens a seguir, acerca de dado, informação, conhecimento e inteligência.

90 Define-se informação como significado, ou seja, como registros icônicos e simbólicos — fonéticos ou numéricos — e signos — linguísticos, lógicos ou matemáticos —, por meio dos quais se representam atos, conceitos ou instruções.

91 O atributo de inteligência depende mais da qualidade da informação disponível do que da sua quantidade, tendo, portanto, natureza qualitativa.

Comentário: Vamos analisar cada uma das alternativas acima.

[90] Vejamos uma definição formal de informação: Informação é todo o conjunto de dados devidamente ordenados e organizados de forma a terem significado. Dados por sua vez pode ser definido como uma representação dos fatos, conceitos ou instruções de uma maneira normalizada que se adapte à comunicação, interpretação e processamento pelo ser humano ou através de máquinas. Os dados são representados por símbolos como por exemplo as letras do alfabeto: a, b, c, etc, mas não são em si a informação desejada. Vejam que a alternativa troca o significado de dados e informação, estando, portanto, incorreta.

[91] Existe uma estrutura baseada nas noções de qualidade versus quantidade, classificando a informação em quatro níveis, na medida em que é agregado valor a cada nível, e nos apresenta essa estrutura através da pirâmide informacional, dado, informação, conhecimento e inteligência. Assim, explica que o atributo da inteligência é mais de caráter qualitativo que quantitativo, uma vez que depende mais da qualidade da informação adquirida do que de sua quantidade, ao contrário dos dados, onde a quantidade de informação é mais importante que sua qualidade aparente. Desta forma a alternativa encontra-se correta.

Gabarito: E C.

DADOS ESTRUTURADOS E NÃO ESTRUTURADOS.

Quando resolvemos classificar os dados, uma das taxonomias possíveis é aquela que divide os dados de acordo com a sua estruturação. Neste contexto segregamos os dados em **estruturados**, **semiestruturados** e **não estruturados**. Vamos tentar definir cada um deles:

> **Dados estruturados** são aqueles quem possuem a mesma estrutura de representação rígida e previamente projetada, ou sejam, existe um esquema que estabelece algumas características dos dados que serão armazenados. Esses dados são organizados e gravado em um banco de dados.

Perceba que os dados estruturados são definidos por um esquema ou um formato padrão para cada um dos atributos. Podemos, portanto, observar que existe uma mesma estrutura de representação

para cada grupo de item de dados (relação) que é rígida e definida previamente por meio de um esquema.

Vamos tentar abstrair um pouco essa definição. Imagine uma tabela com apenas duas colunas nome e telefone. O nome deve possuir no máximo 20 caracteres e o telefone deve ser um valor numérico de 9 algarismos. Veja que para incluir um par, nome/telefone, nesta tabela você deve respeitar o tipo e tamanho dos campos.

Outro ponto interessante é que a tabela tem apenas dois atributos ou colunas, essa característica foi definida pelo esquema. O esquema é semelhante a uma forma de bolo. Não temos como inserir outra informação, como endereço se não reformularmos o esquema para aceitar esse novo atributo. Enfim, perceba que temos uma rigidez quando aos valores que podem ser manipulados quando trabalhamos com dados estruturados.

Para flexibilizar essa situação foram criados os dados semiestruturados.

> **Dados semiestruturados**. Essa segunda categoria de dados geralmente não é mantida em um banco de dados (embora a maioria dos bancos de dados relacionais admita o tipo XML, seguindo a recomendação do padrão SQL). Os dados semiestruturados apresentam uma organização bastante heterogênea, o que pode dificultar as consultas a esses dados.

Os dados não são estritamente tipados[1], mas também não são completamente desestruturados. O que então faz com que ele seja classificado desta forma?

Os dados semiestruturados são dados onde o esquema de representação está presente de forma explícita ou implícita, podendo ser, inclusive, auto descritivo. Uma análise do dado dever ser feita para que a sua estrutura seja identificada e extraída. Como exemplos de dados semiestruturados podemos citar: XML – *eXtensible Markup Language*, RDF – *Resource Description Framework*, OWL – *Web Ontology Language*. A figura abaixo apresenta a diferença entre um texto plano e um conjunto de dados semiestruturado.

[1] Um tipo de dados em uma linguagem de programação ou em um banco de dados são os domínios básicos que podem ser atribuídos a cada uma das colunas de uma tabela. Os tipos de dados mais utilizados são os numéricos e os de caracteres.

Exemplo de texto e seu correspondente XML

Catálogo de endereços João Silva Rua Carijós, 135 Belo Horizonte, MG 30.000 Brasil 31 3335-5556 (preferido) 31 3549-4446 joaosilva@net.com.br José Almeida jalmeida@net.com.br	<?xml version="1.0"?> <catálogo de endereços> <entrada> <nome> João Silva </nome> <endereço> <rua> Carijós, 135</rua> <estado> MG </estado> < cep> 30.000 </cep> <país> Brasil </pais> <endereço> <telefone preferido="true">31 3335-4456</telefone> <telefone> 31 3594-4446 </telefone> <email> joaosilva@net.com.br </email> </entrada> <entrada> <nome><prim>José</prim> <sobren>Almeida</sobren> <email> jalmeida@net.com.br </email> </entrada> </catalogo de endereço>

Como características destes dados podemos listar:

1. **Definição à posteriori** – os esquemas são definidos após a existência dos dados por meio da investigação de suas estruturas particulares.
2. **Estrutura irregular** – não existe um esquema padrão para os dados, as coleções de dados são definidas de maneiras diferentes, podendo conter informações incompletas.
3. **Estrutura implícita** – muitas vezes uma estrutura implícita.
4. **Estrutura parcial** – apenas parte dos dados disponíveis podem ter uma estrutura.

Agora podemos passar para uma análise dos dados não estruturados.

Dados não estruturados. De forma simples são inclusos aqui os dados que não possuem uma estrutura definida. Podem ser listados como exemplo documentos, textos, imagens e vídeos. Esses, geralmente, não possuem descrição para suas estruturas, nem implicitamente. A grande maioria dos dados disponíveis na Web são classificados dentro destes formatos.

Tendo em vista a grande quantidade de dados (**volume**), uma imensa **variedade** de dados e a necessidade de processamento cada vez mais rápida (**velocidade**) surge o conceito de Big Data, que será assunto de uma aula posterior (caso seja um assunto cobrando no seu concurso).

Esse assunto foi cobrado pelo CESPE no concurso do TCE-PB de 2018, vejamos:



5. Ano: 2018 Banca: CESPE Órgão: TCE-PB Cargo: Auditor de Constas Públicas Questão: 96

Com relação a dados estruturados e não estruturados, assinale a opção correta.

A Dados não estruturados podem ser caracterizados por arquivos de diversos tipos — textos, imagens, vídeos, entre outros —, cujas estruturas não são descritas implicitamente.

B Por padrão, documentos do tipo XML (eXtensible Markup Language) são estruturados.

C Dados não estruturados de um mesmo grupo possuem as mesmas descrições e, consequentemente, os mesmos atributos.

D Por padrão, dados não estruturados são organizados em blocos semânticos.

E A alta heterogeneidade facilita as consultas aos dados não estruturados, desde que estes estejam ligados por ponteiros.

Comentário: Vamos analisar cada uma das alternativas acima.

A primeira alternativa trata de dados não estruturados, cujas estruturas não são descritas nem **explicitamente, nem implicitamente.** Veja que o examinador tentou confundir o candidato usando apenas o termo implícito. Isso é o que acontece com os dados **semiestruturados**, onde não existe uma definição a priori dos arquivos, contudo você consegue deduzir características do arquivo. Agora vamos analisar o que está errado nas demais alternativa.

Os documentos XML são considerados semiestruturados. Eles não têm um tamanho definido para seus elementos, mas os termos possuem "tags" que descrevem os atributos presentes nos arquivos. Sendo assim, a alternativa b está incorreta. Já a letra c, inclui atributos nos dados não estruturados. Os dados não estruturados tendem a possuir uma forma mais livre, não tabular, são dispersos e não facilmente recuperáveis. Esses dados requerem uma intervenção deliberada para dar sentido aos mesmos. E-mails, documentos, páginas da web e arquivos (sejam eles texto, áudio e / ou vídeo) em locais dispersos são exemplos de dados não estruturados.

Tente abrir um documento de música ou vídeo em um notepad. Perceba que não existe nenhum bloco semântico facilmente identificado dentro do arquivo. Existe um termo denominado Web Semântica que é uma tentativa de prover soluções e tecnologias adequadas para resolver o problema da integração de dados, em especial quando usamos dados estruturados e não estruturados.

Por fim, a última afirmação é totalmente equivocada. A heterogeneidade é um dos fatores que dificulta as consultas em dados não estruturados. Podemos resumir alguns pontos de dados estruturados e não estruturados na tabela abaixo:

Característica	Dados estruturados	Dados não estruturados
Associação	Organizado	Difuso e disperso
Aparência	Formalmente definido	Definição livre
Acessibilidade	De fácil acesso e consulta	Dificil de cosultar e acessar

Gabarito: A

Antes de continuarmos vamos fazer mais algumas questões do CESPE.



6. Ano: 2016 Banca: CESPE Órgão: TCE-SC Cargo: Auditor de TI

A respeito de dados estruturados, não estruturados e abertos, julgue os itens subsequentes.

93 Em se tratando de dados estruturados, a informação de esquema está mesclada aos valores dos dados, e cada objeto de dados pode ter atributos diferentes, que não são conhecidos com antecedência. Essa característica os diferencia de dados não estruturados.

Comentário: Vamos, mais uma vez, comentar as assertivas acima.

[93] Dados estruturados são organizados em linhas e colunas, geralmente são encontrados em banco de dados relacionais (falaremos sobre eles na aula 02 do nosso curso), são eficientes quanto à recuperação e processamento. Cada coluna deve ter o mesmo tipo de dados para todos os valores diferentes de nulo[2], assim como as linhas devem possuir registros que satisfação as restrições de integridade descritas na definição da tabela. A definição da questão é confusa e se assemelha com o contexto de dados semiestruturados. Assim, podemos concluir que a afirmação está incorreta.

Gabarito: C E.



7. Ano: 2018 Banca: CESPE Assunto: Informática para Polícia Federal Cargo: Agente Conteúdo Banco de dados

Julgue os próximos itens, a respeito da teoria da informação e de sistemas de informação.

79 O conceito de conhecimento é mais complexo que o de informação, pois conhecimento pressupõe um processo de compreensão e internalização das informações recebidas, possivelmente combinando-as.

Comentário: 79. Vejam que essa alternativa coloca o conhecimento como uma etapa posterior a informação dentro de um fluxo de encadeamento e complexidade. O conhecimento, conforme falado na nossa revisão, trata de informações úteis a um propósito, envolve entendimento e análise de um conjunto de informações. Sendo assim, podemos marcar a assertiva como **correta**.

Gabarito: 79. C.

---

8. Ano: 2018 Prova: Perito – Polícia Federal Banca: CESPE Assunto: Conhecimentos básicos – banco de dados e teoria da informação.

Julgue os próximos itens, a respeito de computação na nuvem, sistemas de informações e teoria da informação.

39 A informação se caracteriza pela compreensão e internalização do conteúdo recebido, por meio do seu uso em nossas ações; o dado, por sua vez, é um elemento bruto dotado apenas

[2] Valores nulos correspondem ao conjunto vazio. Suponha que um campo em uma tabela Excel não esteja preenchido, se formos traduzir esse fato para uma tabela em um banco de dados relacional em deve ser preenchido com NULL ou nulo.

de significado e relevância que visem fornecer uma solução para determinada situação de decisão.

Comentário: Essa questão tem um erro na definição de dado. O dado não possui significado e relevância. Logo, temos uma alternativa **incorreta**.

Gabarito: 39. E.

---

9. Ano: 2018 Banca: CESPE Assunto: Informática para Polícia Federal Cargo: Escrivão Conteúdo Teoria da informação

Julgue os itens a seguir, a respeito da teoria da informação e de metadados de arquivos.

79 Em arquivos no formato XML, as tags não são consideradas metadados.

81 Todas as informações geradas a partir do sistema de informação de uma organização são consideradas como dados da organização.

82 O conhecimento é embasado na inteligência das informações que são coletadas e analisadas para uma organização.

Comentários: Vamos falar um pouco sobre cada uma das alternativas.

79. As tags em arquivos XML contêm os descritores dos dados, logo são considerados metadados. Alternativa **incorreta**.

81. As informações geradas pelos sistemas são de fato consideradas um ativo organizacional e, em última instância, são considerados dados da organização armazenados nos bancos de dados e arquivos. Logo, temos uma alternativa **correta**.

82. A expressão "inteligência das informações" mistura conceitos e não faz sentido. Logo, temos uma alternativa errada.

Gabarito: 79. E 81. C 82. E

---

10. Ano: 2018 Banca: CESPE Assunto: Informática para Polícia Federal Cargo: Papiloscopista Conteúdo: Dados e Informação

Acerca da definição de dados e informação e sua representação em sistemas de informação, julgue os itens que se seguem.

82 Na representação da informação, os atributos permitem que entidades e eventos possam ser reconhecidos, referidos e descritos. Um atributo relacional permite relacionar eventos e entidades.

83 Dados são fatos que descrevem os objetos de informação, por exemplo, eventos e entidades.

84 Informação é constituída por um conjunto de dados com características específicas. O ponto de análise é que os dados devem ser irrelevantes para o sistema a que se destinam.

Comentário: Vamos comentar cada um dos itens acima.

82. Essa questão usou uma classificação de atributos de sistemas de informação que é um pouco mais abrangente do que a definição vista nos modelos de dados entidade relacionamento e relacional. Atributos estão associados ao fato das entidades e eventos podem ser reconhecidos, referidos e descritos em termos dos seus atributos. São basicamente fatos caracterizadores dos objetos de informação. Neste contexto temos os seguintes tipos de atributos:

Identificadores: úteis na distinção de objetos.

Descritores: descrição física dos objetos.

Localizadores: espaço físico dos objetos.

Temporais: quando ocorre um evento (tempo).

Relacionais: permitem relacionar eventos e entidades.

Classificadores: fatos que determinam o modo de relacionamento entre os objetos de informação e a organização, por categoria ou tipo.

Condicionais: factos que identificam o objeto de informação, escolhendo uma de várias alternativas. **Logo, a alternativa está correta**.

83. Dados são fatos que descrevem os objetos de informação (eventos e entidades). Os dados referem-se a mais de um fato. Um determinado fato é referido como item. Logo, temos mais uma alternativa correta.

84. Informação é constituída por um conjunto de dados com características específicas, isto é, trata-se de um conjunto de dados significativos e **relevantes** para a componente ou sistema a quem se destinam. Logo, alternativa incorreta.

Gabarito: 82. C 83. C 84. E

CONCEITOS DE BANCO DE DADOS

CONCEITOS BÁSICOS

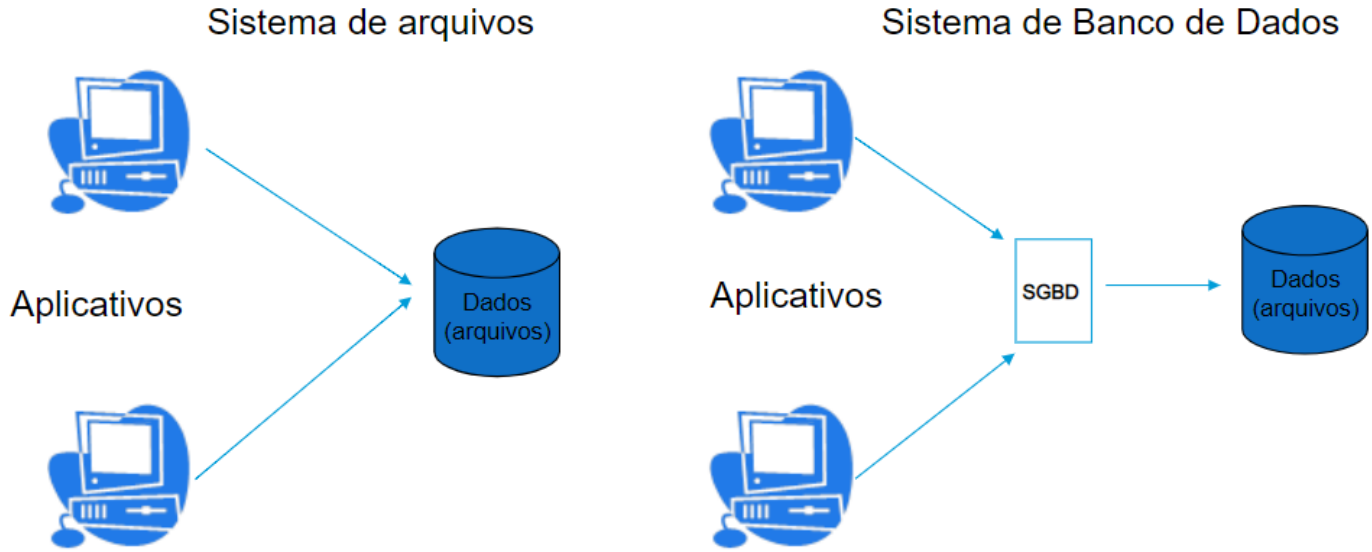


Antes de adentrar nos conceitos de banco de dados gostaria de voltar um pouco no tempo. Preciso que você conheça os sistemas de arquivo. Você deve ter acesso a um sistema de arquivo do computador ou dispositivo que você está usando para ter acesso a esse conteúdo. Agora imagine a seguinte situação: você trabalha em uma empresa e a lista de clientes está armazenada em um arquivo do Excel.

Você e todos os demais funcionários do setor de vendas (isso mesmo, na nossa história você trabalha no setor de vendas) fazem acesso ao mesmo arquivo para incluir e atualizar dados dos clientes. Quando o arquivo começa a ficar sem controle vocês resolvem criar cópias do mesmo para que cada vendedor possa ter seu próprio cadastro. A solução do problema de acesso simultâneo nos leva a outro problema: **redundância**. Várias cópias sem controle podem ainda ter inconsistência entre os dados armazenados.

Para resolver essa situação você e seus colegas de trabalho ouvem falar de **uma solução robusta** para o problema de armazenamento de dados. Essa solução passa por manter apenas uma cópia dos dados em um banco de dados e usar um sistema gerenciador para **controlar o acesso concorrente**.

No universo tecnológico o que aconteceu em um passado recente foi mais ou menos a mesma coisa. Aplicações empresariais faziam acesso a arquivos sem um elemento central para controlar o processo. Eis que surge o Sistema de Gerenciamento de Banco de Dados (SGBD) para atuar como um guardião do banco de dados, que substituiu a abordagem de arquivos. Vejamos uma figura para esclarecer esses contextos:



Perceba que do lado esquerdo não temos o SGBD, já do lado direito existe o elemento. Para entender como e por que esse camarada apareceu aí vamos compreender os conceitos primordiais

para o entendimento do assunto. Em qualquer ciência, o entendimento completo do seu conteúdo deve se basear nos conceitos fundamentais. Nossa primeira aula começa focada nesses conceitos.

Quando iniciamos nosso estudo, precisamos ter a capacidade de responder a seguinte pergunta:

O que é **banco de dados**?

Você já consegue definir esse termo e suas principais características? Se você ainda não tem o entendimento perfeito do que seria um banco de dados, nosso objetivo agora é construir o conceito. Uma forma tradicional de definir banco de dados é começar entendendo o significado das palavras que compõe o termo: **Banco** e **dados**.

Banco tem diversas definições possíveis no dicionário da língua portuguesa. Dentre elas a que melhor se encaixa no nosso contexto é **um conjunto organizado e categorizado de objetos**, por exemplo, podemos ter um banco de fotografias ou um banco de leite.

Dados são **fatos conhecidos** que podem ser registrados e possuem **um significado implícito**. Esse conceito, porém, é um pouco amplo e abstrato para nosso intuito. Quando reduzimos o escopo à tecnologia da informação, temos um conceito mais adequado para dado. Ele é a representação física de um evento no tempo e espaço que não agrega fundamento ou significado para quem o sente ou recebe. É, basicamente, **um registro**!

Imagine que eu fale para você por "32260436, cinco, aprovado, Thiago". Você vai pensar, o professor está ficando louco! Mas em um banco de dados, esses registros são armazenados e chamados de dados. Para representarem uma informação eles precisam de um contexto associado. Se pensarmos em uma agenda telefônica o número "32260436" pode ser o telefone de Thiago. Aprovado pode ser a sua situação no seu próximo concurso.

Agora que temos o entendimento dos termos vamos partir para a definição do banco de dados.



De forma simples e direta: um **Banco de dados** é uma coleção de **dados** relacionados. Vejam que essa definição não estabelece a necessidade de os dados serem armazenados em formato digital. Alguns livros trazem o exemplo de uma agenda telefônica de papel como um exemplo bastante didático do conceito de banco de dados.

*O banco de dados, por si só, pode ser considerado como o equivalente eletrônico de um armário de arquivamento; ou seja, ele é um repositório ou recipiente para uma **coleção de arquivos de dados computadorizados**. – C J Date*

Esta definição, porém, é considerada muito simplista para alguns autores por não contextualizar o termo. O *Navathe*, por exemplo, cita três **propriedades implícitas** que contribuem para o entendimento do termo banco de dados. Primeiramente, ele representa **algum aspecto do mundo real**, às vezes chamado de **minimundo** ou de **universo de discurso** (*UoD – Universe of Discourse*). As mudanças no minimundo devem ser refletidas no banco de dados.

A **segunda** característica implícita diz que a **coleção de dados é logicamente coerente** com algum significado inerente. Uma variedade aleatória de dados **não** pode ser chamada de banco de dados. Um banco de dados pode armazenar as informações de uma empresa, uma faculdade ou um órgão

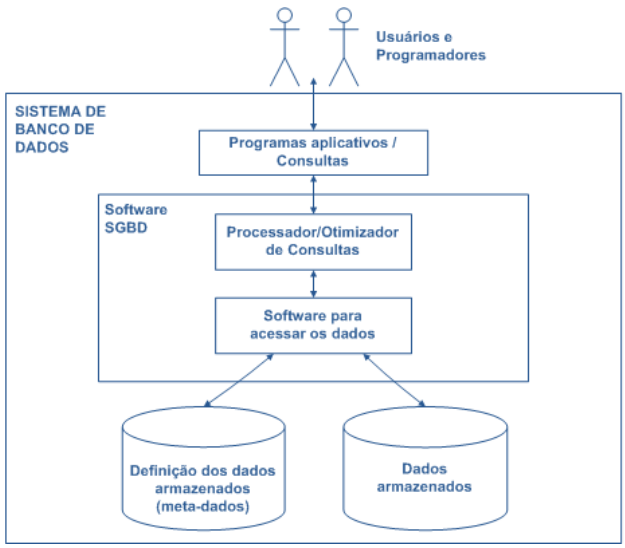
do setor público. Veja que essas informações estão dentro de um contexto, sendo, logicamente coerente.

A **terceira** propriedade afirma que um banco de dados é construído e populado com dados para uma **finalidade específica**. Ele possui um grupo de **usuários** bem definido e algumas aplicações, previamente concebidas, sobre as quais esses usuários interessados fazem acesso aos dados. Não adianta você criar um repositório sem propósito! Ele tem que atender alguma necessidade de acesso ao conjunto dos dados.

Vamos voltar ao nosso exemplo físico da agenda telefônica, você consegue visualizar as propriedades definidas pelo Navathe neste banco de dados. Qual o aspecto do mundo real que ela representa? Os dados estão logicamente relacionados? Possuem uma finalidade específica? Acredito que sim! Tente responder a essas perguntas mentalmente antes de seguir em frente.

As três propriedades definidas pelo Navathe começam a estruturar os conceitos na sua cabeça. Outra definição de **banco de dados** que resume o que apresentamos até agora é: "um conjunto de dados **estruturados** que são confiáveis, coerentes e compartilhados por usuários que têm necessidades de informações diferentes".

O primeiro conceito básico você já dever ter entendido. Você não deve ter dúvidas sobre o que seria um banco de dados. Vamos agora entender a diferença entre **banco de dados**, **sistemas de gerenciamento de banco de dados (SGBD)** e **sistemas de banco de dados (SBD)**. São três conceitos diferentes para os autores dos livros teóricos sobre o assunto. Para entender essas diferenças peço que você observe a figura a seguir:



Seguindo o fluxo de acesso aos dados, podemos observar que os usuários e programadores se comunicam com o sistema de banco de dados. Este, por sua vez, faz acesso ao sistema de gerenciamento do banco de dados. O SGDB usa as informações presentes nos bancos de dados, representados pelos cilindros da figura acima, para ter acesso aos dados armazenados.

Um **Sistema de Gerenciamento de Banco de Dados (SGBD)** é um **conjunto de programas** que permitem armazenar, modificar e extrair informações de um banco de dados. Seu **principal objetivo** é proporcionar um ambiente tanto conveniente quanto eficiente para a **recuperação e armazenamento** das informações do banco de dados.

Contudo, os SGBDs não se restringem apenas a manipulação dos dados. Eles fornecem uma variedade de programas com diferentes funcionalidades.



A figura acima apresenta algumas funcionalidades dos SGBDs. Cada espaço está preenchido com alguma funcionalidade que faz parte do escopo de um software de SGBD. Não se preocupe se você não tiver ideia do que está presente em cada espaço, ao longo desta aula ou do nosso curso, quando você começar a entender melhor o assunto, você perceberá a presença deles dentro do contexto de um SGBD. Vamos em frente!

Outro aspecto interessante sobre banco de dados são as descrições ou definições dos objetos, pense em uma tabela do modelo relacional. Para termos acesso a uma tabela precisamos conhecer sua estrutura, começando pelo seu nome e das suas colunas. Essa tarefa envolve especificar os **tipos, estruturas** e **restrições** dos dados a serem armazenados.

A definição ou informação descritiva do banco de dados também é armazenada pelo SGBD numa estrutura conhecida como **catálogo** ou **dicionário de dados**, que armazena os chamados de **metadados** dos objetos**.** Os metadados carregam consigo um significado. Uma coluna de uma tabela pode ser definida por um tipo de dados, inteiro, uma restrição, *not null*.

É possível ainda fazer o **compartilhamento dos dados** entre diversos usuários e programas, possibilitando o acesso ao banco de dados de forma simultânea. Outras funções importantes também são providas como **proteção** do sistema **contra defeitos** de hardware e software, feitos por meio de redundância ou replicação, e **proteção** de segurança **contra acesso** não autorizados ou maliciosos.

Outros aspectos interessantes estão relacionados com o controle de transações, recuperação após falha, otimização de consultas ou do próprio SGBD, auditoria por meio de logs de sistema, enfim, são várias as **funcionalidades providas** pelos softwares presentes em um **SGBD**. Vamos agora definir o próximo conceito: sistema de banco de dados!

O **sistema de banco de dados (SBD)** é considerado a união entre o banco de dados e o sistema de gerenciamento de banco de dados. Em outras palavras, consiste em uma coleção de dados inter-relacionados e de um conjunto de programas para acessá-los. Partindo da figura que apresentamos anteriormente conseguimos construir a seguinte fórmula:

SBD = BD + SGBD + (Programa de aplicação/consulta)

Antes de continuarmos construindo nosso arcabouço teórico sobre o assunto vamos resolver duas questões de provas passadas, mais uma vez nossa banca de exemplo será o CESPE.



11. Ano: 2016 Banca: CESPE Órgão: TCE-SC Prova: Auditor Fiscal de Controle Externo - Informática

Com relação aos bancos de dados relacionais, julgue o próximo item.

O catálogo de um sistema de gerenciamento de banco de dados relacional armazena a descrição da estrutura do banco de dados e contém informações a respeito de cada arquivo, do tipo e formato de armazenamento de cada item de dado e das restrições relativas aos dados.

Comentário: Perceba que a definição acima está de acordo com o termo dicionários de dados, catálogo de dados ou metadados presentes em um sistema de banco de dados. Lembre-se que essa separação entre a descrição dos dados e os dados propriamente dito é uma das características relevantes que foram apresentadas na evolução de sistemas de arquivos para a **abordagem de banco de dados**. Sendo assim, podemos afirmar que a questão está correta!

Gabarito: C.

---

12. Ano: 2010 Banca: CESPE Órgão: Banco da Amazônia Prova: Técnico Científico - Tecnologia da Informação

O dicionário de dados é uma das principais ferramentas para a administração dos dados corporativos. Por meio da engenharia reversa, pode-se armazenar os modelos de dados, as estruturas de dados, seus relacionamentos e toda a documentação necessária para garantir facilidade na localização e manipulação dos dados. Acerca dos papéis do administrador de dados (AD) e dos dicionários de dados, julgue os itens a seguir.

[1] O dicionário de dados é considerado um subconjunto das funções de um catálogo de sistema.

[2] O catálogo do sistema é um repositório com função de armazenar as definições dos esquemas dos bancos de dados.

Comentário: É importante lembrar que existe uma hierarquia entre os objetos ou elementos em um dicionário de dados. Um dicionário de dados possui a descrição dos esquemas ou catálogo de sistemas. Cada catálogo deve conter a descrição dos objetos que fazem parte do contexto de um sistema, como tabelas, visões e domínios. Dentro das definições das tabelas temos as descrições dos atributos e restrições de integridades dos dados.

Assim, ao analisar as alternativas acima, podemos inferir que a alternativa [1] encontra-se incorreta, já a afirmação [2] está certa!

Gabarito: E C.

CARACTERÍSTICAS DA ABORDAGEM DE BD

Segundo Navathe, são quatro, as principais características da abordagem de banco de dados que a fazem sobressair em relação às abordagens de processamento de arquivo.

Natureza de autodescrição de um sistema de banco de dados
Isolamento entre programas e dados, **abstração de dados**
Suporte a **múltiplas visões** de dados
Compartilhamento de dados e processamento de transação multiusuário.

Esses esforços visam **reduzir a redundância** o que implica em reduzir o desperdício no espaço de armazenamento e os esforços para manter os dados comuns ou duplicados atualizados. Tudo realizado por meio de um único repositório!

Vejam que a lista acima pode ser caracterizada como uma enumeração e, como eu sempre digo, listas fazem parte do rol de questões de prova de concurso. Seja qual for a matéria, sempre gaste um pouco do seu tempo lendo, mais de uma vez, cada uma das listas pertencentes aos assuntos. A verdade é: não importa o grau de relevância dentro do assunto, um examinador preguiçoso sempre está propício a utilizar deste artifício ao elaborar uma questão.

A primeira característica listada pelo Navathe é conhecida por nós como catálogo do SGBD, dicionário de dados ou metadados. Esta propriedade permite ao SGBD gravar as definições das suas estruturas e restrições. E, quais são as descrições que podem ser gravadas? Descrição de tabelas, tamanho do campo, tipo dos dados, propriedade de ser nulo ou não, valores default, restrições de integridade, entre outros. Para facilitar sua visualização pense numa definição de uma tabela em SQL[3]. Veja o exemplo a seguir e observe algumas dessas descrições.

[3] SQL é uma linguagem declarativa usada para enviar comandos aos sistemas de gerenciamento de banco de dados. Esses comandos permitem a criação de tabelas e a manipulação dos dados armazenados. Na criação da tabela podemos observar a definição dos tipos de dados (VARCHAR2, DATE, NUMBER), bem como das restrições de integridade (NOT NULL, DEFAULT).

```
CREATE TABLE PRL_EMPLOYEE(
  ID_EMPLOYEE NUMBER,
  FK_ID_MANAGER NUMBER,
  EMPLOYEE_FIRST_NAME VARCHAR2(100) NOT NULL,
  EMPLOYEE_LAST_NAME VARCHAR2(100) NOT NULL,
  EMPLOYEE_EMAIL VARCHAR2(100) NOT NULL,
  EMPLOYEE_B_DATE DATE NOT NULL,
  EMPLOYEE_START_DATE DATE DEFAULT SYSDATE,
  EMPLOYEE_END_DATE DATE DEFAULT NULL,
  EMPLOYEE_LOCKED NUMBER(1) DEFAULT 1
)
```

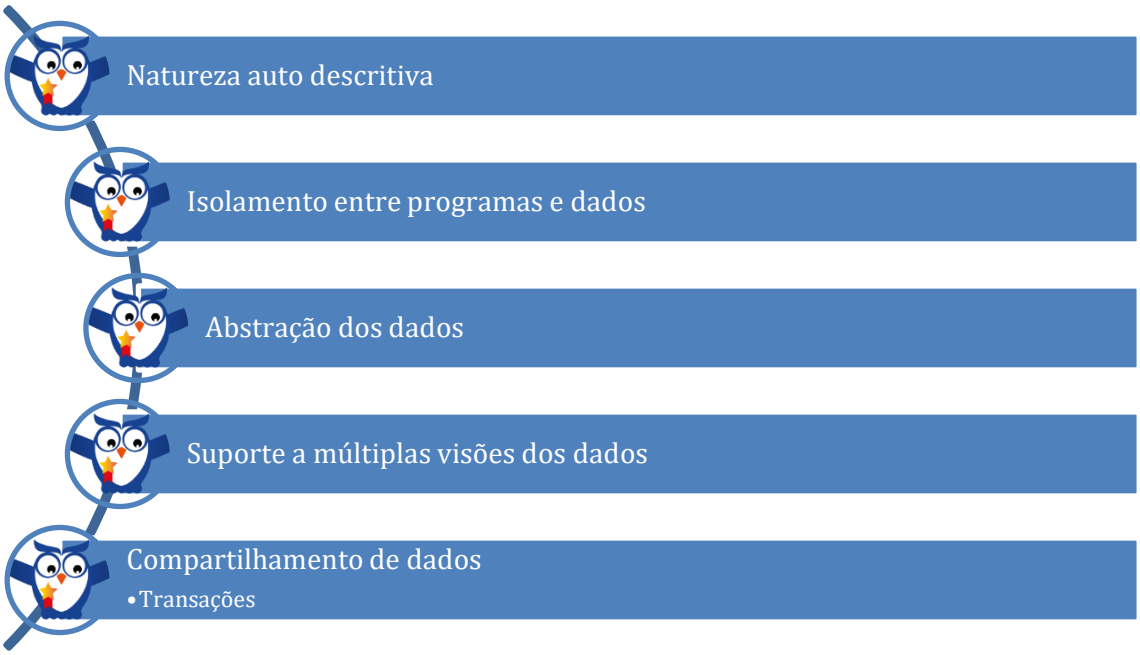
A próxima característica é uma decorrência da anterior. A partir do momento em que temos um dicionário de dados, é possível excluir da estrutura dos programas a definição dos dados presentes nos mesmos. Agora isolados, dados e aplicações, criam um conceito chamado **independência de dados do programa**. Este só é possível por conta da **abstração de dados**. A abstração de dados permite a criação de diferentes níveis de modelos. Cada modelo revela a estrutura dos dados de uma forma específica.

O suporte a **múltiplas visões** parte do princípio que diferentes usuários têm diferentes necessidades sobre os dados. Se pensarmos em SQL, uma VIEW representa um subconjunto de informações referentes a uma ou mais tabelas (ou até a nenhuma tabela). Do ponto de vista mais abstrato, uma visão é a parte do banco de dados ao qual um usuário ou grupo de usuários tem acesso. Porém, existe a possibilidade dessa visão conter um **dado virtual** que é derivado das informações armazenadas. Imagine, por exemplo, a idade calculada a partir da data de nascimento.

Quando falamos de suporte a **múltiplos usuários** queremos, basicamente, permitir que diferentes usuários acessem o banco de dados ao mesmo tempo. Para garantir que isso ocorra é preciso que o SGBD forneça um mecanismo de **controle de concorrência**. As transações efetuadas devem levar o sistema a um estado válido (C), não ter conhecimento umas das outras (I), serem executadas sempre por completo (ou não serem executadas) (A) e, uma vez gravadas na base, devem persistir ao longo do tempo (D).

Você dever estar se perguntando, o que são essas letras ao lado das características das transações citadas anteriormente? Elas se referem às características de uma transação: **atomicidade**, **consistência**, **isolamento** e **durabilidade**. Elas formam uma sigla conhecida como ACID e são bastante conhecidas dentro do modelo de dados relacional/transacional.





Acabamos de tratar das características que o Navathe utiliza para diferenciar sistemas de arquivo dos sistemas de banco de dados. Vamos agora listar as características descritas pelo Date e pelo Silberschatz. Date chama de benefícios da abordagem de banco de dados. Quais sejam:

1. O dado pode ser compartilhado
2. A redundância pode ser reduzida
3. Inconsistências podem ser evitadas
4. Pode-se utilizar o suporte a transações
5. A integridade pode ser mantida
6. A segurança pode ser aperfeiçoada
7. Requisitos conflitantes podem ser balanceados
8. Padrões podem ser utilizados

Já Abraham Silberschatz trata das desvantagens de se utilizar um **sistema de arquivo**:

1. Redundância e inconsistência dos dados
2. Dificuldade de acesso a dados
3. Isolamento dos dados
4. Problemas de integridade
5. Problemas de atomicidade
6. Anomalias de acesso concorrente
7. Problemas de segurança

Lembrem-se, não precisamos decorar todas essas listas, apenas tomar conhecimento da sua existência, pois fazem parte do contexto. Elas procuram sempre expor as características que diferenciam os sistemas de arquivos dos sistemas de banco de dados.

Vejamos mais algumas questões que tratam do assunto. Começando por uma questão do TCE-PB.



13. Ano: 2018 Banca: CESPE Órgão: TCE-PB Cargo: Auditor de Constas Públicas Questão: 97

A respeito de SGBDs, assinale a opção correta.

A Um SGBD, por definição, não é flexível, dada a dificuldade de mudar a estrutura dos dados quando os requisitos mudam.

B Um SGBD é um software que não prevê as funções de definição, recuperação e alteração de dados, sendo essa tarefa a função básica de um sistema de banco de dados.

C A consistência de dados é o princípio que determina a manutenção de determinado dado em vários arquivos diferentes.

D Conforme o princípio da atomicidade, caso ocorra erro em determinada transação, todo o conjunto a ela relacionado será desfeito até o retorno ao estado inicial, como se a transação nunca tivesse sido executada.

E O controle de concorrência é o princípio que garante e permite a manipulação, no mesmo momento, de um mesmo dado por mais de uma pessoa ou um sistema.

Comentário: Vamos analisar as alternativas acima. Elas são relevantes para consolidar nosso conhecimento sobre o assunto. Começando pela alternativa A, o erro aparece quando o examinador afirmar que os SGBDs não são flexíveis. Lembre-se que um SGBD veio solucionar um problema de replicação dos dados em diversos arquivos distintos. Agora, com uma fonte única dos dados, as alterações dos mesmos são centralizadas.

Já a alternativa B nos remonta as funcionalidades do SGBD. O que podemos fazer com tal sistema? É possível, primariamente, armazenar e manipular dados, para tal, é necessário descrever as estruturas das tabelas. Todos os SGBDs possuem estrutura para construção do banco de dados, usando uma linguagem como SQL. Sendo assim, a alternativa B também está incorreta.

As próximas duas alternativas tratam das propriedades das transações, mais especificamente da consistência, que tem por objetivo levar o banco de dados de um estado válido para outro estado consistente, e da atomicidade, que parte do princípio que uma transação é composta por vários comandos de modificação da base de dados e que estes comandos devem ser executados em conjunto completamente ou não serem executados. Desta forma, para garantir a atomicidade, caso uma transação falhe todas as operações já efetivadas precisam ser desfeitas. Desta forma, a alternativa C está incorreta e a alternativa D é a nossa resposta.

Por fim, o controle de concorrência é um mecanismo que permite que apenas um usuário consiga modificar um dado do sistema em um determinado momento. É importante entender que é possível várias pessoas terem acesso aos dados para leitura. Como se todos estivessem lendo um livro. Contudo, na hora de escrever, só existe uma caneta e, se alguma pessoa estiver de posse da caneta, apenas ela poderá fazer alterações na base.

Gabarito: D.

14. BANCA: FCC ANO: 2014 ÓRGÃO: TJ-AP PROVA: ANALISTA JUDICIÁRIO - BANCO DE DADOS - DBA

A redundância controlada de dados em um sistema de banco de dados ocorre quando

A um programa está ciente da múltipla representação de uma dada informação e garante o sincronismo entre as diversas representações.

B a responsabilidade pela manutenção do sincronismo entre as múltiplas representações de uma dada informação é compartilhada entre o programa e o usuário.

C os dados mais importantes são duplicados a cada backup do sistema, visando aumentar a garantia da recuperação da informação em caso de problemas.

D a responsabilidade pela manutenção do sincronismo entre as múltiplas representações de uma dada informação é do usuário.

E um programa está ciente da múltipla representação de uma dada informação, mas não garante o sincronismo entre as diversas representações.

Comentário: Nesta questão vamos entender o que está sendo cobrado antes de avaliarmos as alternativas. Em primeiro lugar, precisamos lembrar que o SGBD vai controlar a redundância dentro do banco de dados. Esse controle é feito quando diferentes usuários compartilham a mesma informação.

Suponha uma tabela de endereços de clientes em um banco comercial. Todos os setores do banco podem fazer uso desta informação. A área de investimento pode mandar o extrato das suas aplicações, o setor de cartão de crédito pode enviar sua fatura e o relacionamento com o cliente pode te enviar um cartão de feliz aniversário. A importância deste ponto único de contato é verificada quando a cliente muda de endereço, a atualização da tabela vai ser efetiva para os diferentes usuários da informação.

Agora vamos analisar as alternativas, ao ler cada uma delas, podemos observar que no SGBD um programa deve estar ciente da múltipla representação de uma dada informação e garantir o sincronismo entre as diversas representações. Essa descrição é a presentes na alternativa A que é a nossa resposta.

Gabarito: A.

PERSONAGEM DO ECOSSISTEMA DE BD

Quando tratamos de grandes organizações, as atividades relacionadas a banco de dados devem ser compartilhadas entre diferentes pessoas. Trataremos agora dos dois principais papéis dentro desse processo: o administrador de banco de dados (DBA) e o administrador de dados (AD).

Só para termos uma ideia, algumas empresas do setor bancário chegam a ter algumas dezenas de ADs dentro da organização. O Bradesco tem por volta de 40 ADs. Vamos então começar falando um pouco sobre esse perfil de trabalhador especializado em banco de dados.

O AD é a pessoa que toma as decisões estratégicas e de normas com relação aos dados da empresa. Os **administradores de dados** também podem ser conhecidos por projetista de dados. Suas tarefas

são realizadas principalmente antes do banco de dados ser realmente implementado e/ou populado.

Eles são responsáveis por identificar os dados a serem armazenados e escolher estruturas apropriadas para representar esses dados. Para isso precisam se comunicar com todos os potenciais usuários a fim de entender suas necessidades e criar um projeto que as atenda suas necessidades. Eles definem então visões para cada grupo de usuários. Podemos listar ainda como atribuições do AD:

- Padronizar os nomes dos objetos criados no BD
- Gerenciar e auxiliar na definição das regras de integridade
- Controlar a existência de informações redundantes
- Trabalhar de forma corporativa nos modelos de dados da organização

Falaremos agora do **DBA – *Database Administrator***, ou, da pessoa que fornece o **suporte técnico** necessário para implementar as decisões. Assim, o DBA é responsável pelo controle geral do sistema em um nível técnico. Tem como **recurso primário** o banco de dados e como **recursos secundários** o SGBD e os softwares relacionados.

O DBA é o responsável por autorizar o acesso ao banco de dados, coordenar e monitorar seu uso, adquirir recursos de software e hardware conforme a necessidade e por resolver problemas tais como falhas de segurança ou demora no tempo de resposta do sistema. Segundo o Date uma lista de atividades associadas ao DBA contém as seguintes tarefas:

- Definir o esquema conceitual (às vezes conhecido como lógico)
- Definir o esquema interno
- Contatar com os usuários
- Definir restrições de segurança e integridade
- Monitorar o desempenho e responder a requisitos de mudanças.
- Definir normas de descarga e recarga (dumping)

Vejamos duas questões rápida sobre o assunto:



15. Ano: 2018 Banca: CESPE Órgão: EBSERH Prova: Analista de Tecnologia da Informação

Com relação a banco de dados, julgue o item seguinte.

Após um banco de dados ser criado, o administrador executa uma série de tarefas para dar permissão de acesso aos usuários que necessitam ler e gravar informações na base de dados. A responsabilidade de gerir os acessos ao banco de dados é do sistema gerenciador de banco de dados (SGBD).

Comentário: Observe que neste texto podemos verificar explicitamente o benefício do sistema de gerenciamento de banco de dados. O administrador de banco de dados vai atribuir aos

diversos usuários as permissões compatíveis com suas necessidades. Contudo, quem vai gerenciar o acesso, no sentido de autorizar o usuário a acessar o banco de dados e visualizar os dados, é o próprio SGBD. Desta forma, temos uma alternativa correta.

Gabarito: C.



16. BANCA: FCC ANO: 2010 ÓRGÃO: TCE-SP PROVA: AGENTE DA FISCALIZAÇÃO FINANCEIRA - PRODUÇÃO E BANCO DE DADOS

Considerando que os dados constituem um dos bens mais valiosos de uma empresa, é necessário que haja um papel que tenha a responsabilidade central pelos dados, principalmente entendendo as necessidades empresariais nos altos níveis da organização. Tal papel é mais adequadamente desempenhado pela

A administração de banco de dados.

B administração de dados.

C alta administração da organização.

D análise de sistemas.

E gerência de sistemas.

Comentário: Percebam que a questão trata da padronização dos dados corporativos. Entender a organização e sua relação com as informações de forma a gerenciar os dados de maneira eficiente é reponsabilidade da equipe de **administração de dados**. Hoje em dia um termo muito utilizado para os dados que permeiam vários setores das empresas ou órgãos públicos é o conceito de **dados mestres**. Eles são armazenados em um repositório central e distribuídos aos diversos usuários.

Gabarito: B.

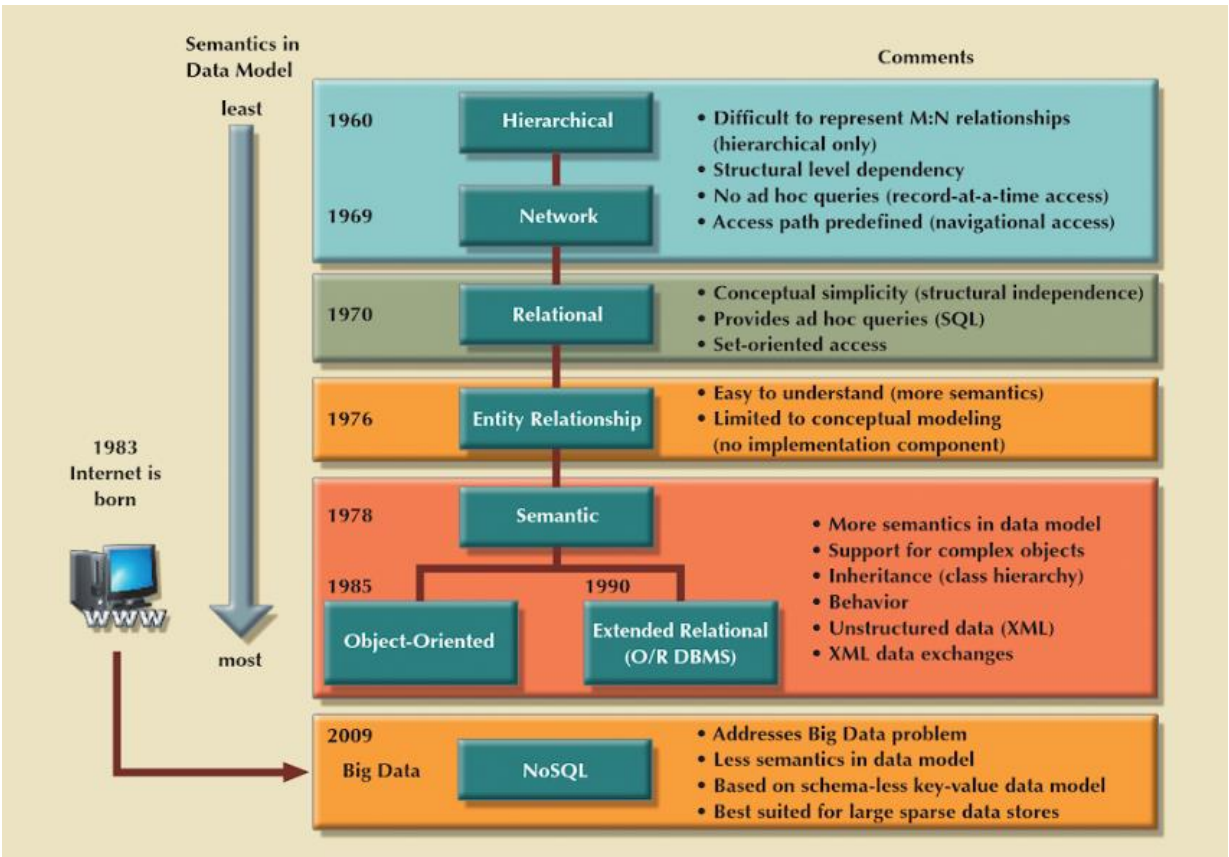
Você precisa entender que são dois os principais papeis presentes no trabalho diário de organização e manutenção de banco de dados. O **administrador de dados** que cuida do contexto estratégico dos dados, sua preocupação está em saber quais os dados serão armazenados em um banco de dados, quem são as pessoas que precisam dos mesmos e ainda, procurar padronizar os nomes dos atributos e outros objetos para que dados redundantes não sejam armazenados.



O outro indivíduo importante neste processo é o **administrador de banco de dados**. Nessa função encontramos todo o suporte técnico necessário para as atividades diárias de um sistema de banco de dados. O DBA distribui as permissões de acesso aos dados, criar os objetos que vão armazenar os dados, faz ajustes para que os mesmos funcionem de forma adequada.

EVOLUÇÃO HISTÓRICA DOS SGBDS

É interessante conhecer a evolução dos modelos de dados até os bancos de dados NoSQL. Afinal, quando começamos a tratar as informações em sistemas como elas eram armazenadas? Veja a figura abaixo para conhecer um pouco mais sobre os modelos de armazenamento de dados.



Os primeiros sistemas de gerenciamento de banco de dados são implementados no final da década de 1960. Charles Bachmann desenvolveu o primeiro SGBD chamado *Integrated Data Store* (IDS) enquanto trabalhava na Honeywell. Esse sistema usava o **modelo de rede** onde as relações de dados são representadas como um **grafo bidirecional**.

Contudo, o primeiro SGBD que obteve sucesso comercial foi desenvolvido pela IBM chamado *Information Management System* (IMS). Ele usava o **modelo hierárquico** no qual as relações entre os dados são representadas como **uma árvore**. Por incrível que pareça, ainda está em uso hoje no sistema de reservas SABRE da IBM na American Airlines. Nesta época a Conference On Data Systems Languages (CODASYL) definiu um modelo de rede mais padronizado.

Esses dois modelos, em rede e hierárquico, apresentavam problemas sérios, entre eles:

- O acesso ao banco de dados feito através de operações com o ponteiro de baixo nível.
- Detalhes de armazenamento dependiam do tipo de dados a serem armazenados.
- Para adicionar um campo no banco era necessário reescrever o esquema subjacente de acesso/modificação, em outras palavras o modelo de dados físico.
- Ênfase nos registros a serem processados, não na estrutura global.
- O usuário tinha que conhecer a estrutura física da BD, para fim de consulta das informações.

No geral os primeiros SGBDs eram muito complexos e inflexíveis, o que tornou cada vez mais difícil o trabalho, quando era necessária a adição de novos aplicativos ou a reorganização dos dados. Para resolver esses e outros problemas **Edgar (Ted) Codd**, conhecido com o pai do **modelo relacional**, trabalhando no laboratório da IBM em San Jose propôs no artigo “A Relational Model of Data for Large Shared Data Banks” a definição do modelo relacional.

Segundo Codd, o modelo fornece um meio de descrição de dados apresentando apenas a sua estrutura natural - isto é, sem sobreposição de qualquer estrutura adicional para efeitos de representação física dos dados. Assim, ele forneceu uma base para uma linguagem de dados de alto nível que permite obter a **independência** máxima **entre dados e programas** com a **representação** de um lado e a **estrutura física** da máquina do outro.

Em outras palavras, o modelo relacional consistiu na independência de dados e, na forma de acesso aos dados definida por uma linguagem. Em vez de processar um registro de cada vez, um programador pode usar o idioma para **especificar operações** individuais que seriam **realizados** em todo o **conjunto de dados**.

Devido à natureza técnica do artigo e a relativa complicação matemática presente no texto, o significado e proposição do artigo não foram prontamente identificados. Entretanto ele levou a IBM a montar um grupo de pesquisa conhecido como **System R** (Sistema R).

O projeto do Sistema R era criar um sistema de banco de dados relacional o qual eventualmente se tornaria um produto. Os primeiros protótipos foram utilizados por muitas organizações, tais como na *Sloan School of Management* (renomada escola de negócios norte-americana). Novas versões foram testadas com empresas de aviação para rastreamento de manufaturas em estoque.

Eventualmente o Sistema R evoluiu para SQL/DS, o qual posteriormente tornou-se **o DB2**. A **linguagem** criada pelo grupo do Sistema R foi a ***Structured Query Language*** (SQL) ou linguagem de consulta estruturada. Esta linguagem tornou-se um padrão na indústria para bancos de dados relacionais e, hoje em dia, é um padrão ISO (*International Organization for Standardization*). A linguagem SQL era originalmente conhecida como SEQUEL (*Structured English QUEry Language*). Depois teve seu nome modificado para SQL por problemas de patentes.

Em meados da década de 80 tornou-se óbvio que existiam várias **áreas onde bancos de dados relacionais não eram aplicáveis**, por causa dos tipos de dados envolvidos. Estas áreas incluíam medicina, multimídia e física nuclear, todas com necessidades de flexibilidade para definir como os dados seriam representados e acessados.

Este fato levou ao início de pesquisas em **bancos de dados orientados a objetos**, nos quais os usuários poderiam definir seus próprios métodos de acesso aos dados e como estes seriam representados e acessados. Ao mesmo tempo, linguagens de programação orientadas a objetos (*Object Oriented Programming* - POO), tais como C++, começaram a surgir na indústria.

No início de 1990, temos a aparição do primeiro Sistema de Gerenciamento de Banco de Dados Orientado a Objetos (SGBDOO), através da companhia *Objectivity*. Isso permitiu que usuários criassem sistemas de banco de dados para armazenar resultados de pesquisas como o CERN (maior laboratório que trabalha com partículas em pesquisas de física nuclear - europeu) e SLAC (Centro de Aceleração Nuclear - norte-americano), para mapeamento de rede de provedores de telecomunicações e para armazenar registros médicos de pacientes em hospitais, consultórios e laboratórios.

A evolução nos levou aos SGBDs orientados a objetos, mas a praticidade nos trouxe de volta para o modelo objeto-relacional, pois a grande maioria das empresas continuou utilizando os bancos de dados relacionais. Contudo esse modelo começou a apresentar outra lista de problemas ou desafios:

1. Dados na ordem de dezenas ou centenas de TB – abordagem de cluster é cara.
2. Poder de crescimento elástico horizontal – controle de transação ACID torna inviável com a elasticidade.
3. Fácil distribuição dos dados e/ou processamento – SGBD paralelos são caros.
4. Tipos de dados variados, complexos e/ou semiestruturados – modelo de dados objeto-relacional não resolve todos os requisitos.

Tivemos então o surgimento de um novo movimento no mercado em busca de uma solução que superasse tais problemas: **o movimento NoSQL**. Este teve sua origem em junho de 2009, para nomear um encontro promovido por **Johan Oskarsson** e **Eric Evans**, que teve como objetivo discutir o surgimento crescente de **soluções *open source* de armazenamento de dados distribuídos** não relacionais.

Podemos considerar NoSQL uma nova onda de SGBDs, pois propõe algumas alternativas ao modelo relacional, porém com uma grande diferença histórica: o movimento NoSQL não tem como objetivo invalidar ou promover a total substituição do modelo relacional, e sim o fim do modelo relacional como bala de prata, como a única solução correta ou válida. Inclusive, é importante entender que NoSQL não significa “no SQL” (não ao SQL), mas sim “not only SQL” (não só SQL).

> ***Curiosidade**: Ao que tudo indica o termo NoSQL foi criado em 1998 por Carlo Strozzi para nomear seu projeto open source, que tinha como objetivo ser uma implementação mais leve de um banco de dados relacional, porém sua principal característica era não expor a interface SQL. Portanto é bem irônico usar o termo NoSQL, criado para nomear um banco de dados relacional, para classificar soluções de armazenamento de dados não relacionais.*

Juntamente com NoSQL surge o conceito de Big Data. A definição mais tradicional usa a equação dos cinco Vs. Nela, Big Data = volume + variedade + velocidade + veracidade + valor, de dados. Cada característica desta é associada aos dados armazenados. É possível justificar cada uma delas. Vejamos:

Volume porque além dos dados gerados pelos sistemas transacionais, temos a imensidão de dados gerados pelos objetos na Internet das Coisas, como sensores e câmeras, e os dados gerados nas mídias sociais por meio de PCs, smartphones e tablets.
Variedade porque estamos tratando tanto de dados textuais estruturados como não estruturados como fotos, vídeos, e-mails e tuites.
Velocidade porque muitas vezes precisamos responder aos eventos quase que em tempo real, ou seja, estamos falando de geração e tratamento de dados em fluxo massivos.

O ponto de vista da **veracidade** também deve ser considerado, pois não adianta muita coisa lidar com a combinação "volume + velocidade + variedade" se não houver dados confiáveis. É necessário que haja processos que garantam a consistência dos dados.

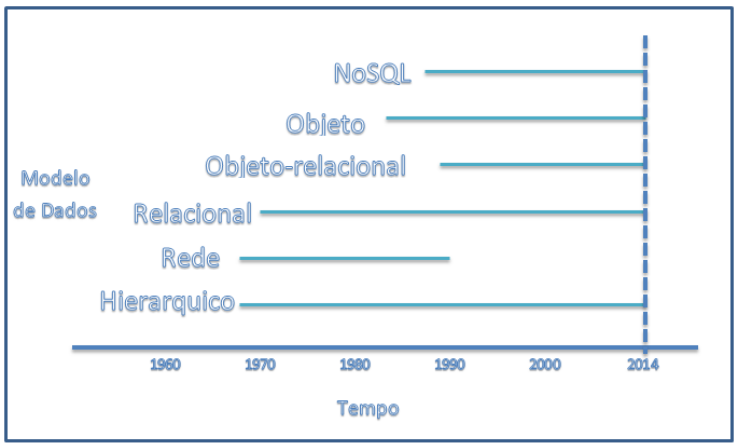
O último V, **valor**, considera que informação é poder, informação é patrimônio. A combinação "volume + velocidade + variedade + veracidade", além de todo e qualquer outro aspecto que caracteriza uma solução de Big Data, se mostrará inviável se o resultado não trouxer benefícios significativos e que compensem o investimento.

O que acontece agora? Diante destas definições, importantes à implementação de SGBDs que suportem a estratégia definida pelo Big Data podemos sugerir diferentes soluções. Bancos do tipo NoSQL são mais flexíveis, sendo inclusive compatíveis com um grupo de premissas que "compete" com as propriedades ACID dos SGBDs tradicionais: **a BASE** (**BAsically Available, Soft state, Eventually consistency** – Basicamente disponível, estado leve, eventualmente consistente).

Exemplos de bancos de dado NoSQL são o Cassandra, o MongoDB, o HBase, o CouchDB e o Redis. Mas, quando o assunto é Big Data, apenas um banco de dados do tipo não basta. É necessário também contar com ferramentas que permitam o tratamento correto do volume de dados. Neste ponto, o Hadoop é, de longe, a principal referência.

O **Hadoop** é uma plataforma *open source* desenvolvida especialmente para processamento e análise de grandes volumes de dados, sejam eles estruturados ou não estruturados. Pode-se dizer que o projeto teve início em meados de 2003, quando o Google criou um modelo de programação que distribui o processamento a ser realizado entre vários computadores para ajudar o seu mecanismo de busca a ficar mais rápido e livre das necessidades de servidores poderosos (e caros). Esta tecnologia recebeu o nome de **MapReduce**.

Vamos ficando por aqui, isso é o que nos interessa para o contexto histórico. Apresentamos abaixo uma figura com uma evolução dos modelos de dados ao longo do tempo.



Vamos agora fazer uma questão sobre o assunto. Desta vez, nossa questão é do CESPE.

CAIU na prova!

17. BANCA: CESPE ANO: 2014 ÓRGÃO: TJ-SE PROVA: ANALISTA JUDICIÁRIO – BANCO DE DADOS

Acerca de bancos de dados semiestruturados e bancos de dados NOSQL, julgue os itens subsecutivos.

[86] Bancos de dados NOSQL orientados a documentos são apropriados para o armazenamento de dados semiestruturados.

[87] Para garantir a eficiência das consultas a bancos de dados semiestruturados, é fundamental a adoção de técnica de indexação que leve em consideração, além das informações, as propriedades estruturais dos dados.

[88] Devido à escalabilidade esperada para os bancos de dados NOSQL, a implementação desses bancos utiliza modelos de armazenamento de dados totalmente distintos dos utilizados em sistemas relacionais.

Comentário: Vimos que um dos desafios que os banco de dados NoSQL tenta resolver tem relação com os tipos de dados variados, complexos e/ou semiestruturados. Assim podemos considerar a alternativa 86 como **correta**.

A questão 87 envolve alguns conceitos interessantes. Começa falando sobre dados semiestruturados, por exemplo, XML ou JSON. Consultas em bancos de dados semiestruturados consideram tanto a estrutura quanto os valores. Outra questão é a criação

de índice sobre um conjunto de dados semiestruturados. Para avaliar se um índice deve ou não ser criado é importante usar as informações sobre a estrutura dos dados e os valores armazenados. Neste caso, considerando a necessidade de um espaço maior para armazenamento e do custo de manutenção, a criação do índice deve melhorar a performance para ser de fato implementado. Logo, a assertiva está **correta**.

A alternativa 88 vai exigir conhecimento sobre os modelos de armazenamento utilizados por bancos de dados NoSQL. Quando tratamos de bases de dados NoSQL podemos classificá-las em quatro diferentes tipos, são eles:

Chave/valor (Key/Value): conhecidos como tabelas de hash distribuídas. Armazenam objetos indexados por chaves, e facilita a busca por esses objetos a partir de suas chaves.

Orientados a Documentos: os documentos dos bancos são coleções de atributos e valores onde um atributo pode ser multivalorado. Em geral, os bancos de dados orientados a documento não possuem esquema, ou seja, os documentos armazenados não precisam possuir uma estrutura em comum. Essa característica faz deles boas opções para o armazenamento de dados semiestruturados.

Colunar: Bancos relacionais normalmente guardam os registros das tabelas contiguamente no disco. Por exemplo, caso se queira guardar id, nome e endereço de usuários em um banco de dados relacional, os registros seriam:

Id1, Nome1, Endereço1;

Id2, Nome2, Endereço2.

Essa estrutura torna a escrita muito rápida, pois todos os dados de um registro são colocados no disco com uma única escrita no banco. Também é eficiente caso se queira ler registros inteiros. Mas para situações onde se quer ler algumas poucas colunas de muitos registros, essa estrutura é pouco eficiente, pois muitos blocos do disco terão de ser lidos.

Para esses casos onde se quer otimizar a leitura de dados estruturados, bancos de dados de famílias de colunas são mais interessantes, pois eles guardam os dados contiguamente por coluna.

O exemplo anterior em um banco de dados dessa categoria ficaria:

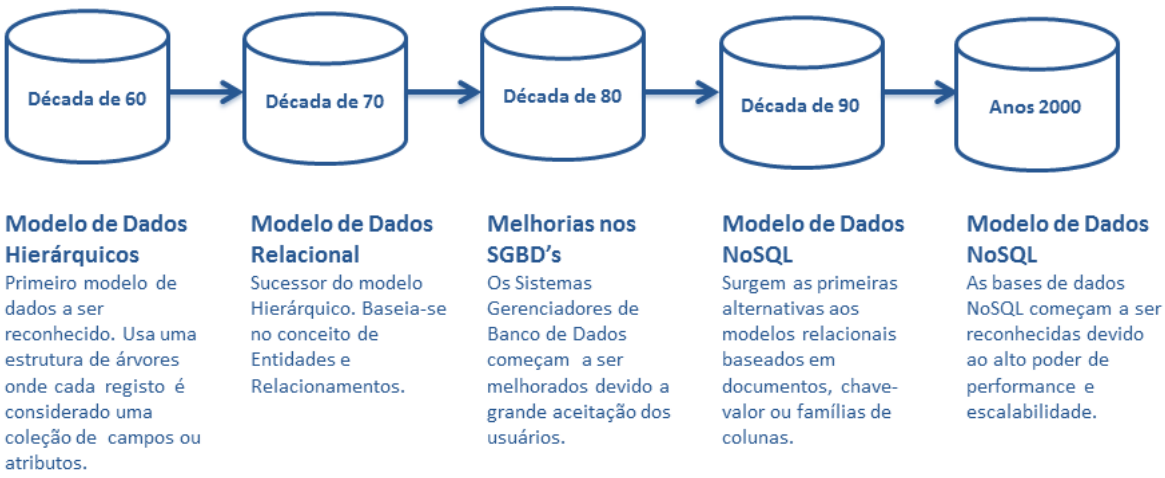
Id1, Id2; Nome1, Nome2; Endereço1, Endereço2.

Os bancos de dados de famílias de colunas são mais interessantes para processamento analítico online (OLAP). ***Bigtable*** é uma implementação da Google dessa categoria de bancos de dados.

Orientado a Grafos: diferente de outros bancos de dados NoSQL, esse está diretamente relacionado a um modelo de dados estabelecido, o modelo de grafos. A ideia desse modelo é representar os dados e/ou o esquema dos dados como grafos dirigidos, ou como estruturas que generalizem a noção de grafos. O modelo de grafos é aplicável quando “informações sobre a interconectividade ou a topologia dos dados são mais importantes, ou tão importante quanto os dados propriamente ditos”. Possui três componentes básicos: os nós (são os vértices do

grafo), os relacionamentos (são as arestas) e as propriedades (ou atributos) dos nós e relacionamentos.

Agora vamos voltar a questão, precisamos responder a seguinte pergunta: nenhum dos modelos acima tem relação com bancos de dados relacionais? Ou ainda, não é possível criar estruturas em SGBDs relacionais que representem esses conceitos? Veja a figura abaixo:



Para finalizar vamos fazer apenas um comentário sobre escalabilidade: A escalabilidade em um banco de dados relacional pode ocorrer de duas formas: horizontal e vertical. A forma horizontal ocorre pela utilização de mais equipamentos e particiona a estrutura de dados de acordo com critérios estabelecidos. A forma vertical ocorre pelo aumento da capacidade do equipamento em que o sistema gerenciador de banco de dados está instalado. Bases de dados NoSQL têm como um de seus motivadores o baixo custo para realizar uma escalabilidade horizontal, o que torna possível o uso de equipamentos mais acessíveis. Além disso, proporciona um modelo de particionamento nativo (Sharding).

Gabarito: C C E.

Vejamos agora uma questão da FCC sobre o assunto.

CAIU na prova!

18. BANCA: FCC ANO: 2017 ÓRGÃO: TST PROVA: TÉCNICO JUDICIÁRIO – PROGRAMAÇÃO

[57] Considere que um determinado tipo de banco de dados organiza os dados na forma de uma pirâmide, onde o registro principal no topo da pirâmide é chamado registro raiz. Os registros são organizados como pai e filho onde um registro filho sempre tem apenas um registro pai ao qual ele está ligado, como em uma árvore familiar normal. Em contrapartida, um registro pai pode ter mais de um registro filho a ele ligado.

Trata-se do Banco de Dados

a) hierárquico.

b) relacional.

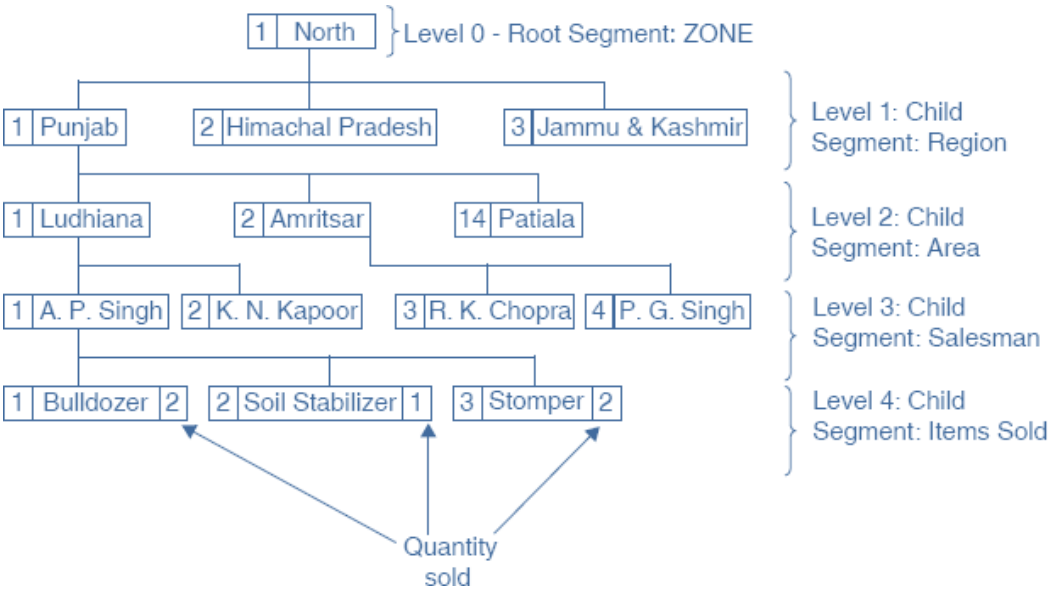
c) orientado a objeto.

d) objeto-relacional.

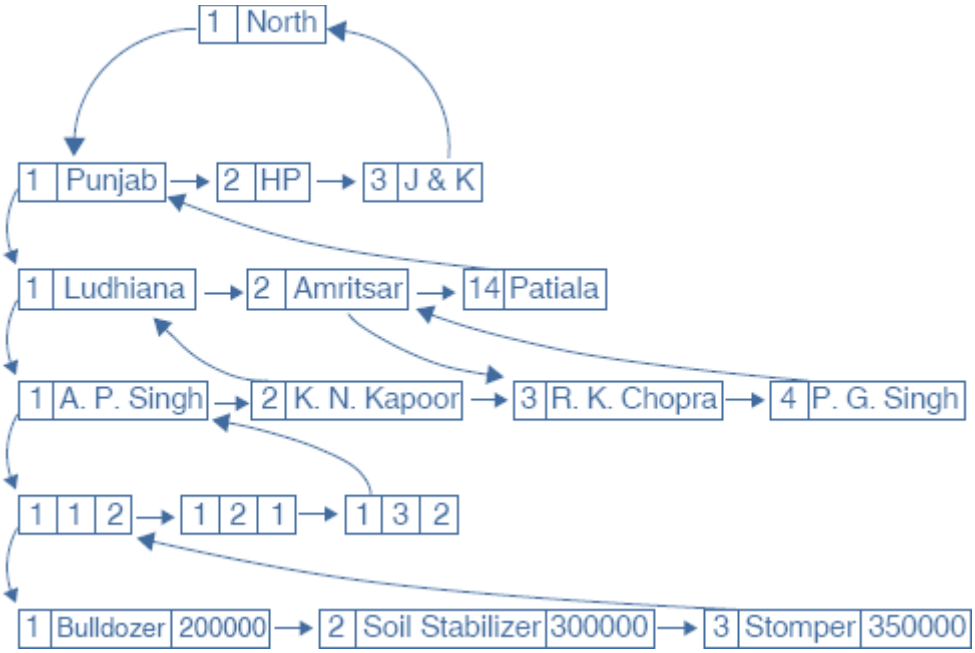
e) de rede.

Comentário: Vamos fazer uma rápida revisão sobre os conceitos associados aos modelos de banco de dados orientados a registros:

O **modelo hierárquico** de dados foi o primeiro modelo de dados desenvolvido na década de 1960. O modelo de dados hierárquico foi chamado de IMS (Information Management System) e foi desenvolvido pela IBM e pela Rockwell Company. Ele foi amplamente utilizado durante os anos 1960 e 1970. As entidades e as relações entre entidades foram estruturadas e gerenciadas com a ajuda de uma **estrutura em forma de árvore**. Nesta árvore, existe uma **raiz** que está relacionada aos seus **filhos**. Uma raiz é conhecida como pai. Podemos observar que essa é a nossa resposta. Veja a figura abaixo com a representação de uma estrutura hierárquica.



Já o **modelo de dados de rede** representa dados usando o **link** entre os registros. O registro pai é chamado de **Registro Proprietário**, e o registro filho é chamada de **Registro de Membro**. Se os registros Proprietário e Membro estiverem relacionados com o relacionamento muitos-para-muitos, eles estão conectados através do registro do conector que é conhecido como **Set**. Agora vejamos um modelo semelhante ao visto acima, agora com as características do modelo em rede:



Temos ainda o conceito de **modelo relacional** foi dado por E. F. Codd, em 1970, em seu documento histórico sobre o modelo de dados relacionais. No modelo relacional, os dados são representados em uma forma tabular denominada, **relação** (tabela), e estão associados a relacionamentos. Portanto, o nome desse modelo é o modelo de dados relacional. Cada **entidade** é convertida em relação e a associação é tratada através de **chaves primárias e estrangeiras**. Cada ocorrência da entidade é conhecida como **tupla** (registro) e a característica de uma entidade é chamada de **atributo** (coluna). É muito fácil representar a relação muitos-para-muitos usando o modelo de dados relacionais. O modelo relacional é amplamente utilizado em todo o mundo, hoje em dia, para armazenar dados. Vejamos agora um exemplo de um banco de dados relacional de uma livraria on-line.

BOOK

ISBN	Book_title	Category	Price	Copyright_date	Year	Page_count	P_ID
001-354-921-1	Ransack	Novel	22	2005	2006	200	P001
001-987-760-9	C++	Textbook	25	2004	2005	800	P001

PUBLISHER

P_ID	Pname	Address	State	Phone	Email_ID
P001	Hills Publications	12, Park street, Atlanta	Georgia	7134019	H_pub@hills.com

REVIEW

R_ID	ISBN	Rating
A002	001-987-760-9	6.0
A006	001-354-921-1	7.5
A008	001-987-760-9	7.2

Gabarito: A

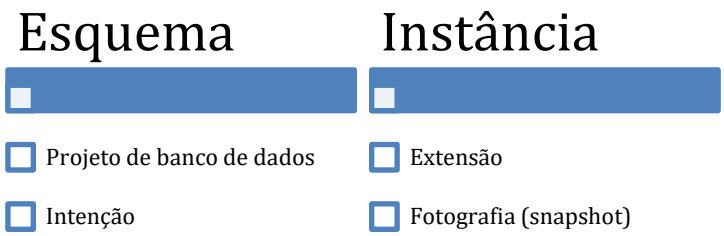
MODELO DE DADOS E ARQUITETURA EM TRÊS ESQUEMAS

MODELO DE DADOS, INSTÂNCIAS, ESQUEMAS

Na parte anterior do nosso curso, tratamos das definições relacionadas aos termos básicos SGBD, BD, SBD e dicionário de dados. Agora vamos tratar de outro grupo de conceitos. Este grupo está relacionado ao entendimento dos diversos participantes do processo de definição de um banco de dados, para definir os diferentes **níveis de modelos de dados** que dão origens aos esquemas. A partir desses esquemas podemos obter instâncias.

Antes de adentrar no estudo dos modelos de dados queria que você entendesse a diferença entre esquema e instância. O **esquema** é a definição da estrutura do modelo e a **instância** se refere aos dados armazenados em um esquema em um momento do tempo.

Para descrever os modelos utilizamos os elementos fornecidos por cada um deles e construímos esquemas. Conhecido com o projeto geral do BD é basicamente a descrição do banco, conhecido também como a **intenção** ou **conotação**. Baseado nesse esquema é possível criar uma instância do BD. Essa coleção de dados armazenados no banco de dados em um determinado instante é o próprio banco de dados, também denominada de **extensão**.



Um **modelo de dados** fornece um significado necessário para permitir a **abstração dos dados**, ocultando detalhes de armazenamento. Pode ser visto como uma coleção de conceitos que são usados para descrever a estrutura de um banco de dados. Cada modelo deve definir uma coleção de ferramentas conceituais para as seguintes tarefas:

(1) descrição de dados
(2) relacionamentos entre eles
(3) a semântica dos dados
(4) restrições de consistência.

Boa parte dos modelos também dá **suporte a operações**, algumas dessas operações podem representar o aspecto dinâmico ou comportamento de uma aplicação de banco de dados.

Os **modelos de dados** podem ser divididos em três categorias de acordo com os tipos de conceitos usados para descrever a estrutura do banco de dados. No nível mais alto temos os modelos de dados **conceituais** que apresentam os dados da forma como os usuários finais percebem. Em um nível intermediário está a classe de modelos de dados **representacionais** (ou de implementação) que fornece um entendimento aos envolvidos no processo de desenvolvimento do BD, mas já introduz

informações sobre a forma pela qual os dados são armazenados dentro de um computador. O último nível apresenta os modelos de dados **físicos** ou de baixo nível. Observem na figura a seguir:

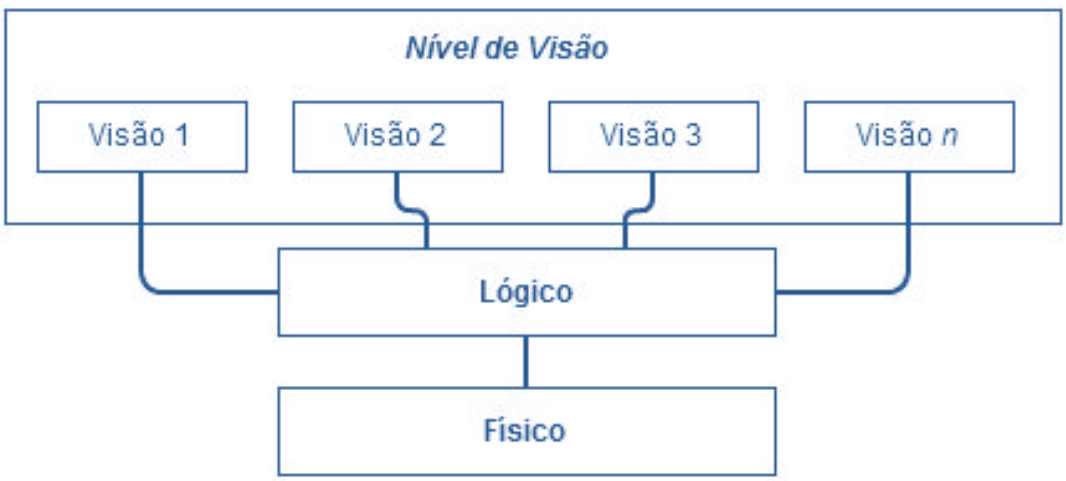


Após conhecermos a divisão de modelo de dados, vamos mudar nosso foco para outra classificação presente no contexto de banco de dados. Silberschatz apresenta um conceito de níveis de abstração. O nível de abstração **mais baixo** ou **físico** descreve **como** os dados realmente são armazenados. Este nível descreve em detalhes estruturas de dados complexas.

O próximo nível de abstração descreve **que** dados estão armazenados no banco de dados e que relações existem entre eles. O **nível lógico** descreve o banco de dados inteiro em termos de um pequeno número de estruturas relativamente simples. Embora a implementação das estruturas simples no nível lógico possa envolver estruturas complexas em nível físico, o usuário do nível lógico não precisa tomar ciência desta complexidade.

O nível de abstração mais alto descreve apenas parte do banco de dados. Muitos usuários de sistema de banco de dos não precisam de todas as informações armazenadas. Em vez disso, eles precisam apenas de uma parte do banco de dados. O **nível de visão** (*view*) existe para simplificar sua interação com o sistema, que pode fornecer muitas visões para o mesmo banco de dados.

A hierarquia de níveis de abstração de dados pode ser vista na figura abaixo:



ARQUITETURA TRÊS ESQUEMAS

Agora que já entendemos de modelos de dados, instâncias e esquemas, vamos passar para a arquitetura em três esquemas. O American National Standards Institute (ANSI) através do Standards Planning and Requirements Committee (SPARC) estabeleceu um padrão para o desenvolvimento de tecnologias de base de dados, definindo uma arquitetura de três níveis independentes: interno, conceitual e externo.

Essa arquitetura tem por objetivo separar o usuário da aplicação do banco de dados físico. Possuem, logicamente, os esquemas definidos em três níveis distintos:

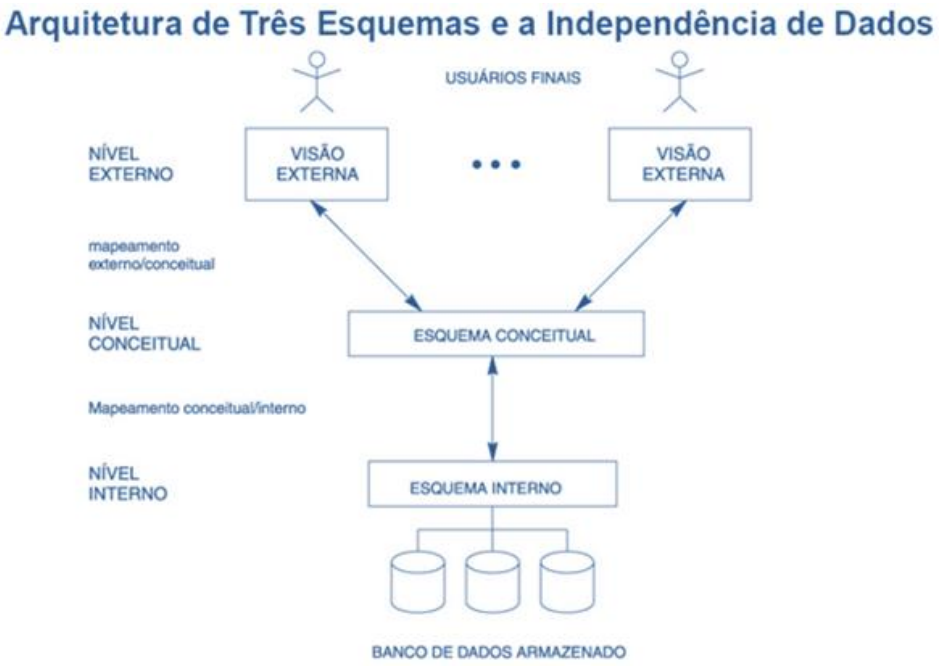
◉ Nível interno - (também conhecido como nível de armazenamento) é o mais próximo do meio de armazenamento físico - ou seja, é aquele que se ocupa do modo como os dados são fisicamente armazenados dentro do sistema.
◉ Nível conceitual - (também conhecido como nível lógico de comunidade, ou às vezes apenas nível lógico, sem qualificação) é um nível "indireto" entre os outros dois.
◉ Nível externo ou visão - (também conhecido como nível lógico do usuário) é o mais próximo dos usuários – ou seja, é aquele que se ocupa do modo como os dados são vistos por usuários individuais.

Podemos usar partir da arquitetura de três esquemas para explicar o último assunto desta nossa parte da aula. A **independência dos dados** que nada mais é do que a capacidade de alterar o esquema em um nível dos sistemas de banco de dados sem alterar o esquema no nível mais alto ou, em outras palavras a habilidade de modificar a definição de um esquema em um nível sem afetar a definição do esquema em um nível mais alto.

Segundo Navathe é possível definir dois tipos de independência de dados:

1. **Independência lógica de dados** - a capacidade de alterar o esquema conceitual sem ter de alterar os esquemas externos ou os aplicativos.
2. **Independência física de dados** - a capacidade de alterar o esquema interno sem ter de alterar o esquema conceitual.

Apresentamos a seguir uma figura que descreve os níveis da arquitetura em três esquemas. Observem que pela definição de independência de dados é necessário a existência de um nível superior ao esquema alterado. Desta forma só temos o conceito para os níveis conceitual e interno da figura.



De uma forma simples, cada um dos níveis possui uma função dentro das suas respectivas abstrações. O nível de visão do usuário determina a parte em que o usuário tem acesso. O nível conceitual identifica os dados armazenados e suas relações. Por fim, o nível físico é o nível mais baixo de abstração, define a maneira como os dados estão armazenados. Vejamos mais algumas questões sobre o assunto:

CAIU na prova!

19. Ano: 2016 Banca: CESPE Órgão: TRE-PI Prova: Analista Judiciário - Análise de Sistemas

A respeito das características de um SGBD e das atividades de administração de banco de dados, assinale a opção correta.

a) Para fins práticos, é necessário distinguir diferentes cardinalidades máximas, que podem ser maiores ou iguais a zero.

b) A característica autodescritiva de um banco de dados define que o banco de dados contém o próprio dado assim como uma descrição desses dados e suas restrições. Essas descrições e restrições estão armazenadas no catálogo (dicionário) do SGBD.

c) A independência física de dados consiste na habilidade de modificar o esquema conceitual sem a necessidade de reescrever os programas aplicativos. As modificações no nível conceitual são necessárias quando a estrutura lógica do banco de dados é alterada.

d) Na linguagem SQL, os comandos DDL GRANT e ROLLBACK permitem a implementação de um controle de acesso discricionário, criando e retirando permissões no banco de dados.

e) A coleção das informações armazenadas em um banco de dados, em determinado momento, corresponde ao esquema do banco de dados.

Comentário: Observe que a descrição presente na alternativa B está perfeitamente adequada. Tente encontrar os erros das demais alternativas. Apenas por curiosidade, na alternativa D, a linguagem SQL possui algumas subdivisões dos comandos. Os comandos GRANT e REVOKE server para implementação do controle de acesso discricionário. O ROLLBACK, juntamente com o COMMIT e SAVEPOINT servem para controle de transações.

Outro ponto importante, é que, na alternativa A, a cardinalidade máxima tem que ser obrigatoriamente maior ou igual a 1. A justificativa para isso será dada na próxima aula.

Gabarito: B.



20. Ano: 2014 Banca: CESPE Órgão: TJ-CE Prova: Analista Judiciário - Ciências Computação

Considerando o sistema gerenciador de banco de dados (SGBD), assinale a opção correta acerca de bancos de dados.

a) Enquanto a DDL (Data Definition Language) é utilizada para definir a estrutura do banco de dados, a SDL (Storage Definition Language) é utilizada para especificar o esquema conceitual e seus mapeamentos com o esquema interno.

b) A informação armazenada no catálogo do SGBD é denominada metamodelo.

c) Na independência de dados do programa, propriedade do SGBD, a estrutura dos arquivos de dados é armazenada no catálogo separadamente dos programas de acesso.

d) Na arquitetura de três esquemas de um banco de dados, o nível conceitual é responsável por descrever de forma detalhada as estruturas de armazenamento físico, incluindo os relacionamentos entre as tabelas.

e) Na arquitetura de três esquemas, a capacidade de alterar o esquema interno sem ter de alterar o esquema conceitual consiste na independência lógica de dados.

Comentário: Nesta questão, tão importante quanto saber que a alternativa C está correta, e esse assunto nós já vimos no início da aula, é entender por que as outras estão erradas. Perceba primeiramente que, se formos teoricamente precisos deveríamos trocas SGBD por SDB. (lembra?!, SGBD + DB = SGBD) A questão é que alguns autores falam da abordagem de sistemas de gerenciamento de banco de dados. Assim, essa acaba sendo a menos errada das alternativas.

Na alternativa A, o examinador sugere que SDL, é utilizada para descrever um esquema conceitual. Vejam que SDL está relacionada com armazenamento, e, quando existe de forma separada em uma linguagem, é utilizada para definir os detalhes do nível interno.

Na alternativa B, no lugar de meta modelo, o correto seria metadados.

Na alternativa D, o nível descrito é o nível interno.

Por fim, a alternativa E trata de independência física e não lógica como descrito no texto.

Gabarito: C.



21. Ano: 2016 Banca: CESPE Órgão: TCE-PA Prova: Auditor de Controle Externo - Área Informática - Analista de Sistema

Julgue o item subsequente, no que se refere a sistemas de gerenciamento de bancos de dados (SGBD).

Independência lógica de dados refere-se à capacidade de alterar o esquema conceitual sem a necessidade de alterar os esquemas externos ou os programas de aplicação.

Comentário: Lembre-se que a independência lógica ou conceitual atual no nível intermediário da arquitetura em 3 esquemas. Ela está relacionada ao fato de que modificações nesta camada da arquitetura não gera necessidade de alteração nos esquemas externos ou programas de aplicação. Desta forma, a alternativa encontra-se correta.

Gabarito: C.



22. Ano: 2016 Banca: CESPE Órgão: TCE-PA Prova: Auditor de Controle Externo - Área Informática - Administrador de Banco de Dados

Com relação a sistemas gerenciadores de bancos de dados (SGBD), julgue o próximo item.

No nível conceitual da arquitetura de três camadas de banco de dados, cada esquema externo descreve a parte do banco que interessa a determinado grupo de usuários e oculta desse grupo o restante do banco de dados.

Comentário: Perceba que essa questão apresenta uma casca de banana das mais malvadas. Ele mistura o nível conceitual da arquitetura em três esquemas com o nível externo. Sabemos que os esquemas externos estão associados as diferentes visões dos grupos de usuários. Tal fato reflete os interesses distintos de cada grupo quanto ao acesso as informações do banco de dados. Assim, podemos marcar nosso gabarito como errado.

Gabarito: E.



23. BANCA: FCC ANO: 2010 ÓRGÃO: TCE-SP PROVA: AGENTE DA FISCALIZAÇÃO FINANCEIRA - CONHECIMENTOS BÁSICOS

As três visões da arquitetura básica de um SGBD, pela ordem, desde a mais próxima do usuário até a mais distante, são:

A externa, conceitual e interna.

B externa, interna e conceitual.

C conceitual, interna e externa.

D conceitual, externa e interna.

E interna, conceitual e externa.

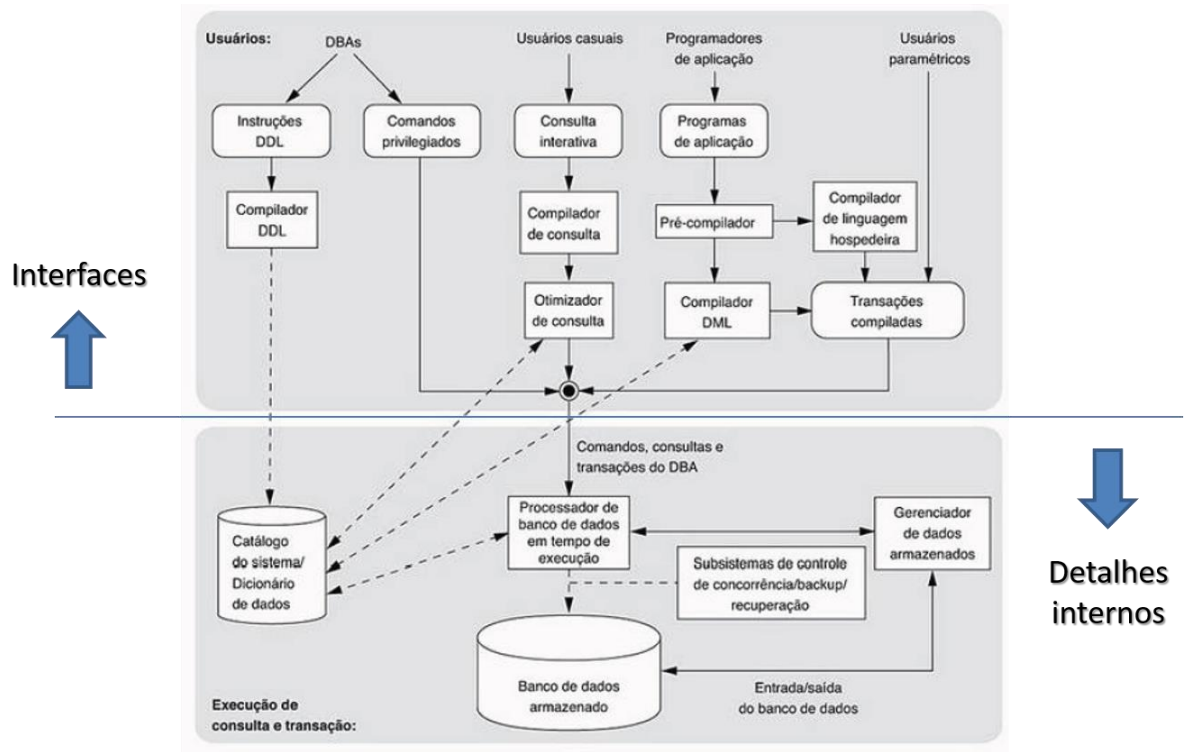
Comentário. Ao analisar as alternativas temos que ter em mente duas informações relevantes de enunciados. Primeiro é solicitado as três visões da arquitetura, desta forma podemos concluir que o examinador está se referindo a arquitetura em três esquemas. A outra informação é que a questão pede para você colocar em ordem crescente, da mais próxima do usuário até a mais distante. Com essas duas informações podemos relembrar dos conceitos vistos anteriormente e descrever a seguinte ordem: **externa, conceitual e interna**.

Gabarito: A.

AMBIENTE DO SISTEMA DE BANCO DE DADOS

A figura a seguir apresenta, de forma simplificada, os componentes típicos de um SGBD. Perceba que temos uma divisão em duas partes. A parte superior mostra os usuários e a forma de interação deles com o SGBD. Cada tipo de usuário possui a sua interface de relacionamento.

A parte inferior ilustra os detalhes internos do SGBD, eles são responsáveis pelo armazenamento dos dados e processamento das transações dos usuários. Vejam que o **banco de dados** e o **catálogo do SGBD** estão representados por um cilindro, representando o fato de serem armazenados permanentemente. Os **discos rígidos** representam uma das diversas possibilidades de armazenamento não volátil ou permanente dos dados.



Nas próximas linhas vou descrever um pouco dos componentes da figura acima. Os conceitos podem ser úteis para que você entenda um pouco da "caixa preta" por trás de um SGBD.

O acesso aos discos rígidos é controlado pelo sistema operacional ou pelo próprio SGBD. O controle vai escalonar as operações de leitura e escrita sobre o disco. Quando consideramos a execução de uma transação sobre o banco de dados, o intervalo de tempo gasto na transferência de dados entre a memória e o disco é considerado um gargalo no processamento. Para otimizar o processo alguns SGBDs podem contar com um **módulo de gerenciamento de buffer** que planeja a troca de dados entre a memória principal e o disco.

Outro módulo, o **gerenciador de dados armazenados**, controla o acesso às informações do SGBD que estão armazenadas, seja no catálogo de dados ou no bando de dados. Ele utiliza os serviços básicos do sistema operacional para executar operações de entrada/saída (leitura/escrita) de baixo nível entre o disco e a memória principal.

Observa-se, na parte superior da figura, as interfaces para os diferentes usuários do sistema. De um lado temos **os usuários casuais** que trabalham com interfaces interativas para formular consultas. Logo em seguida, visualizamos os **programadores de aplicação** que usam uma linguagem de programação hospedeira para ter acesso aos dados. Por fim, temos os **usuários paramétricos** que inserem valores para os parâmetros predefinidos pelas transações.

No parágrafo anterior, falamos apenas das manipulações de dados que podem ser feitas sobre as bases de dados. Essas operações incluem consultas, inserções, deleções e atualizações das informações armazenadas. Contudo, é necessário, antes de manipularmos os dados, construir as estruturas do banco de dados. A linguagem de definição de dados ou *data definition language* é utilizada pelo **DBA** para descrever os objetos presentes na base de dados. Uma tabela do modelo relacional é um exemplo de objeto que podemos criar por meio de uma **instrução DDL** em nossa base de dados.

Quando o DBA digita uma instrução ou comando DDL, essa é enviada ao **compilador DDL.** Um compilador transforma o código fonte (da linguagem de programação) em um código em que o computador entenda. O compilador da DDL processa as definições de esquema especificadas e armazena as descrições de esquema (metadados) no catálogo do SGBD. Esse fluxo pode ser observado no lado esquerdo da figura anterior.

Outra função do DBA é o ajuste fino ou *tuning* do sistema de gerenciamento de banco de dados, bem como a configuração de parâmetros que são feitos por meio dos **comandos privilegiados**. Apenas para exemplificar, um comando presente na maioria dos SGBDs é o *REORG*, serve para reorganizar uma tabela ou índice na estrutura física do banco de dados.

Os usuários casuais interagem usando alguma interface de consulta interativa. Essas consultas são analisadas e validadas pela exatidão da sintaxe da consulta, os nomes de arquivos e elementos de dados, e assim por diante, por um **compilador de consulta**. Essa consulta interna está sujeita a melhorias feitas pelo otimizador de consultas, que se preocupa com o rearranjo e a possível reordenação de operações, com a eliminação de redundâncias e uso dos algoritmos e índices corretos durante a execução.

Ele consulta o catálogo do sistema em busca de informações estatísticas e outras informações físicas sobre os dados armazenados, gerando um código executável. Este por sua vez realiza as operações necessárias para a consulta e faz chamadas ao **processador em tempo de execução** (falaremos mais sobre ele logo mais).

Os programadores de aplicação escrevem programas em linguagens hospedeiras, como Java e C#, que são submetidas a um **pré-compilador**. Este extrai os comandos DML do programa de aplicação. Para entender melhor como funciona essa divisão vamos partir para um exemplo prático. No exemplo abaixo temos um código Java com um comando SQL. O comentário no código delimita o início da instrução SQL dentro do código Java.

```
public boolean verificarUsuario(String login, String senha){
    String sql = "";
    Connection conn = conectarBD();
    //INSTRUÇÃO SQL
    sql += "select nome from usuarios ";
    sql +="where login = " + "'" + login + "'";
    sql += " and senha = " + "'" + senha + "';";
    try{
        Statement st = conn.createStatement();
        ResultSet rs = st.executeQuery(sql);
        if(rs.next()){
            result = true;
            nome = rs.getString("nome");}
    }catch (Exception e) {  }
    return result; }
```

Observem que o comando *select*. Ele está escrito na linguagem SQL considerada uma DML. Esses comandos são enviados ao compilador DML para serem compilados em código objeto com acesso ao banco de dados. O restante do programa é enviado ao compilador da linguagem hospedeira. Os códigos objetos para os comandos DML e o restante do programa são ligados (‘linkados’) formando uma transação programada ou compilada.

As transações programadas são executadas repetidas vezes pelos usuários paramétricos, que apenas fornecem os parâmetros para as transações. No nosso exemplo anterior, as informações necessárias são *login* e senha. Cada execução é considerada uma transação separada. Outro exemplo de transação é o saque no caixa eletrônico, no qual o número da conta e o valor são fornecidos como parâmetros.

Na parte inferior da figura temos o **processador de banco de dados em tempo de execução (PBDTE).** Ele é responsável por executar os comandos privilegiados, os planos de consulta executáveis e as transações programadas. Para isso são utilizadas informações e dados estatísticos do catálogo do sistema. O PBDTE também trabalho com o **gerenciador de dados armazenados**.

Os sistemas de controle de concorrência, backup e recuperação são apresentados como um módulo da figura. Eles são integrados ao processador de banco de dados em tempo de execução para fins de gerenciamento de transações. Você precisa ter em mente que esses controles são necessários para o perfeito funcionamento do SGBD. O backup é utilizado durante a recuperação caso alguma falha aconteça. A concorrência entre transações deve existir dentro de limites bem definidos para evitar que o banco de dados entre em um estado inconsistente.

A figura apresentada não pretende descrever um SGBD específico nem esgotar suas funcionalidades. Nossa ideia foi ilustrar os módulos básicos de um SGBD e estruturar seu raciocínio. Lembre-se que um SGBD é um sistema informatizado. Para executar todas as suas tarefas sua implementação é dividida em diferentes módulos. O SGBD interage ainda com o sistema operacional quando o acesso ao disco rígido é necessário. Vamos fazer uma questão da FGV que trata do assunto.



24. Ano: 2010 Banca: FGV Órgão: BADESC Cargo: Analista de Sistemas

Os objetivos dos compiladores DDL, DML e DCL são, respectivamente:

A) criar os objetos do banco de dados, manipular (recuperação, inserção, remoção e alteração) de dados nos objetos criados pela DDL e fornecer privilégio de acesso às informações.

B) fornecer privilégio de acesso às informações, criar os objetos do banco de dados e manipular (recuperação, inserção, remoção e alteração) de dados nos objetos criados pela DDL.

C) manipular (recuperação, inserção, remoção e alteração) de dados nos objetos criados pela DML, criar os objetos do banco de dados e fornecer privilégio de acesso às informações.

D) fornecer privilégio de acesso às informações, manipular (recuperação, inserção, remoção e alteração) de dados nos objetos criados pela DDL e criar os objetos do banco de dados.

E) criar os objetos do banco de dados, fornecer privilégio de acesso às informações e manipular (recuperação, inserção, remoção e alteração) de dados nos objetos criados pela DDL.

Comentário: Vimos que o compilador DDL permite que o DBA emita comandos para a criação dos objetos do banco de dados, esses vão fornecer informações sobre as estruturas das tabelas como atributos e restrições. O compilador DML é responsável por transformar o código SQL para manipulação dos dados armazenados.

Por fim, temos o ***Data Control Language – DCL,*** essa parte da linguagem SQL vai permitir aos administradores de banco de dados a distribuição de privilégios de acesso sobre a base de dados. Com essa informação, podemos marcar tranquilamente a resposta na alternativa A.

Gabarito: A.

Espero que você tenha entendido nossa proposta de apresentar o ambiente do SGBD. Acabamos aqui o nosso primeiro conjunto de assuntos teóricos envolvidos na introdução dos sistemas de bancos de dados.

QUESTÕES COMENTADAS

Aproveitamos este espaço para complementar seu entendimento sobre o assunto, sempre colocando **pinceladas extras de conteúdo**. Nosso objeto é criar em você a capacidade de utilizar seu conhecimento, atrelado a algumas heurísticas, para resolver de forma rápida e segura as questões na hora da prova. Alguns assuntos abaixo serão detalhados em aulas subsequentes. Aproveite as questões para ir se **familiarizando com o vocabulário** de banco de dados.



25. Ano: 2018 Banca: CESPE Órgão: STM Prova: Técnico Judiciário - Programação de Sistemas

Acerca dos conceitos de normalização de dados e dos modelos de dados, julgue o item subsequente.

Comparativamente aos usados pelos usuários leigos, os modelos de dados utilizados por programadores são considerados menos abstratos, pois contêm mais detalhes de como as informações estão organizadas internamente no banco de dados.

Comentário: Os usuários leigos, por terem um conhecimento mais limitado a respeito da tecnologia de banco de dados, devem ter uma visão mais abstrata dos dados quando comparados com os programadores. Estes, por terem conhecimento e formação específica podem ter mais detalhes da estrutura de armazenamento dos dados.

Gabarito: C



26. Ano: 2018 Banca: CESPE Órgão: CGM de João Pessoa – PB Prova: Auditor Municipal de Controle Interno - Desenvolvimento de Sistemas

A respeito de bancos de dados, julgue o item a seguir.

Nos bancos de dados construídos sob a concepção do modelo hierárquico, os dados são estruturados em hierarquia ou árvores cujos nós contêm ocorrências de registros, e cada registro consiste em uma coleção de atributos.

Comentário: Um **banco de dados hierárquico** consiste em uma coleção de **registros** que são conectados uns aos outros por meio de **ligações**. Um registro é uma coleção de campos, cada qual contendo apenas um valor de dados. Uma ligação é uma associação entre exatamente dois registros. O modelo hierárquico difere do modelo de rede na organização de registros como **coleção de árvores** em vez de como grafos arbitrários.

Um diagrama com estrutura de árvore é um esquema para um banco de dados hierárquico. Tal diagrama consiste em dois componentes básicos: retângulos, que correspondem a tipos de registro, e linhas, que correspondem a ligações. O diagrama com estrutura de árvore serve para

os mesmos propósitos que um diagrama entidade-relacionamento; a saber, ele especifica a estrutura lógica geral do banco de dados.

Após essa rápida reflexão teórica sobre o assunto, podemos marcar nossa reposta como correta.

Gabarito: C



27. Ano: 2018 Banca: CESPE Órgão: CGM de João Pessoa – PB Prova: Auditor Municipal de Controle Interno - Desenvolvimento de Sistemas

A respeito de bancos de dados, julgue o item a seguir.

Um banco de dados é uma coleção de dados que são organizados de forma randômica, sem significado implícito e de tamanho variável, e projetados para atender a uma proposta específica de alta complexidade, de acordo com o interesse dos usuários.

Comentário: Essa questão procura contrapor as propriedades listas pelo Navathe. Um banco de dados precisa representar algum aspecto do mundo real, ser logicamente coerente com algum significado inerente e possuir um grupo de usuários. Enfim, um banco de dados possui

[1] alguma fonte da qual o dado é derivado,

[2] algum grau de interação com eventos no mundo real e

[3] um público que está ativamente interessado em seu conteúdo.

Logo, a afirmação da questão está incoerente, portanto, incorreta.

Gabarito: E



28. Ano: 2017 Banca: CESPE Órgão: SEDF Cargo: Analista de gestão educacional – Especialidade: tecnologia da informação Questão: 109

Julgue os itens seguintes, a respeito de estruturas em programação e de arquiteturas de bancos de dados.

[109] O esquema do nível externo de uma arquitetura de três esquemas oculta os detalhes das estruturas de armazenamento físico e se concentra na descrição de entidades, tipos de dados, conexões, operações de usuários e restrições.

Comentário: O objetivo da arquitetura três esquemas é separar as aplicações do usuário do banco de dados físico. São três níveis:

Nível interno: tem um esquema interno, que descreve a estrutura do armazenamento físico do banco de dados. O esquema interno usa um modelo de dados físico e descreve detalhes completos do armazenamento de dados.

Nível conceitual: tem esquema conceitual, que descreve a estrutura do banco de dados inteiro para uma comunidade de usuários. Oculta detalhes das estruturas de armazenamento físico e se concentra na descrição das entidades, tipos de dados, relacionamentos, operações e restrições. Um modelo representativo é usado para descrever o esquema conceitual.

Nível externo ou de visão: inclui uma série de esquemas externos ou visões do usuário. Cada esquema externo descreve a parte do banco de dados em que um grupo de usuários em particular está interessado e oculta o restante do banco. Como o nível anterior, cada esquema externo é implementado por um modelo de dados representativo.

Após analisarmos as afirmações teóricas acima, podemos concluir que a alternativa 109 da prova do SEDF está incorreta.

Gabarito: E.



29. BANCA: CESPE ANO: 2014 ÓRGÃO: TJ-SE PROVA: ANALISTA JUDICIÁRIO - SUPORTE E INFRAESTRUTURA

Julgue os itens a seguir, relativos à administração de banco de dados e ao sistema de gerenciamento de banco de dados (SGBD).

[69] Os dados físicos de um banco de dados podem ser acessados diretamente por meio de qualquer sistema, sem a necessidade de utilização do SGBD.

[70] Uma das atribuições do administrador de banco de dados é definir a estratégia que determinará como será feito o becape do banco de dados.

Comentário: Aproveitaremos essa questão para fazer um comentário técnico e prático sobre o assunto em cada uma das alternativas.

Na assertiva 69 diz que o acesso aos bancos de dados pode ser feito fisicamente sem a necessidade de um SGBD. Essa afirmação é **falsa**, se você lembrar das conexões que são feitas aos bancos, você precisa passar as informações de endereço (IP ou URL), porta, *schema* e um driver ou conector, que vai permitir uma comunicação correta entre o sistema e o banco de dados, além da autenticação do usuário. Sendo, portanto, **incorreta** a alternativa.

A alternativa 70 faz menção a uma das tarefas técnicas executadas pelo DBA. É necessário definir um roteiro ou procedimento de backup do banco de dados. Neste são definidos a periodicidade, o tipo de backup, a mídia de armazenamento e outras especificidades. Essa é uma das tarefas mais importantes feitas pelo DBA. Sendo assim, a alternativa está **correta**.

Gabarito: E C.



30. BANCA: CESPE ANO: 2013 ÓRGÃO: MC PROVA: ANALISTA DE NÍVEL SUPERIOR - TECNOLOGIA DA INFORMAÇÃO

Julgue os itens a seguir, acerca dos fundamentos e das finalidades do banco de dados.

[51] Atualmente, os bancos de dados são utilizados para armazenar e processar dados de caracteres em geral, não apresentando recursos para tratar dados multimídias, como filmes e fotografias.

[52] Uma característica fundamental do banco de dados e dos antigos sistemas de arquivos é o inter-relacionamento dos dados, sem redundâncias ou duplicação de dados.

[53] Para definir e manter os dados em um banco é necessário o uso de sistemas de aplicação, o que caracteriza a dependência de dados, que é um fundamento do banco de dados.

Comentário: Vamos analisar as alternativas acima.

Começando pelo item 51. Veja que a questão sugere que existe uma limitação nos tipos de dados armazenados em bancos de dados. Sabemos que todos os SGBDs comerciais que implementam SQL possuem o tipo de dados BLOB – *Binary Large Object*. Nele é possível gravar qualquer informação em formato binário como arquivos multimídias. Falaremos mais sobre tipos de dados na aula sobre SQL. Podemos então marcar a alternativas como **incorreta**.

Observem que a alternativa 52 tenta comparar os sistemas de arquivos com os bancos de dados colocando uma das suas principais diferenças como uma similaridade entre eles. A diminuição da redundância e da duplicação ocorre primordialmente com a evolução dos sistemas de arquivo para os sistemas de bancos de dados. Sendo assim, alternativa também está **incorreta**.

Vimos que por estarem isolados, dados e aplicações, criam um conceito chamado independência de dados do programa. Este só é possível por conta da **abstração de dados**. A abstração de dados permite a criação de diferentes níveis de modelos. Falaremos mais sobre os níveis de abstração quando apresentarmos os modelos de dados. Mas, por enquanto, o nosso conhecimento já é suficiente para analisarmos a questão 53 como **errada**.

Gabarito: E E E



31. BANCA: CESPE ANO: 2014 ÓRGÃO: TJ-SE PROVA: ANALISTA JUDICIÁRIO - SUPORTE E INFRAESTRUTURA

Julgue os itens a seguir, relativos à administração de banco de dados e ao sistema de gerenciamento de banco de dados (SGBD).

[71] Um SGBD deve gerenciar o acesso múltiplo aos dados de uma tabela sem ocasionar perda da integridade dessas informações.

Comentário: Vejam que o SGBD possui como uma das suas características fazer o controle de concorrência entre diferentes usuários ou transações que acessam uma mesma tabela. Isso é importante para garantir a integridade dos registros e a consistência das transações executadas. Desta forma, podemos marcar a alternativa como correta!

Gabarito: C.



32. BANCA: CESPE ANO: 2015 ÓRGÃO: MPOG PROVA: ANALISTA - ANALISTA EM TECNOLOGIA DA INFORMAÇÃO

Acerca de sistema de gerenciamento de banco de dados (SGBD), julgue os seguintes itens.

[115] Os dados armazenados em um SGBD são acessados por um único usuário de cada vez, sendo impedido o acesso concorrente aos dados.

[116] O SGBD proporciona um conjunto de programas que permite o acesso aos dados sem exposição dos detalhes de representação e armazenamento de dados, por meio de uma visão abstrata dos dados, conhecida como independência de dados.

Comentário: Vimos na nossa aula que uma das características de SGBDs é o suporte a usuários simultâneos. O SGBD faz o controle de concorrência entre transações que tentam acessar a mesma tabela ao mesmo tempo. A alternativa 115 está incorreta.

Uma das características que já conhecemos é independência entre dados e programas. Vimos que a partir do momento em que temos um dicionário de dados, é possível excluir da estrutura dos programas a definição dos dados presentes nos mesmos. Agora isolados, dados e aplicações, criam um conceito chamado **independência de dados do programa**. Este só é possível por conta da **abstração de dados**. A abstração de dados permite a criação de diferentes níveis de modelos. Por isso a alternativa 116 está correta.

Gabarito: E C.



33. BANCA: CESPE ANO: 2015 ÓRGÃO: DEPEN PROVA: AGENTE PENITENCIÁRIO FEDERAL - TECNOLOGIA DA INFORMAÇÃO

No que diz respeito a linguagens de programação e banco de dados, julgue os itens a seguir.

[101] Os níveis interno, externo e conceitual da arquitetura de um banco de dados são responsáveis, respectivamente, por gerenciar o modo como os dados serão armazenados fisicamente, por gerenciar o modo como os dados serão vistos pelos usuários e por representar todo o conteúdo de informações do banco de dados.

Comentário: Falamos sobre a arquitetura três esquemas. Observem que a alternativa acima está correta. Vá se acostumando com os termos e conceitos.

Gabarito: C.



34. BANCA: CESPE ANO: 2013 ÓRGÃO: MC PROVA: ANALISTA DE NÍVEL SUPERIOR - TECNOLOGIA DA INFORMAÇÃO

Julgue os itens subsequentes, quanto à administração de banco de dados.

[58] O administrador do banco de dados não deve gerenciar a utilização do espaço em disco nos servidores, pois sua função limita-se à utilização de ferramentas de gerenciamento com o objetivo de garantir a disponibilidade dos serviços de banco.

[59] A administração de banco de dados abrange a definição e a alteração de esquema, que, em alguns casos, são tarefas importantes para melhorar o desempenho do banco de dados.

Comentário: Nesta questão tratamos mais uma vez das funções do administrador de banco de dados. Responsável pelo suporte técnico as atividades do SGBD, uma das suas atividades envolve monitorar o crescimento das bases de dados em disco. Quando o espaço disponível se aproxima de zero ele deve trabalhar para alocar mais espaço. Esse serviço faz parte da garantia de disponibilidade do banco de dados. Sendo assim, a alternativa 58, pode ser considerar errada.

O ajuste fino ou *tuning* dos esquemas de banco de dados são de reponsabilidade do DBA. Esses ajustes ajudam a melhorar o desempenho do banco de dados. Temos a alternativa 59 como correta.

Gabarito: E C.



35. BANCA: CESPE ANO: 2013 ÓRGÃO: ANTT PROVA: ANALISTA ADMINISTRATIVO - INFRAESTRUTURA DE TI

No que diz respeito às funções do administrador de dados e à elaboração e implantação de projeto de banco de dados, julgue os itens que se seguem.

[87] Se, durante a elaboração de um projeto de banco de dados relacional, houver dependência entre entidades, trata-se de representação que pode ser concebida utilizando-se o conceito de entidade fraca.

[88] Em um projeto de banco de dados relacional, não é possível representar um autorrelacionamento do tipo N:N.

[89] Uma das funções do administrador de dados é padronizar os dados, documentando as definições e descrições dos itens de dados.

Comentário: Nesta questão começamos com um conceito de entidade fraca. Ele nos diz que se uma entidade depende de outra entidade para existir, ela é denominada entidade fraca ou dependente. Vejam, portanto, que a questão 87 está correta.

Talvez não seja tão fácil de visualizar, mas podemos pensar no atributo gerente em uma tabela de funcionários, ele indica quem é o gerente de um funcionário específico. Agora vamos supor que um funcionário possa ter mais de um gerente. Neste caso teríamos um autorelacionamento N:N. Contudo, quando formos transformar o modelo ER para o modelo Relacional, como não podemos ter atributos multivalorados, teríamos que criar outra tabela para representar esse relacionamento. A alternativa 88 é falsa.

Lembrem-se que autorelacionamento é um conceito relacionado a modelagem conceitual. Quando pensamos em um modelo ER, dizemos que um gerente pode ter vários funcionários e o atributo gerencia da entidade funcionário faz essa função. Contudo quando vamos passar para o modelo relacional, neste caso, de fato temos que criar outra tabela. Um relacionamento N-N iria criar redundância na tabela. Vejam que a pergunta é se podemos representar autorelacionamento, sabemos que é possível. No caso de um relacionamento N-N, temos que criar outra tabela de fato.

Vimos que uma das funções do administrador de dados é trabalhar com os dados e sua organização de forma global. Ele tem a responsabilidade de padronizar valores e definir domínios que sejam coerentes com o negócio. Podemos avaliar a alternativa 89 como correta, pois está de acordo com nosso conhecimento.

Gabarito: C E C.

EXERCÍCIOS



1. Ano: 2016 Banca: CESPE Órgão: TCE-SC Cargo: Auditor de TI

Julgue os itens a seguir, acerca de dado, informação, conhecimento e inteligência.

90 Define-se informação como significado, ou seja, como registros icônicos e simbólicos — fonéticos ou numéricos — e signos — linguísticos, lógicos ou matemáticos —, por meio dos quais se representam atos, conceitos ou instruções.

91 O atributo de inteligência depende mais da qualidade da informação disponível do que da sua quantidade, tendo, portanto, natureza qualitativa.



2. Ano: 2018 Banca: CESPE Órgão: TCE-PB Cargo: Auditor de Constas Públicas Questão: 96

Com relação a dados estruturados e não estruturados, assinale a opção correta.

A Dados não estruturados podem ser caracterizados por arquivos de diversos tipos — textos, imagens, vídeos, entre outros —, cujas estruturas não são descritas implicitamente.

B Por padrão, documentos do tipo XML (eXtensible Markup Language) são estruturados.

C Dados não estruturados de um mesmo grupo possuem as mesmas descrições e, consequentemente, os mesmos atributos.

D Por padrão, dados não estruturados são organizados em blocos semânticos.

E A alta heterogeneidade facilita as consultas aos dados não estruturados, desde que estes estejam ligados por ponteiros.



3. Ano: 2016 Banca: CESPE Órgão: TCE-SC Cargo: Auditor de TI

A respeito de dados estruturados, não estruturados e abertos, julgue os itens subsequentes.

92 Dados abertos são os dados de livre utilização, reutilização e redistribuição, exigindo-se, no máximo, créditos à autoria e compartilhamento pela mesma licença.

93 Em se tratando de dados estruturados, a informação de esquema está mesclada aos valores dos dados, e cada objeto de dados pode ter atributos diferentes, que não são conhecidos com antecedência. Essa característica os diferencia de dados não estruturados.



4. Ano: 2016 Banca: CESPE Órgão: TCE-SC Prova: Auditor Fiscal de Controle Externo - Informática

Com relação aos bancos de dados relacionais, julgue o próximo item.

O catálogo de um sistema de gerenciamento de banco de dados relacional armazena a descrição da estrutura do banco de dados e contém informações a respeito de cada arquivo, do tipo e formato de armazenamento de cada item de dado e das restrições relativas aos dados.



5. Ano: 2010 Banca: CESPE Órgão: Banco da Amazônia Prova: Técnico Científico - Tecnologia da Informação

O dicionário de dados é uma das principais ferramentas para a administração dos dados corporativos. Por meio da engenharia reversa, pode-se armazenar os modelos de dados, as estruturas de dados, seus relacionamentos e toda a documentação necessária para garantir facilidade na localização e manipulação dos dados. Acerca dos papéis do administrador de dados (AD) e dos dicionários de dados, julgue os itens a seguir.

[1] O dicionário de dados é considerado um subconjunto das funções de um catálogo de sistema.

[2] O catálogo do sistema é um repositório com função de armazenar as definições dos esquemas dos bancos de dados.



6. Ano: 2018 Banca: CESPE Órgão: TCE-PB Cargo: Auditor de Constas Públicas Questão: 97.

A respeito de SGBDs, assinale a opção correta.

A Um SGBD, por definição, não é flexível, dada a dificuldade de mudar a estrutura dos dados quando os requisitos mudam.

B Um SGBD é um software que não prevê as funções de definição, recuperação e alteração de dados, sendo essa tarefa a função básica de um sistema de banco de dados.

C A consistência de dados é o princípio que determina a manutenção de determinado dado em vários arquivos diferentes.

D Conforme o princípio da atomicidade, caso ocorra erro em determinada transação, todo o conjunto a ela relacionado será desfeito até o retorno ao estado inicial, como se a transação nunca tivesse sido executada.

E O controle de concorrência é o princípio que garante e permite a manipulação, no mesmo momento, de um mesmo dado por mais de uma pessoa ou um sistema.



7. Ano: 2018 Banca: CESPE Órgão: EBSERH Prova: Analista de Tecnologia da Informação

Com relação a banco de dados, julgue o item seguinte.

Após um banco de dados ser criado, o administrador executa uma série de tarefas para dar permissão de acesso aos usuários que necessitam ler e gravar informações na base de dados. A responsabilidade de gerir os acessos ao banco de dados é do sistema gerenciador de banco de dados (SGBD).



8. BANCA: FCC ANO: 2014 ÓRGÃO: TJ-AP PROVA: ANALISTA JUDICIÁRIO - BANCO DE DADOS - DBA

A redundância controlada de dados em um sistema de banco de dados ocorre quando

A um programa está ciente da múltipla representação de uma dada informação e garante o sincronismo entre as diversas representações.

B a responsabilidade pela manutenção do sincronismo entre as múltiplas representações de uma dada informação é compartilhada entre o programa e o usuário.

C os dados mais importantes são duplicados a cada backup do sistema, visando aumentar a garantia da recuperação da informação em caso de problemas.

D a responsabilidade pela manutenção do sincronismo entre as múltiplas representações de uma dada informação é do usuário.

E um programa está ciente da múltipla representação de uma dada informação, mas não garante o sincronismo entre as diversas representações.



9. BANCA: FCC ANO: 2010 ÓRGÃO: TCE-SP PROVA: AGENTE DA FISCALIZAÇÃO FINANCEIRA - PRODUÇÃO E BANCO DE DADOS

Considerando que os dados constituem um dos bens mais valiosos de uma empresa, é necessário que haja um papel que tenha a responsabilidade central pelos dados, principalmente entendendo as necessidades empresariais nos altos níveis da organização. Tal papel é mais adequadamente desempenhado pela

A administração de banco de dados.

B administração de dados.

C alta administração da organização.

D análise de sistemas.

E gerência de sistemas.



10. Ano: 2016 Banca: CESPE Órgão: TRE-PI Prova: Analista Judiciário - Análise de Sistemas

A respeito das características de um SGBD e das atividades de administração de banco de dados, assinale a opção correta.

a) Para fins práticos, é necessário distinguir diferentes cardinalidades máximas, que podem ser maiores ou iguais a zero.

b) A característica autodescritiva de um banco de dados define que o banco de dados contém o próprio dado assim como uma descrição desses dados e suas restrições. Essas descrições e restrições estão armazenadas no catálogo (dicionário) do SGBD.

c) A independência física de dados consiste na habilidade de modificar o esquema conceitual sem a necessidade de reescrever os programas aplicativos. As modificações no nível conceitual são necessárias quando a estrutura lógica do banco de dados é alterada.

d) Na linguagem SQL, os comandos DDL GRANT e ROLLBACK permitem a implementação de um controle de acesso discricionário, criando e retirando permissões no banco de dados.

e) A coleção das informações armazenadas em um banco de dados, em determinado momento, corresponde ao esquema do banco de dados.



11. Ano: 2014 Banca: CESPE Órgão: TJ-CE Prova: Analista Judiciário - Ciências Computação

Considerando o sistema gerenciador de banco de dados (SGBD), assinale a opção correta acerca de bancos de dados.

a) Enquanto a DDL (Data Definition Language) é utilizada para definir a estrutura do banco de dados, a SDL (Storage Definition Language) é utilizada para especificar o esquema conceitual e seus mapeamentos com o esquema interno.

b) A informação armazenada no catálogo do SGBD é denominada metamodelo.

c) Na independência de dados do programa, propriedade do SGBD, a estrutura dos arquivos de dados é armazenada no catálogo separadamente dos programas de acesso.

d) Na arquitetura de três esquemas de um banco de dados, o nível conceitual é responsável por descrever de forma detalhada as estruturas de armazenamento físico, incluindo os relacionamentos entre as tabelas.

e) Na arquitetura de três esquemas, a capacidade de alterar o esquema interno sem ter de alterar o esquema conceitual consiste na independência lógica de dados.



12. Ano: 2016 Banca: CESPE Órgão: TCE-PA Prova: Auditor de Controle Externo - Área Informática - Analista de Sistema

Julgue o item subsequente, no que se refere a sistemas de gerenciamento de bancos de dados (SGBD).

Independência lógica de dados refere-se à capacidade de alterar o esquema conceitual sem a necessidade de alterar os esquemas externos ou os programas de aplicação.



13. Ano: 2016 Banca: CESPE Órgão: TCE-PA Prova: Auditor de Controle Externo - Área Informática - Administrador de Banco de Dados

Com relação a sistemas gerenciadores de bancos de dados (SGBD), julgue o próximo item.

No nível conceitual da arquitetura de três camadas de banco de dados, cada esquema externo descreve a parte do banco que interessa a determinado grupo de usuários e oculta desse grupo o restante do banco de dados.



14. BANCA: FCC ANO: 2010 ÓRGÃO: TCE-SP PROVA: AGENTE DA FISCALIZAÇÃO FINANCEIRA - CONHECIMENTOS BÁSICOS

As três visões da arquitetura básica de um SGBD, pela ordem, desde a mais próxima do usuário até a mais distante, são:

A externa, conceitual e interna.

B externa, interna e conceitual.

C conceitual, interna e externa.

D conceitual, externa e interna.

E interna, conceitual e externa.



15. Ano: 2017 Banca: CESPE Órgão: SEDF Cargo: Analista de gestão educacional – Especialidade: tecnologia da informação Questão: 109

Julgue os itens seguintes, a respeito de estruturas em programação e de arquiteturas de bancos de dados.

[109] O esquema do nível externo de uma arquitetura de três esquemas oculta os detalhes das estruturas de armazenamento físico e se concentra na descrição de entidades, tipos de dados, conexões, operações de usuários e restrições.



16. BANCA: CESPE ANO: 2014 ÓRGÃO: TJ-SE PROVA: ANALISTA JUDICIÁRIO - SUPORTE E INFRAESTRUTURA

Julgue os itens a seguir, relativos à administração de banco de dados e ao sistema de gerenciamento de banco de dados (SGBD).

[69] Os dados físicos de um banco de dados podem ser acessados diretamente por meio de qualquer sistema, sem a necessidade de utilização do SGBD.

[70] Uma das atribuições do administrador de banco de dados é definir a estratégia que determinará como será feito o becape do banco de dados.



17. BANCA: CESPE ANO: 2013 ÓRGÃO: MC PROVA: ANALISTA DE NÍVEL SUPERIOR - TECNOLOGIA DA INFORMAÇÃO

Julgue os itens a seguir, acerca dos fundamentos e das finalidades do banco de dados.

[51] Atualmente, os bancos de dados são utilizados para armazenar e processar dados de caracteres em geral, não apresentando recursos para tratar dados multimídias, como filmes e fotografias.

[52] Uma característica fundamental do banco de dados e dos antigos sistemas de arquivos é o inter-relacionamento dos dados, sem redundâncias ou duplicação de dados.

[53] Para definir e manter os dados em um banco é necessário o uso de sistemas de aplicação, o que caracteriza a dependência de dados, que é um fundamento do banco de dados.



18. BANCA: CESPE ANO: 2014 ÓRGÃO: TJ-SE PROVA: ANALISTA JUDICIÁRIO - SUPORTE E INFRAESTRUTURA

Julgue os itens a seguir, relativos à administração de banco de dados e ao sistema de gerenciamento de banco de dados (SGBD).

[71] Um SGBD deve gerenciar o acesso múltiplo aos dados de uma tabela sem ocasionar perda da integridade dessas informações.



19. BANCA: CESPE ANO: 2015 ÓRGÃO: MPOG PROVA: ANALISTA - ANALISTA EM TECNOLOGIA DA INFORMAÇÃO

Acerca de sistema de gerenciamento de banco de dados (SGBD), julgue os seguintes itens.

[115] Os dados armazenados em um SGBD são acessados por um único usuário de cada vez, sendo impedido o acesso concorrente aos dados.

[116] O SGBD proporciona um conjunto de programas que permite o acesso aos dados sem exposição dos detalhes de representação e armazenamento de dados, por meio de uma visão abstrata dos dados, conhecida como independência de dados.



20. BANCA: CESPE ANO: 2015 ÓRGÃO: DEPEN PROVA: AGENTE PENITENCIÁRIO FEDERAL - TECNOLOGIA DA INFORMAÇÃO

No que diz respeito a linguagens de programação e banco de dados, julgue os itens a seguir.

[101] Os níveis interno, externo e conceitual da arquitetura de um banco de dados são responsáveis, respectivamente, por gerenciar o modo como os dados serão armazenados fisicamente, por gerenciar o modo como os dados serão vistos pelos usuários e por representar todo o conteúdo de informações do banco de dados.



21. BANCA: CESPE ANO: 2013 ÓRGÃO: MC PROVA: ANALISTA DE NÍVEL SUPERIOR - TECNOLOGIA DA INFORMAÇÃO

Julgue os itens subsequentes, quanto à administração de banco de dados.

[58] O administrador do banco de dados não deve gerenciar a utilização do espaço em disco nos servidores, pois sua função limita-se à utilização de ferramentas de gerenciamento com o objetivo de garantir a disponibilidade dos serviços de banco.

[59] A administração de banco de dados abrange a definição e a alteração de esquema, que, em alguns casos, são tarefas importantes para melhorar o desempenho do banco de dados.



22. Ano: 2018 Banca: CESPE Órgão: STM Prova: Técnico Judiciário - Programação de Sistemas

Acerca dos conceitos de normalização de dados e dos modelos de dados, julgue o item subsequente.

Comparativamente aos usados pelos usuários leigos, os modelos de dados utilizados por programadores são considerados menos abstratos, pois contêm mais detalhes de como as informações estão organizadas internamente no banco de dados.



23. Ano: 2018 Banca: CESPE Órgão: CGM de João Pessoa – PB Prova: Auditor Municipal de Controle Interno - Desenvolvimento de Sistemas

A respeito de bancos de dados, julgue o item a seguir.

Nos bancos de dados construídos sob a concepção do modelo hierárquico, os dados são estruturados em hierarquia ou árvores cujos nós contêm ocorrências de registros, e cada registro consiste em uma coleção de atributos.



24. Ano: 2018 Banca: CESPE Órgão: CGM de João Pessoa – PB Prova: Auditor Municipal de Controle Interno - Desenvolvimento de Sistemas

A respeito de bancos de dados, julgue o item a seguir.

Um banco de dados é uma coleção de dados que são organizados de forma randômica, sem significado implícito e de tamanho variável, e projetados para atender a uma proposta específica de alta complexidade, de acordo com o interesse dos usuários.



25. BANCA: CESPE ANO: 2013 ÓRGÃO: ANTT PROVA: ANALISTA ADMINISTRATIVO - INFRAESTRUTURA DE TI

No que diz respeito às funções do administrador de dados e à elaboração e implantação de projeto de banco de dados, julgue os itens que se seguem.

[87] Se, durante a elaboração de um projeto de banco de dados relacional, houver dependência entre entidades, trata-se de representação que pode ser concebida utilizando-se o conceito de entidade fraca.

[88] Em um projeto de banco de dados relacional, não é possível representar um autorrelacionamento do tipo N:N.

[89] Uma das funções do administrador de dados é padronizar os dados, documentando as definições e descrições dos itens de dados.

GABARITO

1.	E C	10.	B	19.	E C
2.	A	11.	C	20.	C
3.	C E	12.	21. C	21.	E C
4.	C	13.	E	22.	C
5.	E C	14.	A	23.	C
6.	D	15.	E	24.	E
7.	C	16.	E C	25.	C E C
8.	A	17.	E E E		
9.	B	18.	C		

RESUMO

Antes de finalizar de fato a nossa aula vamos revisar alguns tópicos importantes que forma vistos que gostaria que você fixasse. Vamos começar com uma citação ao Silberchatz.

> *Um **sistema de gerenciamento de banco de dados (SGBD)** é uma coleção de dados inter-relacionados e um conjunto de programas para acessar esses dados. A **coleção de dados**, normalmente conhecida como **banco de dados**, contém informações relevantes para uma empresa. O **principal objetivo** de um SGBD é proporcionar uma forma de armazenar e recuperar informações de um banco de dados de maneira **conveniente e eficiente**.*

Outro ponto importante são os conceitos hierárquicos de banco de dados vistos ao longo da aula. Reconhecemos que em determinado momento parece que elas estão tratando do mesmo conceito.

Eu gosto sempre de dizer que os conceitos acabam se complementando. Os **modelos de dados** podem ser associados a forma de representação dos dados. Eles variam de acordo com a percepção das pessoas que interagem com o sistema em cada um dos níveis. Sendo assim, temos o **modelo conceitual**, que representa os dados para os usuários de negócio; o **modelo representativo ou lógico**, que estrutura os dados para implementação em um SGBD; e os **modelos físicos,** que descrevem a forma como os dados são armazenados nos dispositivos de armazenamento, por exemplo, disco rígido.

Por outro lado, temos as hierarquias definidas pelos **níveis de abstração** e pela **arquitetura em 3 esquemas**. Uma característica importante em ambas é que no nível mais alto de abstração elas separam o banco de dados de acordo com grupos de usuários. Imagine que cada setor de uma empresa tenha acesso a apenas uma parte dos dados. O **nível de visão** ou a **visão externa** descreve exatamente essa perspectiva.

Abaixo deste nível temos o nível de lógico na hierarquia de abstração e o nível conceitual na arquitetura em três esquemas. Você deve estar se perguntando como esses níveis se juntam com os modelos de dados? Os modelos de dados (conceitual e lógico) são unificados nesta camada. Para resolver esse problema alguns autores chamam os modelos conceituais de modelos lógicos baseados em objetos e os modelos representativos de modelos lógicos baseados em registros.

Por fim, temos o nível físico ou interno. Esse é igual em todas as hierarquias. Este nível é responsável pela forma de organização dos dados no disco e por descrever os métodos de acesso para o mesmo.

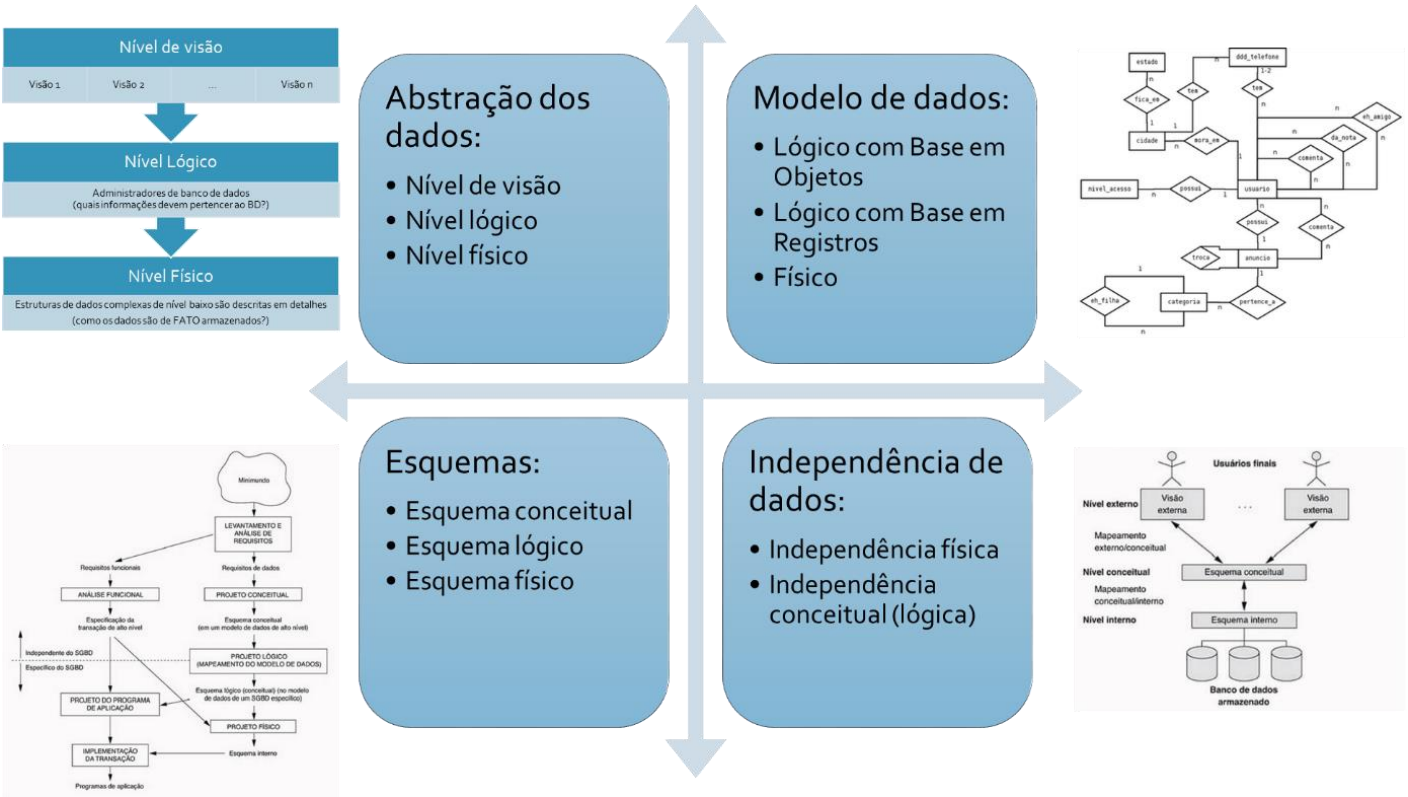
Para finalizar temos que lembrar do conceito de independência de dados que trata da capacidade de alterar o esquema em um nível do sistema de banco de dados sem ter que alterar o esquema nos outros níveis. A independência de dados lógicos é a capacidade de alterar o esquema conceitual sem afetar os esquemas externos ou os programas de aplicativos. O esquema conceitual pode ser alterado devido a mudanças nas restrições ou adição de novos itens de dados ou a remoção de itens de dados existentes.

Já a independência de dados físicos é a capacidade de alterar o esquema interno sem afetar o esquema conceitual ou externo. Um esquema interno pode ser alterado por vários motivos, como por exemplo, para criar uma estrutura de acesso adicional, alterar a estrutura de armazenamento, etc. A separação do esquema interno do esquema conceitual facilita a independência física dos dados.

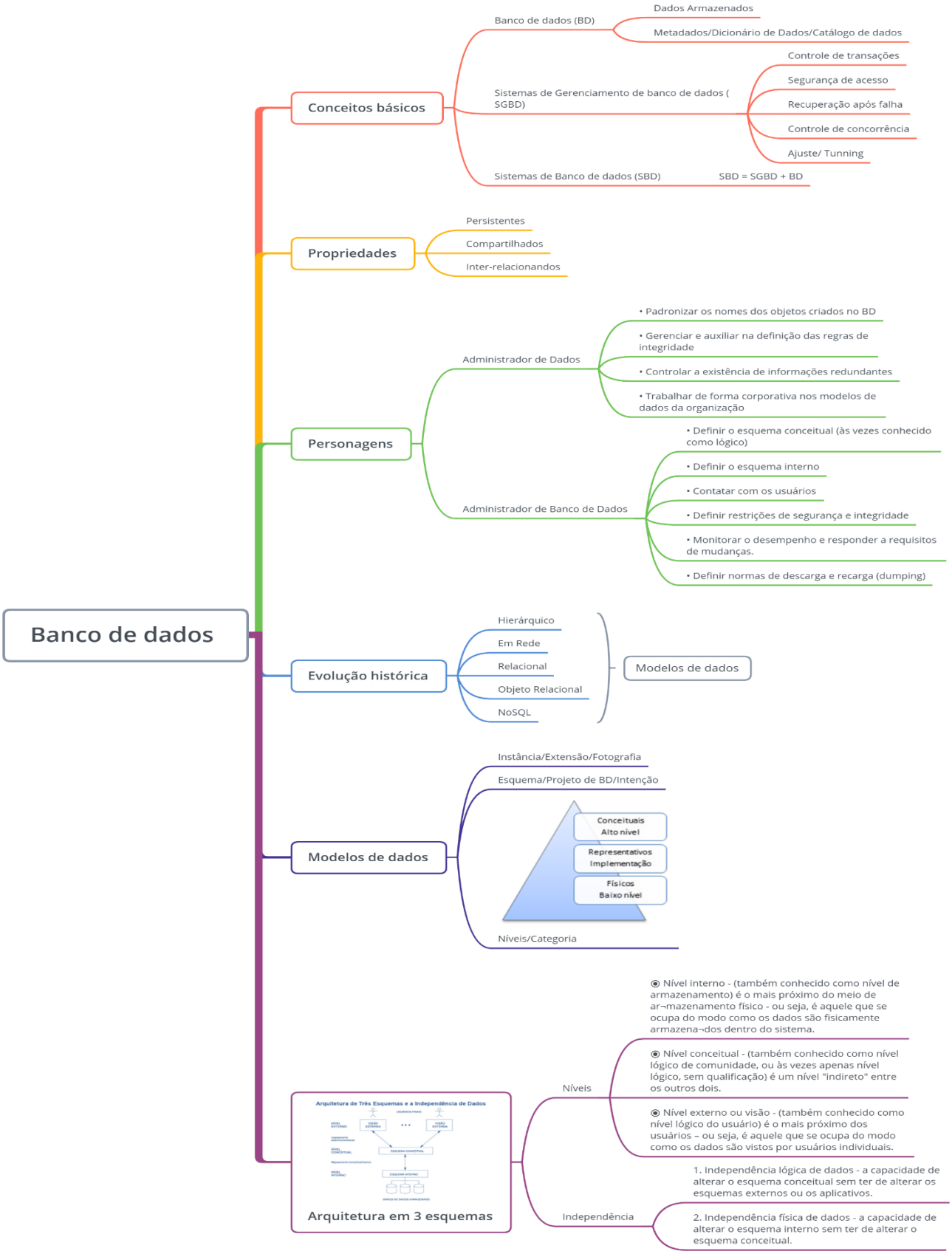
Vejamos então as listas de tópicos e a figura com o relacionamento entre elas:

Modelo de dados: conceitual (alto nível), representativos (lógico ou de implementação) e físicos
Níveis de abstração de dados: nível de view, nível lógico e nível físico.
Arquitetura 3 esquemas: visão externa, esquema conceitual e esquema interno.
Arquitetura 3 esquemas (níveis): nível externo, nível conceitual e nível interno.
Independência de dados: lógica e física.

ESQUEMATIZANDO



MAPA MENTAL


Banco de dados
Conceitos básicos
Banco de dados (BD)
Dados Armazenados
Metadados/Dicionário de Dados/Catálogo de dados
Sistemas de Gerenciamento de banco de dados (SGBD)
Controle de transações
Segurança de acesso
Recuperação após falha
Controle de concorrência
Ajuste/ Tunning
Sistemas de Banco de dados (SBD)
SBD = SGBD + BD
Propriedades
Persistentes
Compartilhados
Inter-relacionandos
Personagens
Administrador de Dados
• Padronizar os nomes dos objetos criados no BD
• Gerenciar e auxiliar na definição das regras de integridade
• Controlar a existência de informações redundantes
• Trabalhar de forma corporativa nos modelos de dados da organização
Administrador de Banco de Dados
• Definir o esquema conceitual (às vezes conhecido como lógico)
• Definir o esquema interno
• Contatar com os usuários
• Definir restrições de segurança e integridade
• Monitorar o desempenho e responder a requisitos de mudanças.
• Definir normas de descarga e recarga (dumping)
Evolução histórica
Hierárquico
Em Rede
Relacional
Objeto Relacional
NoSQL
Modelos de dados
Modelos de dados
Instância/Extensão/Fotografia
Esquema/Projeto de BD/Intenção
Conceituais Alto nível
Representativos Implementação
Físicos Baixo nível
Níveis/Categoria
Arquitetura em 3 esquemas
Níveis
◉ Nível interno - (também conhecido como nível de armazenamento) é o mais próximo do meio de ar¬mazenamento físico - ou seja, é aquele que se ocupa do modo como os dados são fisicamente armazena¬dos dentro do sistema.
◉ Nível conceitual - (também conhecido como nível lógico de comunidade, ou às vezes apenas nível lógico, sem qualificação) é um nível "indireto" entre os outros dois.
◉ Nível externo ou visão - (também conhecido como nível lógico do usuário) é o mais próximo dos usuários – ou seja, é aquele que se ocupa do modo como os dados são vistos por usuários individuais.
Independência
1. Independência lógica de dados - a capacidade de alterar o esquema conceitual sem ter de alterar os esquemas externos ou os aplicativos.
2. Independência física de dados - a capacidade de alterar o esquema interno sem ter de alterar o esquema conceitual.

CONSIDERAÇÕES FINAIS

Prezados Alunos,

Chegamos, pois, ao final da aula demonstrativa do curso de Banco de Dados! As videoaulas referentes ao conteúdo visto até aqui já estão gravadas e devidamente disponibilizadas na área do aluno. Espero que gostem!

A continuação deste assunto encontra-se na próxima aula. Espero reencontrar você como um aluno efetivo.

Forte abraço, bons estudos e até breve! Que Deus abençoe!

Thiago Cavalcanti

REFERÊNCIAS

Fiz uma lista com alguns livros que são referências do conteúdo apresentado na aula caso você queria se aprofundar um pouco.

1. Fundamentals of Database Systems - Ramez Elmasri, Sham Navathe Addison-Wesley, 2011 - Computers - 1172 pages
2. Introdução a sistemas de bancos de dados - By C. J. Date - Elsevier Brasil, 2004 - 865 pages
3. Sistema de Banco de Dados - Abraham Silberschatz, Henry F. Korth, S. Sudarshan - Editora: ELSEVIER BRASIL

THIAGO CAVALCANTI
PROFESSOR

ESSA LEI TODO MUNDO CONHECE:
PIRATARIA É CRIME.

Mas é sempre bom revisar o porquê e como você pode ser prejudicado com essa prática.

1. 

Professor investe seu tempo para elaborar os cursos e o site os coloca à venda.

2. 

Pirata divulga ilicitamente (grupos de rateio), utilizando-se do anonimato, nomes falsos ou laranjas (geralmente o pirata se anuncia como formador de "grupos solidários" de rateio que não visam lucro).

3. 

Pirata cria alunos fake praticando falsidade ideológica, comprando cursos do site em nome de pessoas aleatórias (usando nome, CPF, endereço e telefone de terceiros sem autorização).

4. 

Pirata compra, muitas vezes, clonando cartões de crédito (por vezes o sistema anti-fraude não consegue identificar o golpe a tempo).

5. 

Pirata fere os Termos de Uso, adultera as aulas e retira a identificação dos arquivos PDF (justamente porque a atividade é ilegal e ele não quer que seus fakes sejam identificados).

6. 

Pirata revende as aulas protegidas por direitos autorais, praticando concorrência desleal e em flagrante desrespeito à Lei de Direitos Autorais (Lei 9.610/98).

7. 

Concurseiro(a) desinformado participa de rateio, achando que nada disso está acontecendo e esperando se tornar servidor público para exigir o cumprimento das leis.

8. 

O professor que elaborou o curso não ganha nada, o site não recebe nada, e a pessoa que praticou todos os ilícitos anteriores (pirata) fica com o lucro.



Deixando de lado esse mar de sujeira, aproveitamos para agradecer a todos que adquirem os cursos honestamente e permitem que o site continue existindo.