

Lista de Revisão - Capítulo 6

Resumo

1 Como um banco de dados relacional organiza os dados?

O banco de dados relacional é, atualmente, o método mais usado para organizar e manter dados em sistemas de informação. Ele organiza-os em tabelas bidimensionais, com linhas e colunas chamadas relações. Cada tabela contém dados sobre uma entidade e seus atributos. Cada linha representa um registro e cada coluna, um atributo ou campo. Cada tabela também contém um campo-chave, que confere identificação exclusiva a cada registro para posterior recuperação ou manipulação. O diagrama entidade/relacionamento descreve graficamente a relação entre as entidades (tabelas) em um banco de dados relacional. O processo de fragmentar agrupamentos complexos de dados e simplificá-los, a fim de minimizar a redundância e as relações muitos para muitos inadequadas, é denominado normalização. Bancos de dados não relacionais estão se tornando populares para gerenciar os tipos de dados que não podem ser facilmente manipulados pelo modelo de dados relacional.

2 Quais os princípios de um sistema de gestão de banco de dados?

Um sistema de gestão de banco de dados (DBMS) é um software que permite a centralização e a gestão de dados, de maneira que as empresas disponham de uma única fonte consistente para todas as necessidades de informações. Um único banco de dados atende a múltiplas aplicações. O DBMS separa as visões lógica e física dos dados, a fim de que o usuário não precise se preocupar com a localização física deles. Entre os principais recursos de um DBMS estão um recurso de definição, outro de dicionário de dados e uma linguagem de manipulação dessas informações.

3 Quais as principais ferramentas e tecnologias para extrair informações de bancos de dados e, assim, melhorar o desempenho da empresa e a tomada de decisão?

A tecnologia contemporânea de gestão de dados dispõe de um conjunto de ferramentas para extrair informações úteis de todos os diversos tipos de dados utilizados pelas empresas hoje, incluindo big data semiestruturado e não estruturado em grandes quantidades. Esses recursos incluem data warehouses e data marts, Hadoop, computação em memória e plataformas analíticas. O processamento analítico on-line (Olap) representa os relacionamentos entre os dados como uma estrutura multidimensional que pode ser visualizada como cubos de dados uns dentro dos outros. O data mining analisa grandes quantidades de dados, incluindo o conteúdo dos data warehouses, a fim de encontrar padrões e regras utilizados para prever o comportamento futuro e orientar a tomada de decisão. Ferramentas de mineração de texto ajudam as empresas a analisar grandes conjuntos de dados formados por textos não estruturados. Ferramentas de mineração na Web concentram-se na análise de padrões e informações úteis na World Wide Web a partir da avaliação da estrutura dos sites, das atividades dos usuários na Web e dos conteúdos das páginas. Bancos de dados convencionais podem ser conectados à Web ou a uma interface dela para facilitar o acesso do usuário aos dados internos da organização.

4 Qual o papel da política de informação e da gestão de dados na gestão dos recursos de dados organizacionais? Desenvolver um ambiente de banco de dados exige políticas e procedimentos para gerenciar os dados organizacionais, assim como um bom modelo de dados e uma boa tecnologia de banco de dados. Uma política de informação formal regula a manutenção, a distribuição e o uso da informação na organização. Em grandes corporações, uma função normal da administração de dados é ser responsável pela política de informação, assim como pelo planejamento, desenvolvimento do dicionário e monitoração do uso da empresa.

5 Por que é importante para a empresa garantir a qualidade dos dados?

Dados inexatos, incompletos ou inconsistentes criam sérios problemas financeiros e operacionais para as empresas, porque podem conduzi-las a decisões errôneas sobre as ações que devem tomar.

Assegurar a qualidade dos dados envolve o uso de padrões válidos para toda a empresa, bancos projetados para minimizar a inconsistência e a redundância, auditorias de qualidade, e software de limpeza e padronização.

Exercícios

- Defina e explique os conceitos de entidades, atributos e campos-chave.
- Defina banco de dados relacional e explique como ele organiza e armazena as informações.
- Explique o papel dos diagramas entidade/relacionamento e da normalização no projeto de banco de dados.
- Defina sistema de gestão de banco de dados não relacional (NoSQL) e explique como ele difere de um DBMS relacional.
- Defina sistema de gestão de banco de dados (DBMS), descreva como ele funciona e explique como ele beneficia as organizações.
- Defina e descreva as três operações de um sistema de gestão de banco de dados relacional.
- Nomeie e descreva os três principais recursos de um DBMS.
- Defina big data e descreva as tecnologias para a gestão e análise de big data.
- Descreva os recursos de processamento analítico on-line (Olap).
- Defina data mining e descreva os tipos de informação que podem ser obtidos a partir do data mining.
- Defina política de informação e administração de dados e explique como elas ajudam as empresas a gerenciar suas informações.
- Liste e descreva os problemas mais comuns de qualidade de dados.
- Descreva as ferramentas e técnicas mais importantes para garantir a qualidade dos dados.