

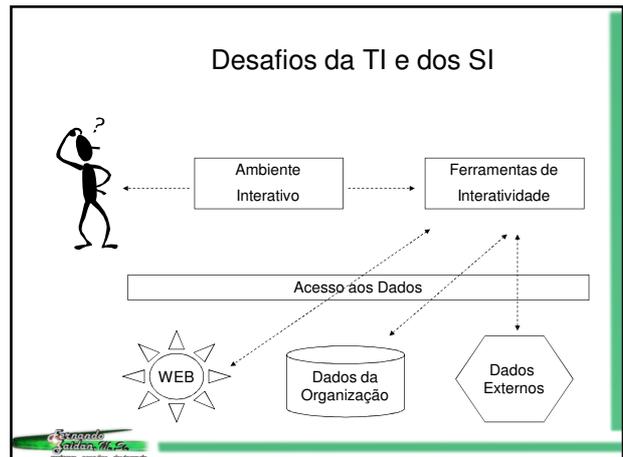
CEFET-MG

PÓS-GRADUAÇÃO LATO SENSU

Curso: Banco de Dados

Disciplina: Data Warehouse e Business Intelligence
Professor: Fernando Zaidan

Unidade 1
2012



Cenários anteriores

- **Sistemas legados;**
- **Banco de Dados cheios e que não agregavam valor;**
- **Sem suporte à decisão;**
- **Sistemas particionados (ilhas de informação);**

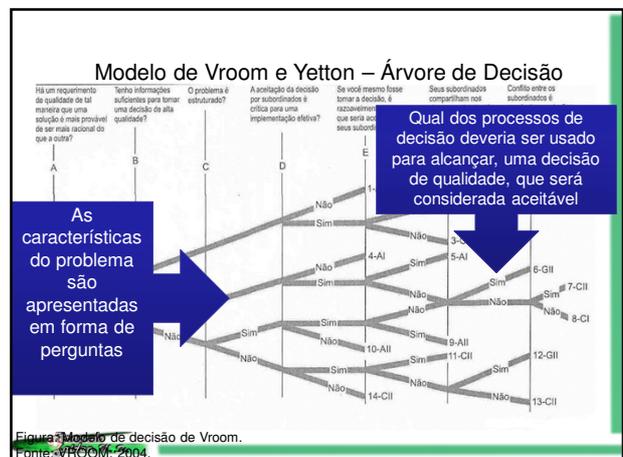
→ **Vale a pena acumular dados sem analisá-los?**

Pode-se dizer que as informações estão armazenadas em três níveis:

- **Estruturado**, aquele que está necessariamente consolidado/sistematizado em algum tipo de suporte;
- **Estruturáveis**, são produzidos internamente, nos diversos setores da organização, porém sem seleção, tratamento e acesso;
- **Não Estruturados**, são produzidos externamente à organização, porém não estão sistematizados, nem filtrados a partir de qualquer tipo de intervenção, bem como estão sem tratamento de qualquer espécie.

Tomada de Decisão

- Decisão é o ato ou efeito de determinar ou resolver questões, solucionar, dar deliberação, dar sentença ou ter preferência;
- Ação explícita neste ato de decidir → ser humano como o elemento que dará a palavra final em uma decisão;
- A tomada de decisão é algo bem mais complexo, sofrendo interferências de variáveis e de fatores internos e externos e o envolvimento de diversos atores que participam do processo organizacional.



Incertezas, riscos e confiança

- Em decisões gerenciais existem componentes de incerteza e desconhecimento que conduzem os indivíduos a fazerem escolhas correndo riscos;
- Ao assumir um risco, o indivíduo confia que sua escolha é a mais adequada e acertada, mesmo que ainda tenha incertezas sobre seu ato;
- Confiança → mecanismo que impulsiona a decisão, oferecendo segurança aos atores responsáveis por ela.

8

Sistemas de Informação Organizacionais e os Níveis

Ilhas de Informação

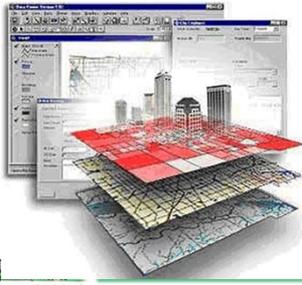
Tipos

STP: Sistemas de Processamento de Transações
 SIG: Sistemas de Informações Gerenciais.
 SATD: Sistemas de Apoio à Tomada de Decisão;
 SIE: Sistemas de Informações Estratégicos;
 SAE: Sistema de Apoio aos Executivos;
 KBS: Sistemas baseados em Conhecimento;
 Sistemas IA: Sistemas Especialistas, Redes Neurais, Robótica;
 Etc.

TIPOS DE SISTEMAS	Sistemas do nível estratégico				
Sistemas de apoio executivo (SAEs)	Previsão quinzenal da tendência de vendas	Plano operacional	Previsão quinzenal de orçamento	Planejamento de lucros	Planejamento de pessoal
Sistemas de informações gerenciais (SIGs)	Sistemas do nível gerencial				
Sistemas de apoio à decisão (SADs)	Gerenciamento de vendas	Controle de estoque	Orçamento anual	Análise de investimento de recursos	Análise de realocação
Sistemas de trabalhadores do conhecimento (STCs)	Sistemas do nível do conhecimento				
Sistemas de automação de escritório	Análise das vendas por região	Programação da produção	Análise de custo	Análise de preços e lucratividade	Análise de custo de contratos
Sistemas de processamento de transações (SPTs)	Sistemas do nível operacional				
	Estações de trabalho de engenharia	Estações de trabalho gráficas	Estações de trabalho administrativas		
	Edição de texto	Treatmento de imagens (digitalização) de documentos	Agendas eletrônicas		
	Sistemas do nível operacional				
	Acompanhamento de pedidos	Controle do maquinário	Negociação de seguros	Folha de pagamento	Remuneração
	Processamento de pedidos	Programação industrial	Contas a pagar	Contas a receber	Treinamento e desenvolvimento
		Controle de movimentação de materiais	Gerenciamento do caixa	Contas a receber	Manutenção do registro de funcionários
	Vendas e marketing	Fabricação	Finanças	Contabilidade	Recursos humanos

GeoProcessamento como Apoio à Decisão

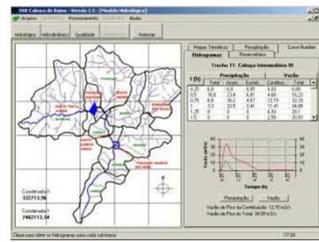
Áreas como cartografia, gestão dos recursos naturais, planejamento urbano, turismo, agricultura, dentre inúmeras outras. Instituições públicas e privadas têm se apercebido de suas vantagens.



Fonte: http://clickgeo.blogspot.com/2009_05_01_archive.html

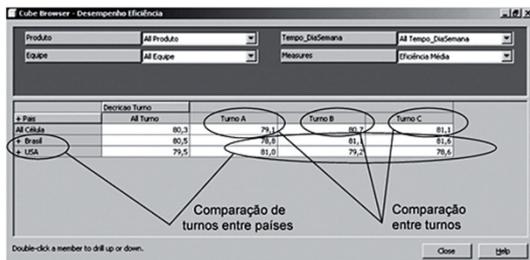
SATP para Gestão de Água Urbana

Software que permite elaborar estratégias preventivas em bacias urbanas, incorporando uma metodologia que permite ações preventivas em regiões de risco.

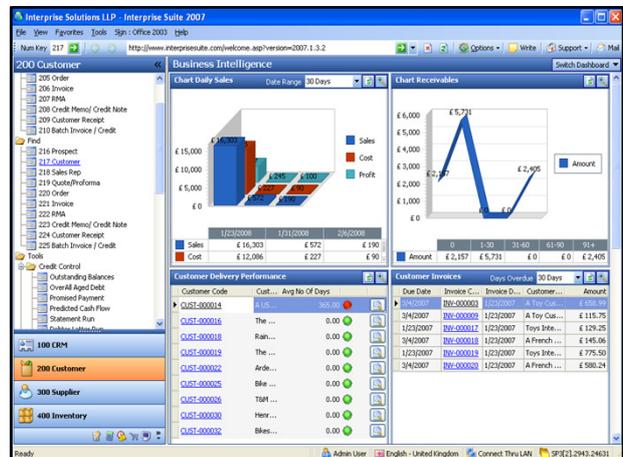
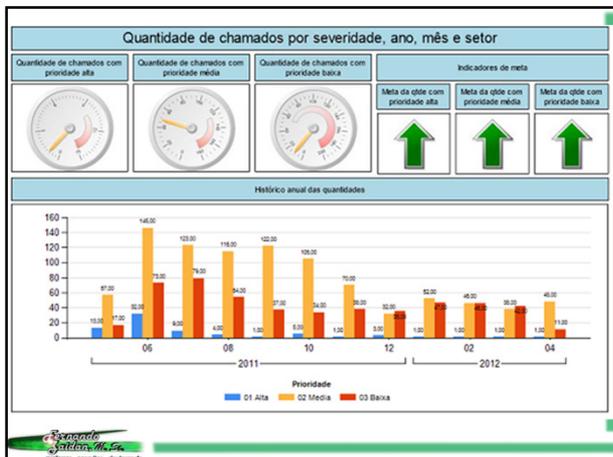


Fonte: <http://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=programa-ajuda-prevenir-cheias-fios-urbanos&id=010125091229>

SATP para Chão de Fábrica



530X2005000100006&script=sci_arttext



Um comparativo

Característica	Sistemas Transacionais(OLTP)	Sistemas Analíticos(OLAP)
Atualizações	Mais frequentes	Menos frequentes
Tipo de Informação	Detalhes	Agrupamento
Quantidade de Dados	Poucos	Muitos
Precisão	Dados atuais	Dados históricos
Complexidade	Baixa	Alta
Consistência	Microscópica	Global
Exemplos	CRM, ERP, Supply Chain	EIS (Executive Information System)
Terminologia	Linhas e Colunas	Dimensões, Medidas e Fatos

Data Warehouse



Introdução

- A teoria de banco de dados foi difundida e está disponível há alguns anos.
- Nesta teoria os **dados eram armazenados com o propósito maior de realizar operações transacionais**, sejam em tempo real, sejam em um intervalo de tempo programado. Mas também eram utilizados, em menor frequência, para a execução de processos analíticos.

Introdução

- Na década de 70 no Instituto de Tecnologia de Massachusetts, MIT, estudos eram realizados com o objetivo de desenvolver uma arquitetura que tivesse **mais eficácia aos objetivos dos sistemas de informação**.
- Surgiu, pela primeira vez, a necessidade de separar sistemas operacionais e aplicações analíticas criando-se o princípio de separar este dois tipos de processamento em **projetos e armazém de dados distintos**.
- Por isto, Hainsten (1999) credita a este período o surgimento do conceito de **data warehouse**.

Introdução

- Para Ballard e Herreman (1998) e Teresko (1999) o **conceito surgiu no início da década de 80** quando os sistemas gerenciadores de banco de dados (SGBD) apareceram como facilitadores da computação de dados, e conseqüentemente, **facilitadores a tomada de decisão**.
- Teresko (1999) expõem que **Inmon é o “pai do Data Warehouse”**, isto pelo fato de Inmon ter observado que os dados poderiam ser organizados em um ativo corporativo que o mesmo nomeou de Data Warehouse.

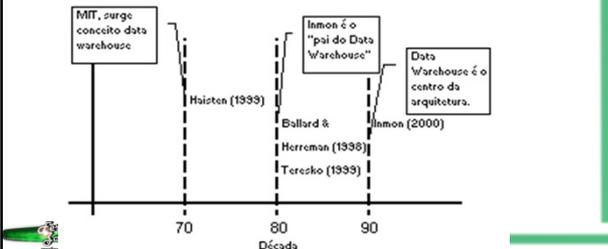
Introdução

- A criação de um **data warehouse é para a integração dos dados de toda a corporação**.
- Proporcionando, ao invés de investir em diversas aplicações de inteligência do negócio, um **foco em ferramentas analíticas e técnicas em um repositório integrado de dados transacionais do negócio**.

Introdução

• O Data Warehouse é o centro da arquitetura para os sistemas de informação na década de 90.

História do Data Warehouse:



Definições DW

“Um Data Warehouse é um repositório de informações integradas, disponíveis para análises e consultas. Dados e informações são extraídos de fontes heterogêneas de onde elas são geradas. A principal vantagem é que se torna muito mais fácil e eficiente rodar consultas sobre esses dados que sobre os sistemas originais.”. Inmon (1997).

“Data Warehouse é uma coleção de dados orientada por assuntos, integrada, variante ao tempo, e não volátil, que tem por objetivo dar suporte aos processos de tomada de decisão.” Inmon (1997).

Termos do DW

• **Orientado por assunto**, a primeira característica do armazém de dados é que ela é orientada em torno dos grandes assuntos da empresa.

• Exemplo: uma instituição financeira na qual existem aplicações como empréstimo, poupança e cartão de crédito. E o data warehouse é organizado sobre os principais assuntos como cliente, vendedor, produto e atividade.

Termos do DW

• **Integrado**, os dados encontrados dentro do armazém de dados serão sempre, sem possibilidades de exceção, íntegros. A integração pode ocorrer ao consistir nomes, consistir medidas de variáveis, consistir estrutura de dados, consistir atributos físicos de dados, e assim por diante.

• **Não volátil**, existem dois tipos de carga que ocorrem no data warehouse, a carga inicial dos dados, e o acesso aos dados. Atualizações não ocorrem no data warehouse. Diferentemente, no ambiente operacional, inclusões, exclusões e alterações ocorrem com frequência nos registros.

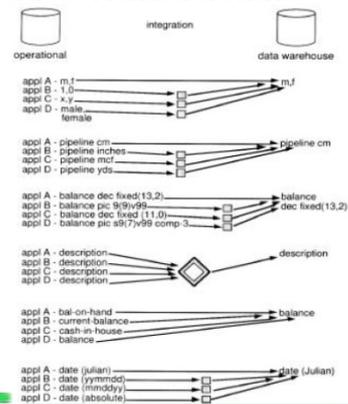
Termos do DW

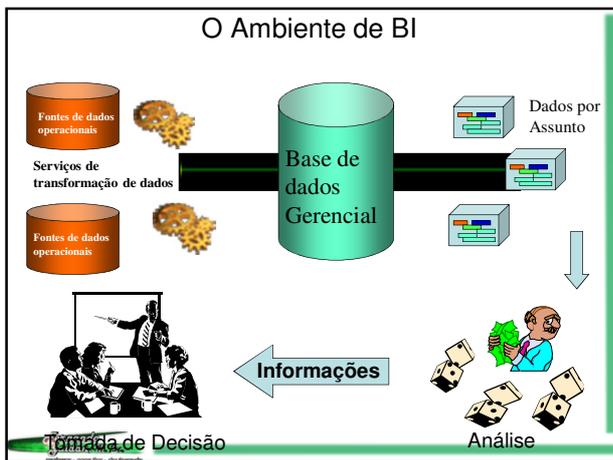
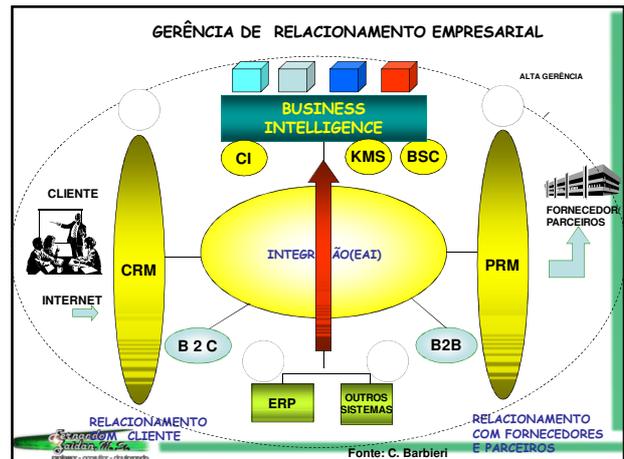
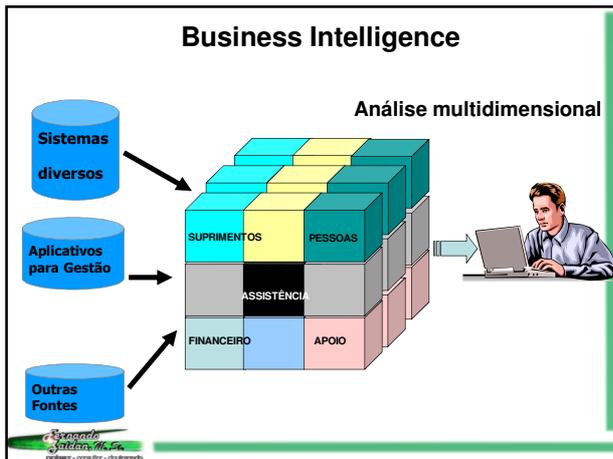
• **Variante no tempo**, os dados no Data Warehouse podem ser apresentados em tempos determinados.

• Em contrapartida nos sistemas operacionais espera-se que os dados sejam a representação exata do tempo no momento do acesso.

• **O tempo representado no horizonte de dados de um Data Warehouse são de vários anos.** A variação no tempo nos dados de um Data Warehouse também pode ser percebida pelo fato de que toda estrutura contém, implicitamente ou explicitamente, um elemento de tempo, seja semana, mês e ano.

Dados Transacionais -> DW





Business Intelligence

- Conjunto de ferramentas e técnicas que objetivam dar suporte à tomada de decisões;
- BI extrai grandes quantidades de dados e os analisa de forma eficiente, para que sejam transformados em informações e possam ser compartilhados de forma clara e transparente, auxiliando as tomadas de decisões e melhorando o desempenho gerencial e operacional.

Fonte: Associação Saitens M. S.

Benefícios do BI

- Alinhar projetos de tecnologia com as metas estabelecidas pelas empresas na busca do máximo retorno do investimento;
- Ampliar a compreensão das tendências dos negócios, propiciando melhor consistência no momento de decisão de estratégias e ações;
- Facilitar a identificação de riscos e gerar segurança para migração de estratégias, criando maior efetividade nas implementações dos projetos;
- Permitir um planejamento corporativo mais amplo, substituindo soluções de menor alcance por resultados integrados pela informação consistente;

Fonte: Associação Saitens M. S.

Benefícios do BI

- Gerar, facilitar o acesso e distribuir informação de modo mais abrangente para obter envolvimento de todos os níveis da empresa;
- Ligar e consolidar dados de diferentes sistemas de modo a oferecer um visão comum do desempenho da empresa;
- Automatizar tarefas eliminando os erros ao colocar as pessoas no fim dos processos;
- Oferecer dados estratégicos para análise com um mínimo de atraso em relação a uma transação ou evento dentro da empresa;

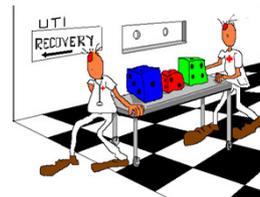
Fonte: Associação Saitens M. S.

BI Aplicativos Operacionais x BI

Aplicativos Operacionais	Business Intelligence
Visão do atual e do real	Visão histórica e de tendência
Solução para requisitos conhecidos	Permitir a identificação de fatos desconhecidos
Abrangência restrita	Abrangência ampla
Informação produzida por profissionais de informática	Informação produzida pelo próprio usuário
Custo e tempo para obtenção da informação altos	Informação obtida com baixo custo e em tempo real
Informação disponível a poucos usuários	Informação democratizada

Escuela Saldan M. S.
 www.com.br/database

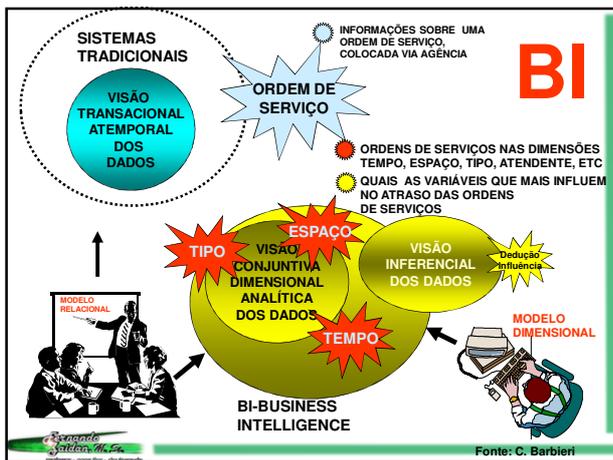
PODER DA INFORMAÇÃO-BI



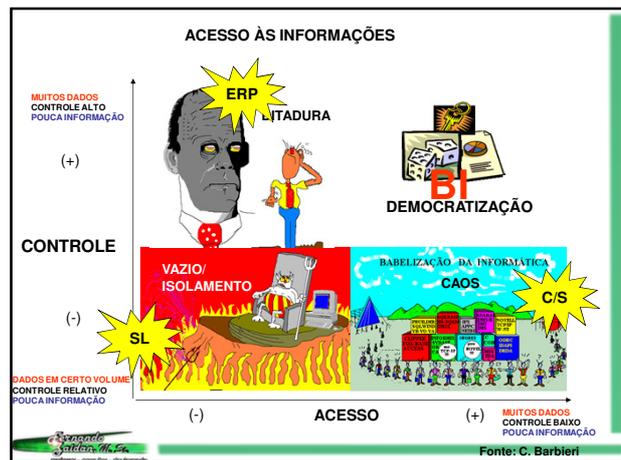
TRATAR BEM DOS DADOS:
 • DADO COMO RECURSO
 • DADOS ENTENDIDOS
 • DADOS ESTRUTURADOS
 • DADOS MODELADOS
 • DADOS SEGMENTADOS
 • DADOS DISPONIBILIZADOS
 • DADOS PROTEGIDOS
INSTRUMENTO DE INTELIGÊNCIA NOS NEGÓCIOS

Escuela Saldan M. S.
 www.com.br/database

Fonte: C. Barbieri



Fonte: C. Barbieri



Fonte: C. Barbieri

BI



- **CONSULTAS PADRÕES**
 - LISTAR TODOS OS CLIENTES QUE TIVERAM UM DECRÉSCIMO DE CONSUMO DE 20% OU + NA CONTA DE ENERGIA
- **ANÁLISES MULTIDIMENSIONAIS**
 - LISTAR TODOS OS CS DO TIPO **residencial** QUE TIVERAM CONSUMO ENTRE 400 E 500 kwh NA REGIÃO **sul**, DURANTE O PERÍODO de **MAIO A JULHO DE 2000**
- **DATA MINING**
 - LISTAR QUAIS CS TEM > PROBABILIDADE DE RESPONDER A CAMPANHA DE REDUÇÃO

Escuela Saldan M. S.
 www.com.br/database

Fonte: C. Barbieri

OLTP E OLAP



- DADOS ORIENTADOS A **TRANSAÇÕES**
 - NORMALIZAÇÃO
- AMBIENTE COM ALTA TAXA DE TRANSAÇÕES
- DADOS SÃO ATUALIZADOS
- DADOS ESTÃO TOTALMENTE UP-TO DATE
- DADOS SÃO ESTRUTURADOS PARA TUDO ACIMA
 - RELACIONAL



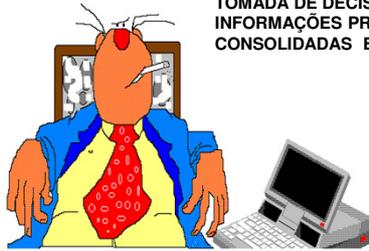
- DADOS ORIENTADOS A **ANÁLISE** DE INFORMAÇÃO
- AMBIENTE DE BAIXA TAXA DE TRANSAÇÕES
- DADOS SÃO CONSULTADOS
- DADOS SÃO "REFRESHED" EM PERÍODOS DE TEMPO
- DADOS SÃO REFORMATADOS PARA TUDO ACIMA
 - DIMENSIONAL

Escuela Saldan M. S.
 www.com.br/database

Fonte: C. Barbieri

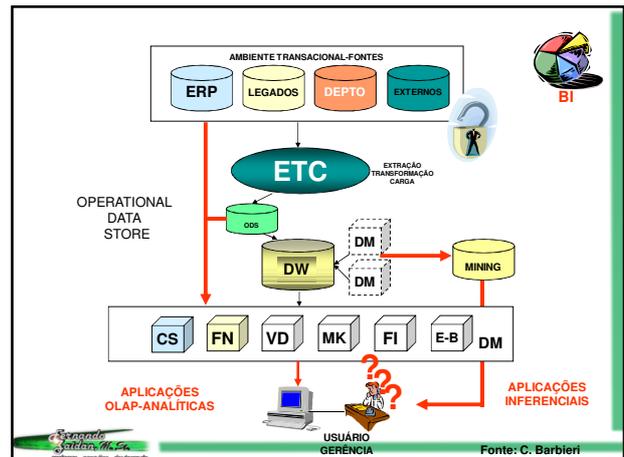
DATA WAREHOUSE

BANCO DE DADOS PARA ANÁLISE E TOMADA DE DECISÕES, BASEADA EM INFORMAÇÕES PREPARADAS, CONSOLIDADAS E SUMARIADAS



DATA MART – Pequeno DW – depósitos que atendem a certas áreas específicas – Custo mais baixo – veículo de testes – tempo menor de implantação

Fonte: C. Barbieri



Fonte: C. Barbieri

BI E DATA WAREHOUSE O QUE NÃO É

- NÃO RESOLVE TODOS OS PROBLEMAS
- **50% DOS PROJETOS DE DW FALHAM POR FALTA CLARA DE OBJETIVOS E GIGANTISMO DE ESCOPO**
- NÃO É UM ÚNICO PROJETO-SÃO VÁRIOS
- NÃO SE IMPLEMENTA NUM ÚNICO RELEASE-SÃO PRODUTOS GRADATIVAMENTE DISPONIBILIZADOS
- NÃO É UM BD COMUM DE DADOS OPERACIONAIS-NOVA FORMA DE ORGANIZAR OS DADOS(INFORMACIONAIS)

Fonte: C. Barbieri

CONCLUSÃO

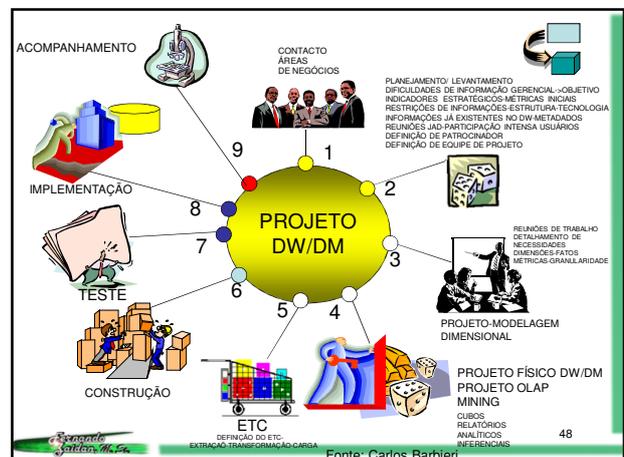
- **BI NÃO É MODISMO**
- **APRIMORAMENTO DA ETERNA BUSCA POR INFORMAÇÕES QUE SEMPRE EXISTIU NA INFORMÁTICA**
- **REFORMATAÇÃO DE CONCEITOS**
 - BD RELACIONAIS => ESTRUTURAS DIMENSIONAIS
 - APLICAÇÕES TRANSACIONAIS => APLICAÇÕES ANALÍTICAS
 - INCLUSÃO DE APLICAÇÕES INFERENCIAIS
- **ENTENDER O CONCEITO DE VALOR DA INFORMAÇÃO, AGORA COM MÉTRICAS PLAUSÍVEIS**
- **DADO=>INFORMAÇÃO PRECISA NO MOMENTO DEMANDADO POR QUEM EFETIVAMENTE NECESSITA**

Fonte: C. Barbieri

Pesquisa BI

- Realizou uma pesquisa em 1997 com 121 participantes de uma conferência sobre Data Warehouse patrocinada pelo Data Warehousing Institute.
- Quando foram solicitadas a responder sobre as motivações de suas companhias para Data Warehouse, 38% responderam que gostariam de ter melhor acesso a informação, 21% ter uma informação melhor e mais precisa, e 20% ter uma única fonte de dados.

Fonte: C. Barbieri



Fonte: Carlos Barbieri

LEVANTAMENTO DE NECESSIDADES



- **DEFINIR AS NECESSIDADES DE INFORMAÇÃO PARA O NEGÓCIO:**
 - **DIFICULDADES-RESTRIÇÕES SUGEREM OBJETIVOS**
 - **INDICADORES, MÉTRICAS, COMPARATIVOS**
 - **SEMPRE OBSERVAR “DIMENSÕES”/COMBINAÇÕES**
 - **MODELO DE DADOS EXISTENTES:**
 - ENTIDADES-RELACIONAMENTOS-ATRIBUTOS- (KEYS-ORIGINAIS-DERIVADOS)
 - ARQUIVOS, DADOS MANUAIS EXISTENTES

Fonte: Carlos Barbieri

49

Erros Comuns a Evitar em Modelagem Dimensional

- Colocar atributos de texto usados para restrições e agrupamento numa tabela de fatos.
- Limitar atributos em dimensões para economizar espaço.
- Ignorar a necessidade de cuidar de mudanças em atributos de dimensões.
- Resolver todos os problemas de desempenho de consultas adicionando mais hardware.
- Usar chaves operacionais ou “inteligentes” para junções de tabelas de dimensão com tabela de fatos.
- Projetar o modelo dimensional baseado em um relatório específico.

Fonte: Marcos André Gonçalves

50

Erros Comuns a Evitar em Projetos de Data Warehouse

- Aceitar a premissa de que os responsáveis pelos sistemas operacionais fontes mais relevantes da organização são muito importantes e ocupados para gastar tempo com a equipe de DW.
- Assegurar para o pessoal de suporte do DW escritórios agradáveis no prédio da TI, que fica próximo dos usuários de negócio, e providenciar um número de telefone de suporte de DW com várias opções de menu.
- Treinar cada usuário em cada característica da ferramenta de acesso a dados na primeira aula de treinamento, adiar o treinamento sobre conteúdo de dados porque a aula usa dados falsos (os dados reais não estarão prontos nos próximos dois meses) e declarar sucesso ao término da primeira aula de treinamento já que o DW foi disponibilizado para os usuários de negócio.

Fonte: Marcos André Gonçalves

51

Erros Comuns a Evitar em Modelagem Dimensional

- Assumir que os usuários de negócio vão naturalmente gravitar em direção a dados robustos e desenvolver suas próprias “killer applications” analíticas.
- Evitar a ilusão sedutora de desenvolvimento iterativo, que é somente uma desculpa para não fazer certo da primeira vez. Antes de implementar o DW, fazer uma análise completa descrevendo todos os possíveis ativos de dados da empresa e todos os usos desejados de informação, e

Fonte: Marcos André Gonçalves

52

Erros Comuns a Evitar em Modelagem Dimensional

- Não encorajar os usuários de negócio a lhe dar feedback contínuo ao longo do ciclo de desenvolvimento sobre novas fontes de dados e métricas chaves de desempenho que eles gostariam de acessar, e não assegurar a inclusão desses requisitos na release em desenvolvimento.
- Não conversar com os usuários de negócio; ao invés disso, confiar em consultores ou especialistas internos para lhe dar interpretação dos requisitos de usuários do DW.

Fonte: Marcos André Gonçalves

53

FCS-EM PROJETOS DE DW

- **DEFINIR UMA ORIENTAÇÃO METODOLÓGICA**
 - DW OU DMART GRADATIVO-DMART ISOLADO
 - RALPH KIMBALL - BILL INMON
- **DEFINIR UMA ARQUITETURA TECNOLÓGICA CONSISTENTE, MODERNA, EVOLUTIVA**
 - INTERFACE WEB É UM DIFERENCIAL
- **DEFINIR UMA METODOLOGIA PRÁTICA, ENXUTA, INTERATIVA, REFINAMENTOS SUCESSIVOS E PRODUTOS ENTREGUES EM PRAZO RAZOÁVEL**
- **DEFINIR EQUIPE, COM PRESERVAÇÃO DO CONHECIMENTO APÓS O TÉRMINO DO PROJETO**
 - CUIDADO COM TERCEIROS-CONSULTORES
- **OBTER PATROCINADORES FORTES PARA O PROJETO, COM OBJETIVOS DIRETOS NO NEGÓCIO DA EMPRESA**
- **DESENVOLVER UM FORTE ESQUEMA DE DEMONSTRAÇÃO DOS PRODUTOS DESENVOLVIDOS (VENDER BEM)**

Fonte: Carlos Barbieri

Bibliografia

BARBIERI, Carlos. BI - Business Intelligence: Modelagem e tecnologia. Rio de Janeiro, Axcel Books, 2001.

CAMPOS, M. L. Data Ware Housing. UFRJ, 2007.

COME, Gilberto de. **Contribuição ao Estudo da Implementação de Data Warehousing: um caso no setor de telecomunicações** – São Paulo : FEA/USP, 2001. 133 p

FANTAUZZI, F. A. C.; ROCHA, Rogério Morais. **Diretório de Softwares para Inteligência Competitiva** Monografia apresentada ao Departamento de Ciência da informação como requisito para a conclusão do curso de especialização em Gestão Estratégica da Informação da Universidade Federal de Minas Gerais - UFMG, Belo Horizonte, ano de 2006.

FARIA, João Marcos Bonadio de. **Artefatos da Semiótica Organizacional na Elicitação de Requisitos para Soluções de Data Warehouse** Trabalho final (mestrado profissional) - Universidade Estadual de Campinas, Instituto de Computação, fevereiro de 2006.

Bibliografia

INMON, William. **What is Data Warehouse ?** UNJobs, acessado em 19 de abril de 2009, disponível em < <http://unjobs.org/authors/w.-h.-inmon>>

SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S. Sistema de banco de dados. Tradução da 5ª Edição. São Paulo: Campus, 2006.

TERESKO, John. **Information Rich, Knowledge Poor ?** IndustryWeek.com, acessado em 19 de abril de 2009, disponível em < <http://www.industryweek.com/PrintArticle.aspx?ArticleID=245> >



Obrigado e bom trabalho,

Zaidan

www.fernandozaidan.com.br
fhzaidan@gmail.com

"Aí está o mérito do êxito de meus projetos: sempre fui muito exigente e rigoroso com procedimentos que aparentemente não faziam muito sentido na época. Mais tarde viu-se que esse rigor fez a diferença entre afundar ou não, concluir ou não um projeto".

Amyr Klunk

