**UNIVERSIDADE FEDERAL DE SERGIPE**

DEPARTAMENTO DE SISTEMAS DE INFORMAÇÃO

Sistemas de Apoio à Decisão

Professor: André Vinicius Rodrigues Passos Nascimento

Estudo Dirigido I

1. Na década de 80, programas de extração tornaram-se extremamente comuns para a tarefa de extrair dados do ambiente OLTP. A proliferação desses programas deu origem ao que chamamos hoje de *Ambiente de Desenvolvimento Espontâneo*. Quais os principais problemas dessa arquitetura para o processo de tomada de decisão?

Resp.: falta de credibilidade dos dados, problemas de produtividade, impossibilidade de transformar dados em informações e falta de dados históricos armazenados.

1. Explique o que é um ambiente de *Data Warehouse*. Apresente, através de um desenho (arquitetura) como o Ambiente de *Data Warehouse* se relaciona com outros ambientes, sistemas e tecnologias em uma organização. Explique o desenho apresentado.

Resp.: É um banco de dados de suporte à decisão mantido separado dos bancos de dados operacionais, o qual suporta o processamento de informações através de uma plataforma sólida de dados históricos e consolidados para análises.

1. Quais as principais características de um *Data Warehouse* segundo Inmon ? Explique cada uma delas.

Resp.:

1. Orientado por assunto: provê uma visão simples e consistente de um assunto através da exclusão de dados não relevantes para o processo de tomada de decisão.
2. Integrado: assegura a consistência de convenção de nomes, codificação de atributos e medidas entre as diferentes fontes de dados.
3. Variante no tempo: todo DW contém um elemento de tempo, seja explícito ou implícito.
4. Não volátil: repositório de dados fisicamente separados do ambiente operacional, as atualizações operacionais não ocorrem no DW, além disso, não requer processamento de transações, recuperação ou controle de concorrência.
5. Christopher Adamson, no livro Star Schema: The Complete Reference, Capítulo 02, detalha três principais arquiteturas no contexto de ambientes de suporte à decisão baseados em *Data Warehouse*. Descreva, em linhas gerais, as principais caraterísticas de cada uma dessas arquiteturas.

Resp.:

1. Fábrica de informações corporativas: estes sistemas alimentam um processo denominado ETL; É um repositório integrado de dados atômicos, integrado a partir de diversos sistemas operacionais, contém uma representação definitiva e consistente das atividades empresariais em um único lugar.
2. Data Warehouse Dimensional: é projetado de acordo com os princípios da modelagem dimensional, isto consiste em uma série de esquemas estrela ou cubo. Além disso, pode ser acessado diretamente por sistemas analíticos.
3. Data Marts Independentes: é um armazenamento de dados analíticos que não foi projetado de uma forma contexto empresarial, é focado exclusivamente em uma área temática. O data mart pode empregar design dimensional , um modelo entidade-relacionamento ou alguma outra forma de design.
4. No contexto de um ambiente de suporte à decisão baseado em *Data Warehouse*, o que significa a sigla ETL ? Explique o que acontece em cada atividade ou fase principal de um processo ETL. (Não precisa detalhar muito)

Resp.: Extrair, transformar e carregar. É o processo que as organizações orientadas a dados usam para coletar dados de várias fontes e reuni-los para dar suporte à descoberta, à geração de relatórios, à analise e à tomada de decisões.

1. Extrair: Durante a extração, o ETL identifica os dados e os copia de suas origens, de forma que possa transportar dados para o armazenamento de dados de destino.
2. Transformar: O ETL valida, autentica, desduplica e/ou agrega os dados de formas que tornam os dados resultantes confiáveis e consultáveis.
3. Carregar: O ETL move os dados transformados para o armazenamento de dados de destino. Esta etapa pode implicar o carregamento inicial de todos os dados de origem ou pode ser o carregamento de alterações incrementais nos dados de origem.