

USP

Campus de São Carlos

**UMA ANÁLISE DA APLICABILIDADE DO DATA
WAREHOUSE NO COMÉRCIO ELETRÔNICO,
ENFATIZANDO O CRM ANALÍTICO**

Fabricia Vancim Frachone Neves

Orientador: Prof. Dr. Edson Walmir Cazarini

UNIVERSIDADE DE SÃO PAULO



ESCOLA DE ENGENHARIA DE SÃO CARLOS

Uma análise da aplicabilidade do Data
Warehouse no comércio eletrônico,
ênfatizando o CRM analítico.

Serviço de Pós-Graduação EESC/USP
EXEMPLAR REVISADO
Data de entrada no Serviço: 06 / 12 / 01
Ass.: <i>Guilherme</i>

Fabricia Vancim Frachone Neves



Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Produção.

ORIENTADOR: Prof. Dr. Edson Walmir Cazarini

DEDALUS - Acervo - EESC



31100036851

São Carlos

2001

Class.	TESE-EESC
Cutt.	13843
Tombo	T0005/02

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

N518a Neves, Fabricia Vancim Frachone
Uma análise da aplicabilidade do data warehouse no
comércio eletrônico, enfatizando o CRM analítico /
Fabricia Vancim Frachone Neves. -- São Carlos, 2001.

Dissertação (Mestrado) -- Escola de Engenharia de
São Carlos-Universidade de São Paulo, 2001.

Área: Engenharia de Produção.

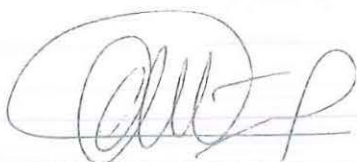
Orientador: Prof. Dr. Edson Walimir Cazarini.

1. Comércio eletrônico. 2. CRM. 3. Data warehouse.
4. OLAP. 5. Data mining. I. Título.

FOLHA DE JULGAMENTO

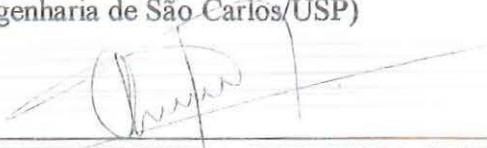
Candidata: Bacharela **FABRICIA VANCIM FRACHONE NEVES**

Dissertação defendida e julgada em 22-11-2001 perante a Comissão Julgadora:



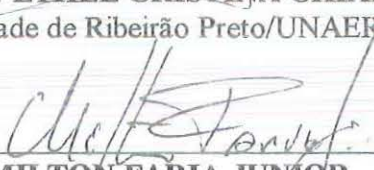
Prof. Dr. **EDSON WALMIR CAZARINI (Orientador)**
(Escola de Engenharia de São Carlos/USP)

Aprovada



Profa. Dra. **ETHEL CRISTINA CHIARI DA SILVA**
(Universidade de Ribeirão Preto/UNAERP)

Aprovada

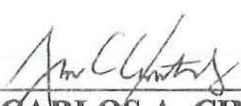


Prof. Dr. **MILTON FARIA JUNIOR**
(Universidade de Ribeirão Preto/UNAERP)

Aprovada



Prof. Doutor **EDMUNDO ESCRIVÃO FILHO**
Coordenador do Programa de Pós-Graduação
em Engenharia de Produção



JOSÉ CARLOS A. CINTRA
Presidente da Comissão de Pós-Graduação da EESC

Dedicatória

Especialmente aos meus pais, pessoas a quem amo muito, que estiveram sempre ao meu lado, amparando-me nas horas de sofrimento, incentivando-me nas horas de desespero e, acima de tudo, proporcionando condições para que eu pudesse lograr mais esta conquista. A minha vitória é a deles também. Obrigada!

Ao meu irmão Fernando e às minhas irmãs Fabiana e Flávia que sempre acreditaram em mim, me apoiando, incentivando e dando forças para continuar o meu caminho.

“Aspiras a vencer e vencerás, mas lembra-te de que vencer sem abrir os caminhos da vitória para os outros é avançar para o tédio da inutilidade sob o frio da solidão”.

Emmanuel

Agradecimentos

À Deus, pela saúde, disposição e vontade de trabalhar que tenho todos os dias.

Ao professor Dr. Edson Walmir Cazarini, pelo apoio, orientação, amizade e respeito as minhas idéias.

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior), pelo apoio financeiro, fundamental para conclusão deste trabalho.

Aos funcionários do Departamento de Engenharia de Produção da Escola de Engenharia de São Carlos da USP.

Ao professor Dr. Edmundo Escrivão Filho e à professora Dra. Ethel Cristina Chiari da Silva, pela importante contribuição na Banca de Qualificação, pelos conhecimentos que me foram passados, pela dedicação que fora dada e pelas orientações que me foram passadas que seguirão comigo tanto em minha vida pessoal quanto profissional.

Ao professor Dr. João do Espírito Santo Batista (ICMC), pelo intenso apoio que tenho recebido desde a graduação até os dias atuais

À bibliotecária Eleninha, da Biblioteca Central da Escola de Engenharia de São Carlos EESC-USP, pela grande ajuda na correção das referências bibliográficas.

Aos funcionários da biblioteca da Universidade de Ribeirão Preto – UNAERP, pelo auxílio às pesquisas.

À professora Sandra Ântico, coordenadora do LIAPE - UNAERP, pela disponibilização dos equipamentos para utilização e incentivo na pesquisa.

Aos monitores do LIAPE, pelo auxílio e disposição em me ajudar.

Aos professores Paulo Sprovieri, Osvaldo Lázaro, Osvaldo Takai, Fernando Cozin e Milton Faria da Universidade de Ribeirão Preto, pelo incentivo científico.

Aos amigos João Batista, Neli Silveira, Sandra Ântico, Lauro Osiro, Ariadna Garrafoli, Luciana Castro, Sílvia de Pádua, Bispo, Etienne Cardoso, Rodrigo Cambiaghi e Maria Emília, pela convivência e troca de experiências que com certeza, ajudaram-me a evoluir nesta pesquisa.

Finalmente ao meus pais a minha eterna gratidão pelo apoio e pelos ensinamentos morais que me foram passados e essenciais durante esta pesquisa.

SUMÁRIO

LISTA DE FIGURAS	III
LISTA DE TABELAS	IV
LISTA DE ABREVIATURAS	V
RESUMO	VII
ABSTRACT	VIII
1. INTRODUÇÃO	1
1.1 CONTEXTO E JUSTIFICATIVA DO TRABALHO	1
1.2 OBJETIVO DO TRABALHO	4
1.3 LIMITAÇÕES DO TRABALHO	4
1.4 MÉTODO DE PESQUISA	5
1.4.1 PERGUNTAS DE PESQUISA	5
1.4.2 ESCOLHA DO MÉTODO DE PESQUISA	5
1.4.3 FASES DA PESQUISA	6
1.4.4 ETAPAS DO TRABALHO	7
1.5 SELEÇÃO DA FERRAMENTA DATA WAREHOUSE	14
1.6 ORGANIZAÇÃO DO TEXTO	15
2. COMÉRCIO ELETRÔNICO	16
2.1 CONSIDERAÇÕES INICIAIS	16
2.2 OS PRINCIPAIS CONCEITOS DE COMÉRCIO ELETRÔNICO	16
2.3 FORMAS DE COMÉRCIO ELETRÔNICO: BUSINESS-TO-BUSINESS E BUSINESS-TO-CONSUMER	18
2.4 ASPECTOS DO COMÉRCIO ELETRÔNICO	19
2.5 SEGURANÇA NO COMÉRCIO ELETRÔNICO	23
2.6 PROBLEMAS COM A SEGURANÇA	24
2.7 PRINCIPAIS MEDIDAS DE SEGURANÇA	25
2.8 RELACIONAMENTO COM O CLIENTE NO COMÉRCIO ELETRÔNICO	29
2.9 CONSIDERAÇÕES FINAIS DO CAPÍTULO	31
3. CRM – CUSTOMER RELATIONSHIP MANAGEMENT	33

3.1	CONSIDERAÇÕES INICIAIS	33
3.2	RELACIONAMENTO COM OS CLIENTES ANTES DO CRM	33
3.3	CRM COMO ESTRATÉGIA DE NEGÓCIO	37
3.4	GERENCIAMENTO DO RELACIONAMENTO COM O CLIENTE NO COMÉRCIO ELETRÔNICO	44
3.5	CONSIDERAÇÕES FINAIS DO CAPÍTULO	48
4.	DATA WAREHOUSE: PRINCIPAIS ASPECTOS	50
<hr/>		
4.1	CONSIDERAÇÕES INICIAIS	50
4.2	DATA WAREHOUSE - O SURGIMENTO DO CONCEITO	50
4.3	PRINCIPAIS FERRAMENTAS UTILIZADAS EM UM DATA WAREHOUSE	57
4.4	CARACTERÍSTICAS DO DATA WAREHOUSE	59
4.5	METADADOS	61
4.6	ARQUITETURA DO DATA WAREHOUSE	64
4.7	MODELAGEM DOS DADOS	66
4.7.1	MODELO DIMENSIONAL	67
4.7.1.1	MODELO ESTRELA	68
4.7.1.2	VANTAGENS DO MODELO ESTRELA	70
4.7.2	COMPARAÇÃO ENTRE O MODELO RELACIONAL E O MODELO DIMENSIONAL	71
4.8	DATA MART	74
4.9	A IMPORTANCIA DO DATA WAREHOUSE NO COMÉRCIO ELETRÔNICO	76
4.10	EXTRAINDO AS INFORMAÇÕES DE UM DATA WAREHOUSE	80
4.10.1	DATA MINING	81
4.10.2	OLAP - PROCESSAMENTO ANALÍTICO ON-LINE	86
4.10	CONSIDERAÇÕES FINAIS DO CAPÍTULO	90
5.	MODELO INTEGRADO	93
<hr/>		
5.1	CONSIDERAÇÕES INICIAIS	93
5.2	INTRODUÇÃO	93
5.3	DIMENSÕES DA SEQÜÊNCIA DE CLIQUES	94
5.3.1	DIMENSÃO DATA DO CALENDÁRIO	95
5.3.2	DIMENSÃO DO HORÁRIO DO DIA	97
5.3.3	DIMENSÃO CLIENTE	98
5.3.4	DIMENSÃO DA PÁGINA	102
5.3.5	DIMENSÃO DO EVENTO	103
5.3.6	DIMENSÃO DA SESSÃO	104
5.3.7	DIMENSÃO DE REFERÊNCIA	105
5.3.8	DIMENSÃO DO PRODUTO (OU SERVIÇO)	106
5.3.9	DIMENSÃO CAUSAL	108
5.3.10	DIMENSÃO DA ENTIDADE DE NEGÓCIO	109
5.4	CHAVES DE MONITORAMENTO DE SEQÜÊNCIA DE CLIQUES	110
5.5	DATA MART DE SEQÜÊNCIA DE CLIQUES	112
5.6	GERENCIANDO O DATA WEBHOUSE	118
5.7	CONSIDERAÇÕES FINAIS DO CAPÍTULO	130
6.	CONCLUSÃO	132
<hr/>		
7.	REFERÊNCIAS BIBLIOGRÁFICAS	136
<hr/>		

Lista de Figuras

<i>Figura 1: Fases da Pesquisa e etapas do trabalho</i>	9
<i>Figura 2: Fluxo de Criptografia</i>	26
<i>Figura 3: Exemplo de um Firewall</i>	28
<i>Figura 4: Ciclo de como servir melhor os clientes</i>	30
<i>Figura 8: Valor do CRM</i>	47
<i>Figura 9: Processo CRM</i>	48
<i>Figura 10: Diferença entre o Ambiente Operacional e Data Warehouse</i>	53
<i>Figura 11: Exemplo de Integração no Data Warehouse</i>	54
<i>Figura 12: Exemplo de não volatilidade no Data Warehouse</i>	55
<i>Figura 13: Exemplo de Variação em relação ao tempo no Data Warehouse</i>	55
<i>Figura 14: Arquitetura Genérica do Data Warehouse</i>	64
<i>Figura 15: Modelo Estrela</i>	69
<i>Figura 16: Faces de um cubo</i>	72
<i>Figura 17: Dimensão Tempo</i>	73
<i>Figura 18: Dimensão Produto</i>	73
<i>Figura 19: Dimensão Cliente</i>	73
<i>Figura 20: Informações Cruzadas</i>	74
<i>Figura 21: Datamart</i>	75
<i>Figura 22: O relacionamento único entre o ISP e o usuário conectado</i>	77
<i>Figura 23: Dimensões e Fato de um DM – FC</i>	78
<i>Figura 24: Exemplo de uma Rede neural</i>	85
<i>Figura 25: Exemplo de um OLAP</i>	87
<i>Figura 26: Uma tabela de fatos de seqüência de cliques</i>	113
<i>Figura 27: Papéis e Organizações de Webhouse</i>	121

Lista de Tabelas

<i>Tabela 1: Estrutura de Análise do Comércio Eletrônico</i>	20
<i>Tabela 2: Estágios no ciclo de vida do consumidor na Internet</i>	22
<i>Tabela 3: Mudanças no papel do Marketing</i>	35
<i>Tabela 4: Ferramentas do Data Warehouse</i>	59
<i>Tabela 5: Comparação entre Banco de Dados Operacional e Data Warehouse</i>	61
<i>Tabela 6: OLAP versus OLTP</i>	88
<i>Tabela 7: Dimensão de Data do Calendário</i>	96
<i>Tabela 8: Dimensão de Horário do Dia</i>	98
<i>Tabela 9: Dimensão do Cliente – grupo 1</i>	99
<i>Tabela 10: Dimensão do Cliente – grupo 2</i>	99
<i>Tabela 11: Dimensão do Cliente – grupo 3</i>	101
<i>Tabela 12: Dimensão do Cliente – grupo 4</i>	102
<i>Tabela 13: Dimensão da Página</i>	103
<i>Tabela 14: Dimensão do Evento</i>	103
<i>Tabela 15: Dimensão da Sessão</i>	105
<i>Tabela 16: Dimensão da Referência</i>	106
<i>Tabela 17: Dimensão do Produto</i>	107
<i>Tabela 18: Dimensão do Serviço</i>	107
<i>Tabela 19: Dimensão Causal</i>	108
<i>Tabela 20: Dimensão da Entidade do Negócio</i>	110
<i>Tabela 21: Chaves de Monitoramento</i>	112

Lista de Abreviaturas

B2B	Business to Business
AS	Administrados de Sistemas
B2C	Business to Consumer
CMP	Clientes de Maior Potencial
CMV	Clientes de Maior Valor
CRM	Customer Relationship Management
DBA	Data Base Administrator
DBM	Data Base Marketing
DMFC	Data Mart de Fluxo de Clientes
DW	Data Warehouse
EIS	Executive Informatio Systems
E-R	Entidade-Relacionamento
ETL	Extrair-Transformar-Carregar
HTML	Hypertext Markup Language
ID	Identificação de usuário
ISP	Internet Service Provider
KDD	Knowledge Discovery in Data Base
OLAP	On line Analytical Processing
OLTP	On line Transaction Processing
PKI	Public Key Infastructure
SAD	Sistemas de Apoio à Decisão
SET	Secure Eletronic Transation
SKU	Stock Keeping Unit
SQL	Structured Query Language

TI	Tecnologia de Informação
URL	Universal Resource Locator
VPN	Virtual Private Network
XML	Extended Markup Language

Resumo

NEVES, F.V.F (2001). *Uma análise da aplicabilidade do Data Warehouse no comércio eletrônico, enfatizando o CRM analítico*. São Carlos. 144p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

Atualmente, um dos grandes desafios enfrentados pelas empresas que operam no comércio eletrônico é de garantir a fidelidade de seus clientes. Devido a esse fato, as empresas têm se preocupado cada vez mais na utilização de uma ferramenta capaz de integrar os dados distribuídos em banco de dados operacionais, para transformar suas informações em diferencial competitivo. Desta forma, a aplicabilidade da ferramenta Data Warehouse no comércio eletrônico, juntamente com a estratégia do *Customer Relationship Management* (CRM), surge como um elemento central da empresa focada no cliente, pois é capaz de fornecer informações competitivas e essenciais aos responsáveis pelas decisões estratégicas. Assim, através da ferramenta Data Warehouse sustentada pela estratégia de CRM, toda a “seqüência de cliques” realizada pelos clientes do comércio eletrônico, é armazenada historicamente, possibilitando então, através de análises, a identificação de clientes potenciais, buscando a retenção destes com a empresa. Portanto o objetivo do presente trabalho é analisar a aplicabilidade do Data Warehouse no comércio eletrônico, visando o CRM analítico. O trabalho contribui também para mostrar às empresas que a integração entre tecnologia de informação (Data Warehouse) e *marketing* (CRM) em um ambiente *Web* oferece diferentes oportunidades de coleta de dados a respeito dos clientes, promovendo desta forma a personalização de produtos e serviços, garantindo maior lealdade destes para com a empresa e, conseqüentemente, maior lucratividade.

Palavras- chaves: Comércio Eletrônico, CRM, Data Warehouse, OLAP, Data Mining.

Abstract

NEVES, F.V.F (2001). *An analysis of applicability of Data Warehouse in Electronic Commerce, with emphasis on the analytical CRM*. São Carlos. 144 p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

Currently, one of the greatest challenges faced by companies that operate in the electronic commerce is to guarantee their customers' fidelity. Due to this fact, these companies have been considering the employment of tools capable of integrating distributed data onto operational databases in order to transform information into a competitive and distinguishing asset. Thus, the applicability of Data Warehouse tools in Electronic Commerce, together with the strategy of Customer Relationship Management (CRM), appear as a key element for customer-oriented companies , since they are capable of supplying the ones responsible for strategic decisions with competitive and essential information . Hence, through Data Warehouse tools supported by the CRM strategy, all the “clickstream” carried out by the customers in electronic commerce, can be historically stored, making it possible, through further analyses, to identify potential customers, as well as keeping them with the company. The objective of the present research is to consider the applicability of Data Warehouse in Electronic Commerce, aiming at the analytical CRM. The research also shows companies that Data Warehouse, when combined with the CRM in a Web environment, offers alternative means for customer-related data collection, thus, providing the customization of products and services, reassuring clients' loyalty to the company and, consequently, greater profits.

Key-Words: Electronic Commerce, CRM, Data Warehouse, OLAP, Data Mining.

1. INTRODUÇÃO

Neste capítulo são apresentados o contexto no qual se insere este trabalho e as justificativas que motivaram a sua realização, o objetivo proposto, as limitações consideradas, a metodologia de pesquisa adotada, a seleção da ferramenta proposta : o Data Warehouse e, por fim, a forma de organização desta dissertação.

1.1 CONTEXTO E JUSTIFICATIVA DO TRABALHO

Desde os primórdios da sociedade existe o comércio. Entende-se por comércio o ato de trocar, vender e comprar valores, produtos naturais, artificiais ou industriais. No começo era praticado pelos primeiros povos de maneira rudimentar, pois apenas trocavam um objeto pelo outro (escambo). Após as grandes expedições que culminaram com a descoberta dos novos mundos, a utilização de pedras e metais preciosos, o grande desenvolvimento científico e industrial, juntamente com as vias de comunicação possibilitou ao comércio o incentivo à produção (MAIA & MAIA, 1981).

Com o avanço da tecnologia, juntamente com a popularidade da Internet, ao falar-se em comércio, deixa-se de imaginar um estabelecimento comercial onde o cliente compra o produto pessoalmente com o vendedor, como na forma tradicional onde havia uma maior facilidade de conhecer cada cliente, seus gostos, condições financeiras e suas expectativas. Neste tipo de comércio tradicional os produtos podiam ser tocados, analisados e os clientes podiam até mesmo sentir o cheiro, entre outras coisas. Atualmente o comércio realizado via Internet, também conhecido como comércio eletrônico ou *e-commerce*, o cliente se depara com uma vitrine eletrônica, tão atraente ou mais do que o tradicional. No entanto inúmeras razões fazem com que a permanência ou a realização de uma compra em um determinado *site* não seja realizada com êxito.

Preocupados com a retenção de clientes em um *site* de comércio eletrônico, especialistas da área de *marketing* e informática vêm se absorvendo em atender às expectativas dos seus principais clientes, ou seja, conhecê-los ao máximo a fim de oferecer produtos e serviços de forma personalizada.

De acordo com PEPPERS & ROGERS (2000), a Internet trouxe mudanças em todas as áreas de negócios. A utilização de um meio tão poderoso e que está sempre (ou quase sempre) disponível causou uma revolução na maneira de ter acesso à informação e de fazer negócios.

Essa revolução fez com que muitas empresas se apressassem a disponibilizar seus produtos e serviços via Web. Porém, a grande maioria dessas empresas que ainda não estavam preparadas, desenvolveram *sites* comerciais sem nenhum planejamento e sem nenhuma integração com os seus outros sistemas.

Integrar os dados é um dos grandes desafios que as empresas enfrentam. Porém ao se conceber um *site* de comércio eletrônico, deve-se estar consciente que se abre um canal extremamente poderoso para a criação e o desenvolvimento do gerenciamento de relações com os clientes, também conhecido como CRM (*Customer Relationship Management*).

O CRM analítico, fonte de toda a inteligência do processo, serve para o ajuste das estratégias de diferenciação de clientes, bem como para o acompanhamento de seus hábitos, com o objetivo de identificar suas necessidades e os eventos que possam ocorrer na vida de cada um deles. Tudo com o objetivo de tornar, para os melhores clientes, a relação com a empresa algo fácil e conveniente, buscando sua satisfação e fidelidade. Entende-se por “melhores clientes” os clientes que representam um valor real para empresa, ou seja, os que geram as mais altas margens de lucro e os que possuem maior potencial de crescimento.

“Manter um cliente significa desenvolver um empreendimento que tornará mais conveniente para ele permanecer fiel a um determinado produto. O sucesso desse empreendimento depende de um sistema no qual o cliente ensinará a uma empresa como suprir suas necessidades individuais. Uma vez incorporadas às necessidades individuais desse determinado cliente, será mais fácil para uma empresa mantê-lo” (PEPPERS, 2000, p50).

Além disso, é através das ferramentas analíticas que identifica-se o cliente de menor e maior valor para a empresa a fim de conseguir uma redução de custo de atendimento a esses dois tipos de clientes a partir de tratamentos diferenciados.

Segundo TAURION (2000), o estágio analítico do CRM envolve o tratamento das informações sobre clientes para gerar um maior conhecimento sobre seus hábitos e interesses. Com base nesse conhecimento, a empresa pode refletir suas ações de *marketing* refinando a segmentação de seus clientes. Os componentes tecnológicos deste nível envolvem um Data Warehouse (DW) e as ferramentas de análise como OLAP (*On line Analytical Processing*) e Data Mining (Minerando dados).

Para PACE (2000), as empresas precisam implementar DWs para poderem descobrir qual a lucratividade efetivamente proporcionada aos seus clientes. A partir disso, utiliza-se sofisticados programas de CRM para lograr proveito dessas informações.

No entanto observa-se que as transações de negócios eletronicamente feitas entre empresas e clientes, oferecem às empresas excelentes oportunidades de coleta de dados a

respeito dos clientes. Utilizando-se desses dados com o DW, as empresas serão capazes de obter um conhecimento sobre o comportamento de cada um de seus clientes. Assim elas promoverão a personalização de seus produtos e serviços de acordo com o perfil de cada cliente, contribuindo para uma maior lealdade desses com a empresa e o aumento de lucratividade.

↳ A forma mais desenvolvida de utilizar dados de comportamento de um *site* da Web é construir um modelo completo de lucratividade do canal de comércio pela Web de uma empresa. Esta é uma área em que o conhecimento acumulado sobre DW da indústria existente auxilia enormemente (KIMBALL & MERZ, 2000).

De acordo com KIMBALL & MERZ (2000), a revolução da Web certamente não substituiu a necessidade de DW – ela elevou significativamente as expectativas de que todos os tipos de informações seriam publicados de forma transparente através de interfaces de navegadores da Web. A audiência para dados de DW cresceu do gerenciamento interno para a inclusão de clientes, parceiros e um grupo muito maior de funcionários internos. O foco da Web na “experiência do cliente” fez com que muitas empresas se tornassem muito mais conscientes da necessidade de aprender sobre o cliente e de fornecer informações úteis a este cliente.

Para PEPPERS & ROGERS (2000), os melhores *sites* da Web são organizados ao redor das distintas necessidades dos clientes em vez de produtos. A melhor forma de realizar isto é através da observação do comportamento do cliente no *site*, de suas transações e interações.

↳ Desta forma, observa-se que pesquisadores e executivos passam a considerar que a utilização do DW no comércio eletrônico torna-se um elemento central e coesivo da empresa moderna focada no cliente, fornecendo informações competitivas e essenciais a gerentes e responsáveis pelas decisões estratégicas.

Este trabalho apresenta uma proposta abrangente sobre os mais importantes aspectos relacionados com a aplicação do DW no comércio eletrônico para apoiar o gerenciamento do relacionamento com o cliente (CRM). Pelo grande interesse das organizações no que se refere à implementação de CRM e pela importância que o comércio eletrônico vem adquirindo ultimamente, considera-se que este trabalho será de grande utilidade para o entendimento e utilização da ferramenta Data Warehouse envolvida neste cenário.

1.2 OBJETIVO DO TRABALHO

O objetivo do trabalho é analisar a aplicabilidade do DW no comércio eletrônico, visando o CRM analítico, com base no modelo proposto por KIMBALL & MERZ (2000).

Segundo AURÉLIO (1999), a palavra analisar significa:

1. Decompor (um todo) em suas partes componentes; fazer análise de;
2. Observar, examinar com minúcia; esquadrinhar;
3. Submeter a crítica; examinar criticamente;
4. Decompor (uma oração ou período) nos seus elementos, a fim de classificar cada um destes, de acordo com as regras gramaticais.

A análise proposta neste trabalho baseia-se em examinar criticamente a elaboração de um DW para a Web e explicar suas ligações com o CRM analítico no comércio eletrônico.

O trabalho apresenta uma proposta abrangente sobre os mais importantes aspectos relacionados com a aplicação do DW no comércio eletrônico para apoiar o CRM de forma eficaz.

1.3 LIMITAÇÕES DO TRABALHO

A limitação da proposta desta pesquisa, utilizando a ferramenta DW, deve-se às características que a mesma apresenta, ou seja, o DW se refere ao processo de criação de um armazém de dados pela depuração e o armazenamento desses por assunto, tornando-os disponíveis para a recuperação on-line. Este armazém contém dados históricos integrados por assunto, para utilização em ambientes de análise de dados e de Sistemas de Apoio à Decisão (SAD).

Uma característica importante do DW é que ele é um banco de dados “desenhado” para ferramentas analíticas usando dados de diferentes aplicações.

A definição dos aspectos do DW no comércio eletrônico visando o CRM analítico neste trabalho é baseado principalmente nas explicações de como construir o DW para a Web apresentado por KIMBALL & MERZ (2000) e, nas observações sobre CRM analítico no comércio eletrônico essencialmente apresentadas por PEPPER S & ROGERS (2000), dentre outros autores.

Para que seja possível realizar uma análise da aplicabilidade do DW no comércio eletrônico, houve a necessidade de delimitar o assunto somente ao CRM analítico devido ao fato de que o DW no comércio eletrônico possui uma generosa abrangência: dentre eles pode-se citar o impacto tecnológico, o comprometimento organizacional, a segurança e a

privacidade, os aspectos legais, o gerenciamento da cadeia de suprimentos, a resposta eficiente ao consumidor, a inteligência de negócios, entre outros.

Além de enfatizar a importância do DW no comércio eletrônico visando o CRM analítico, são estudados também alguns aspectos de tecnologia de análise e de mineração de dados como as ferramentas OLAP e o Data Mining sem se aprofundar em alguma técnica específica.

Aspectos de investimento e custo do DW no comércio eletrônico visando o CRM analítico, por sua vez, não são considerados no escopo deste trabalho.

1.4 MÉTODO DE PESQUISA

A seguir são discutidos os aspectos metodológicos considerados na realização desta pesquisa. Partindo-se do objetivo estabelecido, são elaboradas as perguntas de pesquisa, é escolhido o método de pesquisa empregado e são definidas as fases da pesquisa e as etapas do trabalho.

1.4.1 PERGUNTAS DE PESQUISA

A partir do objetivo do trabalho, foram definidas as seguintes perguntas da pesquisa as quais seguem uma seqüência lógica:

1. Como trazer a Web para o DW a fim de direcionar ao CRM ?
2. Como monitorar as ações dos usuários de um *site* da Web ?
3. Como utilizar a seqüência de dados para exercer o CRM ?
4. Como compreender a seqüência de cliques como uma fonte de dados necessária ao CRM?
5. Como construir Data Marts a partir de uma seqüência de cliques ?
6. Como fazer para gerenciar o DW ?

1.4.2 ESCOLHA DO MÉTODO DE PESQUISA

Para se atingir o objetivo proposto e responder as perguntas estabelecidas para este trabalho, é empregado uma abordagem de pesquisa bibliográfica.

De acordo com CERVO & BERVIAN (1996), a pesquisa bibliográfica procura explicar um problema a partir de referências teóricas publicadas em documentos. Busca conhecer e analisar as contribuições culturais ou científicas do passado existente sobre um

determinado assunto, tema ou problema. Ela também evita a duplicidade de pesquisa para a base de conhecimentos e, obtém fundamentação teórica para o desenvolvimento do projeto.

KÖCHE (2000)¹ *apud* SILVA & SACOMANO (2001) define na página 122 que “a pesquisa bibliográfica é a que se desenvolve tentando explicar um problema, utilizando o conhecimento disponível a partir das teorias em livros ou obras congêneres. Na pesquisa bibliográfica o investigador irá levantar o conhecimento disponível na área, identificando as teorias produzidas, analisando-as e avaliando a sua contribuição para auxiliar a compreender ou explicar o problema objeto da investigação.”

Segundo GIL (1991)² *apud* SILVA & MENEZES (2001), a pesquisa bibliográfica é elaborada a partir de material já publicado, constituído principalmente de livros, artigos de periódicos e atualmente com material disponibilizado na Internet.

THIOLLENT (1983)³ *apud* SILVA & SACOMANO (2001) explana na página 65, “em matéria de investigação científica, pesquisar não consiste apenas em observar as coisas através de instrumentos, mas também exige do pesquisador uma formação intelectual bastante ampla para poder colocar as coisas em perspectivas, perceber ou aprender os aspectos do real de modo que seja relevante em função de determinado propósito científico que só pode ser definido à luz de uma teoria. (...) O pesquisador deve conhecer as diversas teorias relativas ao assunto e, a partir delas, selecionar os elementos a serem aplicados para explicar ou resolver os problemas reais.”

A abordagem de pesquisa bibliográfica é considerada adequada a este trabalho, pois o que se pretende é analisar a aplicabilidade do DW no comércio eletrônico visando o CRM analítico, determinando a integração entre estes.

1.4.3 FASES DA PESQUISA

Este trabalho é subdividido em três fases: a Definição do Trabalho, a Pesquisa Bibliográfica e a Exposição dos Resultados.

O emprego dessas três fases da pesquisa, visa à obtenção de uma abrangente análise da aplicabilidade do DW no comércio eletrônico, visando o CRM analítico. As fases da pesquisa são descritas a seguir.

1 KÖCHE, J. C. (2000). Fundamentos de metodologia científica: teoria da ciência e prática da pesquisa, , 18ª ed. Petrópolis. Vozes

2 GIL, A. C. (1991). Métodos e técnicas de pesquisa social. São Paulo. Atlas.

3THIOLLENT, M. (1983). Problemas de metodologia. In: A. C. C. Fleury e N. Vargas. Organização do trabalho: uma abordagem interdisciplinar - sete casos brasileiros para estudo . São Paulo. Atlas.

- Fase 1:

Definição do Trabalho

A definição do trabalho abrange o estabelecimento do objetivo a ser cumprido, a especificação das perguntas de pesquisa, a escolha do método de pesquisa, a definição das fases da pesquisa e das etapas do trabalho e por fim a escolha da ferramenta DW.

- Fase 2:

Pesquisa Bibliográfica

A pesquisa bibliográfica neste trabalho é realizada de acordo com os objetivos estabelecidos por CERVO & BERVIAN (1983): procura explicar um problema a partir de referências teóricas publicadas em documentos; evita duplicidade de pesquisa; evita problemas ocorridos em trabalhos anteriores; obtém fundamentação teórica.

Esta fase da pesquisa incorpora a identificação dos dados comportamentais de consumidores realizados no comércio eletrônico, a ferramenta DW para o armazenamento desses dados distribuídos e, uma estratégia de negócio (CRM) para então, propor o modelo integrado.

Inicialmente, nesta etapa será realizado um levantamento bibliográfico do Comércio eletrônico, do CRM e do DW.

A partir desta pesquisa bibliográfica, será apresentada uma análise do DW no comércio eletrônico visando o CRM analítico a partir de um modelo integrado proposto por KIMBALL & MERZ (2000).

Essa fase também tem como objetivo apresentar o gerenciamento do modelo integrado, ou seja, gerenciar pessoas e atribuir atividades a elas de acordo com suas habilidades.

- Fase 3:

Exposição dos Resultados

São expostos finalmente, os resultados observados em todas as fases anteriores, são apresentadas as conclusões obtidas e são realizadas sugestões para trabalhos futuros nesta área de pesquisa.

1.4.4 ETAPAS DO TRABALHO

As etapas do trabalho são definidas a partir das fases que configuram a pesquisa e do método adotado.

As etapas definidas para esse trabalho em cada uma das fases da pesquisa são apresentadas na FIGURA 1. Os retângulos maiores e pontilhados, no fundo da FIGURA 1, representam as três fases da pesquisa, discutidas anteriormente. O nome de cada uma dessas fases está inserido no canto superior esquerdo de cada um desses retângulos. Os retângulos menores representam as etapas do trabalho. As flechas que interligam os retângulos, representam a seqüência lógica entre as etapas. O resultado das etapas é representado por um retângulo com parte inferior curvilínea. As etapas do trabalho são apresentadas detalhadamente na FIGURA 1.

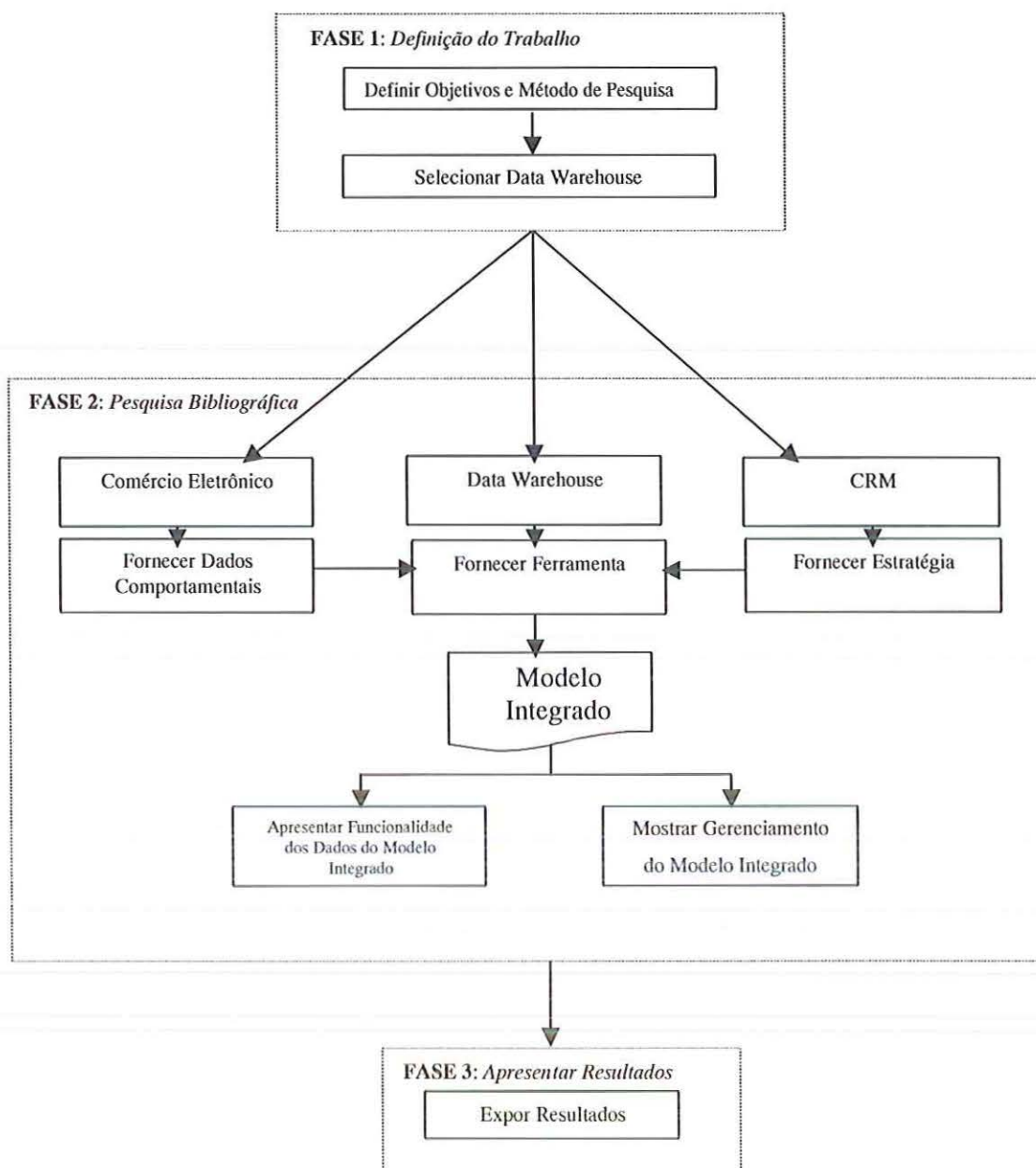


Figura 1: Fases da Pesquisa e etapas do trabalho

- **FASE 1: Definição do Trabalho**

- Etapa 1: Definir Objetivo e Método de Pesquisa**

Esta etapa abrange o estabelecimento do objetivo a ser cumprido, a determinação das perguntas, a escolha do método de pesquisa e a definição das fases da pesquisa e das etapas do trabalho.

- Etapa 2: Selecionar Data Warehouse**

Logo, é selecionado a ferramenta DW. Sabendo-se que o Data Webhouse é a instanciação de Web do DW, a ferramenta deve refletir a realidade encontrada nas empresas que realizam o comércio eletrônico a fim de obter uma considerável expectativa em relação ao gerenciamento do relacionamento com o cliente. Em outras palavras, a ferramenta selecionada deve ser capaz de armazenar um grande volume de dados, principalmente a respeito do comportamento dos clientes na Web, a fim de serem analisados para então haver a tomada de decisões de acordo com as necessidades das empresas situadas neste contexto.

- **FASE 2: Pesquisa Bibliográfica**

Durante o desenvolvimento de todo o trabalho, a pesquisa bibliográfica é utilizada como abordagem metodológica para o desenvolvimento deste.

Para direcionar a revisão bibliográfica são definidos três temas principais: Comércio Eletrônico, CRM e DW.

O primeiro tema abrange diversas áreas do conhecimento. Por esse motivo, o trabalho limitou-se a realizar a pesquisa bibliográfica de seus principais conceitos, formas, aspectos, segurança e a questão do relacionamento humano neste cenário. Devido à proposta inicial, o trabalho limitou-se também a conceder enfoque a um tipo de comércio eletrônico: o business-to-consumer (negócios para consumidores).

O segundo tema, também por possuir uma característica de assunto muito estudado por diversos pesquisadores e executivos, o trabalho procurou limitar-se principalmente em seus principais conceitos, estratégias e o papel do CRM na Web.

Já o terceiro tema, existe poucas pesquisas científicas e material publicado a respeito do DW na Web. Inicialmente o trabalho expõe os conceitos, ferramentas utilizadas, características, arquitetura do DW e a modelagem em um ambiente tradicional. Logo então, o papel da ferramenta em um outro cenário, a Web, é explicado. Devido a essa limitação, a pesquisa bibliográfica é complementada através de periódicos não científicos publicados por profissionais da área que estudam o assunto em questão, *sites* voltados para a área de tecnologia da informação, livros de referência de DW e um livro específico de Data Webhouse. Este assunto restringe principalmente aos dados referentes ao comportamento dos clientes.

Etapa 3: Comércio Eletrônico

A pesquisa bibliográfica é iniciada com a seleção do comércio eletrônico, para gerar dados para o DW.

O estudo do comércio eletrônico deve ser bastante amplo. Engloba desde os tipos de comércio eletrônico até os dados comportamentais que são armazenados em bancos distribuídos – dados esses que serão vitais para a revelação do perfil dos principais clientes.

Etapa 4: CRM

Da mesma forma ao tema anterior, o gerenciamento do relacionamento com o cliente (CRM), é pesquisado com o intuito de explicar a necessidade de um relacionamento personalizado.

Devido o trabalho estar inserido em um contexto Web, o CRM será explorado inicialmente em seus diversos conceitos, para então ser enfocando em seu contexto analítico na Web.

Esta etapa procura fornecer as estratégias para a fidelização de um cliente na Web.

De acordo com PEPPERS & ROGERS (2000), os melhores *sites* incentivam a permanência dos clientes e os desenvolvem ao longo do tempo. A melhor forma de realizar isso é observando o comportamento do cliente no *site*, de suas transações e interações. Dessa forma pode-se antecipar as necessidades dos clientes e apresentar ofertas, informações e recomendações personalizadas. Os melhores *sites* da Web armazenam dados que permitem a antecipação de necessidade.

Etapa 5: Data Warehouse

Esta etapa explica os principais conceitos e características do DW - uma ferramenta capaz de integrar e consolidar dados para fins de exploração e análise.

Segundo BISPO (1998), esta ferramenta deve tornar-se imprescindível para os gerentes conseguirem administrar seus negócios nos próximos anos.

BRETZKE (1999), argumenta que uma das principais justificativas para desenvolver um DW está em atender às necessidades de informação dos tomadores de decisão e especialistas em negócio.

Como o cenário estudado no trabalho é a Web, após explicar o DW, a ferramenta será adaptada em um ambiente Web.

Devido ao intensivo volume de dados oriundos de uma seqüência de cliques de cada usuário da Web, a necessidade de minerar esses dados surgiu como uma ferramenta extremamente útil às organizações. É desta forma que o trabalho descreve esse assunto sem o intuito de aprofundar em algum método ou técnica específica.

Após “minerado” os dados, deve ser realizado um processamento analítico de suporte *on line*, através de uma arquitetura de aplicação interativa, ou seja, o analista pode

“jogar” um valor para simular algo ou até mesmo descobrir padrões encobertos. A necessidade de receber um intenso número de dados de um grande banco de dados, são os motivos de existir o OLAP.

A mineração e a análise dos dados proporcionam uma melhora no uso do enorme volume de dados armazenados no DW, compreendendo principalmente o comportamento e o perfil de clientes praticantes do comércio eletrônico. Este assunto é descrito nesta etapa.

Etapa 6: Fornecer Dados Comportamentais

Esta etapa compreende uma seqüência de cliques realizadas em um *site* de comércio eletrônico como uma fonte de dados. Descreve como clientes e *sites* da Web se comunicam um com os outros. Mostra também com alguns detalhes, quantas informações podem ser derivadas de um *cookie*⁴ e quais podem ser as limitações que mesmo um “bom” *cookie* pode ter. Expõe as informações de referência, que é uma fonte de idéias sobre as razões pelas quais o usuário chegou a um determinado *site* da Web, tanto por motivos certos ou por razões erradas.

Será portanto, a fonte de dados que “alimentará” o DW no comércio eletrônico através de registros de *log*⁵ produzidos pelo servidor da Web toda vez que uma solicitação for satisfeita.

Etapa 7: Fornecer Estratégia

Esta etapa abrange o ajuste das estratégias de diferenciação de clientes, bem como para o acompanhamento de seus hábitos, com o objetivo de identificar suas necessidades e os eventos que possam ocorrer na vida de cada um deles.

Basicamente as estratégias fornecidos pelo CRM para dar suporte ao comércio eletrônico são:

- ✓ Identificar os clientes;
- ✓ Diferenciar os clientes;
- ✓ Interagir com os clientes;
- ✓ Personalizar produtos e serviços.

⁴ Nota da autora: *cookie* são pequenos arquivos textos gravados no disco rígido contendo informações que o próprio usuário forneceu ao *site*.

⁵ Nota da autora: Log consiste em uma identificação do usuário para ter acesso ao servidor da rede.

Etapa 8: Fornecer Ferramenta

Paralelamente ao tema anterior, o DW é apresentado como uma ferramenta capaz de integrar os dados distribuídos em banco de dados operacionais.

Como o cenário estudado é a Web, o DW será adaptado para um ambiente Web mas, seguindo as mesmas características: orientado por assunto, integrado, não volátil e variável em relação ao tempo.

Como o tema é abrangente, a ferramenta DW focará principalmente nos dados relacionados ao comportamento dos clientes que realizam transações econômicas em um *site* da Web.

Etapa 9: Apresentar Funcionalidade do Modelo Integrado

Entende-se por modelo integrado, o DW no comércio eletrônico enfocando o CRM analítico.

Nesta etapa são apresentados os dados de seqüência de cliques, a exploração e análise desses dados. Como se trata de um modelo integrado, os dados apresentados são de acordo com o comportamento dos usuários.

Se o armazenamento dos dados de uma seqüência de cliques for feito de uma forma correta, terá uma plataforma para entender o comportamento de um visita nte da Web. Mostra transitoriamente como é feita exploração dos dados através de uma ferramenta analítica.

Etapa 10: Mostrar Gerenciamento do Modelo Integrado

Esta etapa descreve um estudo de como um gerente de projeto de Data Webhouse deve portar-se. Mostra que além de executar uma grande quantidade de tarefas diferentes, o gerente de projeto deve também gerenciar pessoas com diferentes habilidades e diferentes perspectivas.

- **FASE 3: Apresentar Resultados**

Etapa 13: Expor Resultados

Nesta etapa são observados e discutidos os resultados obtidos na pesquisa bibliográfica. São também apresentadas as conclusões obtidas relacionadas com o tema da pesquisa.

1.5 SELEÇÃO DA FERRAMENTA DATA WAREHOUSE

A Seleção da ferramenta DW neste trabalho é baseada no princípio de que a ferramenta é capaz de integrar dados distribuídos em um ambiente Web. Devido ao grande volume de informações geradas em um ambiente Web, os métodos tradicionais de retenção de clientes estão se tornando cada vez mais ineficazes.

As empresas que realizam transações comerciais eletronicamente, devem “ter em mente” que essas informações quando integradas e bem gerenciadas, causam um diferencial notável em termos de retenção de clientes. O comércio eletrônico pode beneficiar -se de um DW para realizar análises cruzadas dessas informações a respeito dos clientes.

É através de ferramentas analíticas como o OLAP utilizada pelos usuários finais para análise dos dados extraídos do DW, que a possibilidade de retenção de clientes ficará mais fácil, pois os dados comportamentais estarão sendo extraídos e analisados.

Em outras palavras, os dados oriundos de diversas fontes de dados, são armazenados no DW, os quais são analisados e cruzados, fornecendo informação integrada e consistente sobre a empresa, o negócio e é capaz de consolidar dados importantes a respeito do comportamento de clientes.

Tal ferramenta no comércio eletrônico traz benefícios às organizações, como:

- Melhorar a comunicação com os clientes;
- Direcionar atividades de *marketing* pelo reconhecimento da identidade do cliente e agrupando o cliente por atributos de seqüências de cliques;
- Identificar quando os clientes utilizam um determinado *site* da Web, e qual é o horário local que o fazem;
- Visualizar se um cliente está prestes a deixar o *site*;
- Decidir se um anúncio personalizado está funcionando;
- Decidir se uma promoção está sendo lucrativa;
- Decidir quais clientes são lucrativos;
- Conquistar a fidelidade de um consumidor em potencial, sendo que este é mais lucrativo que a obtenção de um novo consumidor.

É por essa razão que a ferramenta DW foi selecionada, pois faz com que o grande volume de dados gerados diariamente no comércio eletrônico seja armazenado, analisado e utilizado para satisfazer seus clientes de forma individual, resultando então em diferencial competitivo.

1.6 ORGANIZAÇÃO DO TEXTO

Este texto está organizado em sete capítulos.

CAPÍTULO 1 - Apresenta a definição do trabalho. Em princípio discute-se o contexto no qual se insere o trabalho e as justificativas que motivaram o seu desenvolvimento. São apresentados também o objetivo e as limitações do trabalho, as perguntas da pesquisa, o método de pesquisa e as fases da pesquisa.

CAPÍTULO 2 - Apresenta a pesquisa bibliográfica do comércio eletrônico, bem como suas respectivas contribuições.

CAPÍTULO 3 - Apresenta a pesquisa bibliográfica do CRM, bem como suas respectivas contribuições.

CAPÍTULO 4 - Apresenta a pesquisa bibliográfica do DW, das ferramentas de análise e mineração de dados e suas contribuições.

CAPÍTULO 5 - Trata-se da apresentação, análise e gerenciamento do modelo integrado.

CAPÍTULO 6 - São apresentadas as conclusões obtidas e são feitas algumas considerações finais e sugestões para trabalhos futuros nesta área de pesquisa.

2. COMÉRCIO ELETRÔNICO

2.1 CONSIDERAÇÕES INICIAIS

Atualmente, as empresas têm se deparado com uma nova forma de se realizar as transações comerciais: o comércio eletrônico. É por meio do ambiente da Web que as empresas vendem seus produtos e serviços aos seus clientes de forma virtual. Mas para operar neste novo cenário, as empresas devem entender primeiramente o que vem a ser comércio eletrônico, bem como seus aspectos e seus benefícios.

Além da questão conceitual, as empresas também se deparam com outras problemáticas que são consideradas fatores-chaves de sucesso, tais como a segurança de dados dos clientes que possivelmente comprarão os produtos e serviços em um determinado *site* de comércio eletrônico, além da questão do relacionamento com os clientes, uma vez que o contato não é realizado fisicamente.

Esse capítulo tem como objetivo apresentar os principais conceitos, formas e aspectos do comércio eletrônico. Por apresentar grande importância, mas não ser o foco deste trabalho, a segurança e suas principais medidas no comércio eletrônico serão comentadas. Finalmente será apresentada uma introdução sobre o relacionamento com os clientes neste cenário.

2.2 OS PRINCIPAIS CONCEITOS DE COMÉRCIO ELETRÔNICO

Tradicionalmente, quando se fala em comércio eletrônico, imagina-se a venda para consumidores finais. No entanto, mais do que isso, no comércio eletrônico são realizadas transações comerciais através de computadores entre empresa-empresa e empresa-consumidor.

De acordo com GONÇALVES (2000), quando a Internet surgiu, era um instrumento acadêmico para a troca de informações; com a abertura às empresas, não demorou muito até que alguém pensasse em como poderia lucrar com essa nova tecnologia.

O primeiro uso comercial da rede foi a criação de *sites* para divulgação de produtos, serviços e história. A expansão foi rápida, e hoje acredita-se que todas as empresas de grande, médio e até de pequeno porte já possuem o seu *site*, podendo divulgar seus produtos/serviços de forma globalizada. Hoje, considera-se sinônimo de sobrevivência as empresas que possuem transações comerciais via Internet.

Mas para entender o que é de fato o comércio eletrônico, é preciso conhecer alguns conceitos vitais para se obter sucesso neste ramo promissor.

Para MAY (2000), existem muitas definições de comércio eletrônico e todas são úteis para algumas dimensões.

Segundo CAMERON (1997), o comércio eletrônico inclui qualquer negócio transacionado eletronicamente, sendo que essas transações ocorrem entre dois parceiros de negócio ou entre um negócio e seus clientes.

De acordo com ALBERTIN (2000) o comércio eletrônico é a realização de toda a cadeia e valores dos processos de negócios em um ambiente eletrônico, por meio da aplicação intensa das tecnologias de comunicação e de informação, atendendo aos objetivos de negócio.

TAURION (1998) afirma que o comércio eletrônico é um processo irreversível – seus impactos serão grandes e as práticas comerciais serão modificadas. Um exemplo será a mudança do tradicional esquema de preços fixos por preços variáveis modificados de acordo com cada cliente, coletando, ao mesmo tempo, informações bastante valiosas sobre esses consumidores, como seus hábitos de consumo.

Basicamente, ao analisar o comércio eletrônico, deve atentar-se ao desenvolvimento de produtos, à cadeia de suprimentos, à produção e certamente ao atendimento a clientes.

APPLEGATE *et al* (1996), descrevem que o comércio eletrônico envolve mais do que compra e venda, inclui também todos os tipos de esforços no pré e pós-venda, assim como um conjunto de atividades auxiliares, novos enfoques para pesquisa de mercado, geração de condução qualificada de vendas, anúncios, distribuição de produtos, compra, apoio ao cliente, recrutamento, operações de negócios, relações públicas, administração da produção, distribuição de conhecimento e transações financeiras. De acordo com os autores, essas atividades afetam o planejamento estratégico, oportunidades empreendedoras, projeto e desempenho organizacional, leis de negócio e políticas de transação.

Nota-se portanto que o comércio eletrônico não se restringe apenas ao acesso de lojas através da Internet, abrange também todo o serviço relacionado ao processo de venda de produtos e serviços.

Embora o comércio eletrônico signifique uma revolução na oferta de produtos e serviços, ele não surgiu para substituir a forma tradicional de comércio: o comércio eletrônico surgiu para proporcionar agilidade ao processo de aquisição de produtos e serviços.

Entendendo-se os principais conceitos de comércio eletrônico é necessário detalhar as formas de realizá-lo, sendo elas: *business-to-consumer* e *business-to-business*, discutidos no próximo item.

2.3 FORMAS DE COMÉRCIO ELETRÔNICO: *BUSINESS- TO-BUSINESS* e *BUSINESS-TO-CONSUMER*

Business-to-business ou B2B é o nome que se dá às transações comerciais entre empresas. Este tipo de comércio eletrônico é realizado por empresas que se relacionam com seus fornecedores, realizando pedidos, recebendo e pagando faturas e pagamentos, trocando dados e até mesmo capturando novos parceiros. Essa forma de comércio eletrônico visa aumentar o grau de integração entre empresas e buscar maior flexibilidade nas transações entre os fornecedores. Em outras palavras, essa forma de comércio eletrônico compreende a ligação entre empresas através da rede.

De acordo com ALBERTIN (2000a), nessa perspectiva (B2B), o comércio eletrônico facilita as seguintes aplicações:

- Gerenciamento de fornecedor;
- Gerenciamento de estoque;
- Gerenciamento de distribuição;
- Gerenciamento de canal de vendas;
- Gerenciamento de pagamento.

Já no *Business-to-consumer* ou B2C é o nome que se dá às transações comerciais realizadas entre a empresa e o consumidor.

De acordo com MAY (2000), B2C é um termo que evidencia a direção de entrega, ou seja, o comércio eletrônico realizado pela forma B2C efetua negócios para os consumidores.

Para ALBERTIN (2000a), nesse tipo de transação os clientes aprendem sobre os produtos por meio de publicação eletrônica, compram produtos com dinheiro eletrônico e outros sistemas de pagamento seguros. Aqui o comércio eletrônico facilita as seguintes transações econômicas:

- Interação social;
- Gerenciamento de finanças pessoal;
- Informações e compra de produtos.

2.4 ASPECTOS DO COMÉRCIO ELETRÔNICO

Para ALBERTIN (2000a), os aspectos do comércio eletrônico podem ser entendidos da seguinte maneira:

- **Adoção:** A chave para o sucesso do sistema do comércio eletrônico para clientes, certamente é uma grande adoção desses tipos de tecnologias por parte dos clientes;
- **Relacionamento:** Com o comércio eletrônico houve uma grande mudança no relacionamento entre cliente e fornecedor, devido a eliminação de intermediários, diminuição da interação *face-to-face*, maior informação sobre o cliente, facilidade de acesso, integração, entre outros;
- **Adequação:** O comércio eletrônico faz com que novas maneiras de produtos e serviços se adequem às novas formas de divulgação;
- **Estratégia:** Para que o comércio eletrônico ofereça uma grande quantidade de novas oportunidades é preciso que a organização altere suas estratégias;
- **Comprometimento Organizacional:** Este é imprescindível para que o comércio eletrônico tenha sucesso – incluem significativos investimentos para criação e manutenção dos sistemas, administração de mudanças organizacionais profundas, gerenciamento de riscos, etc.;
- **Privacidade e Segurança:** Esta é a grande preocupação dos clientes, os quais ainda se sentem intimidados quanto à utilização dos sistemas do comércio eletrônico;
- **Sistemas Eletrônicos de Pagamento:** Transações eletrônicas entre vendedores e compradores só terão sucesso num ambiente simples, universalmente aceito, seguro e barato;
- **Aspectos Legais:** Estes podem restringir a aplicação de comércio eletrônico (por exemplo, a regulamentação de um *site*);
- **Aspectos de implementação:** A tecnologia precisa estar alinhada com a organização e adequada à tecnologia externa. Sozinha, a tecnologia não cria vantagem.

Ainda para o autor, a análise do comércio eletrônico deve ser feita a partir de duas dimensões: Aspectos e Contribuições do comércio eletrônico, ilustradas na TABELA 1.

Dimensões	Aspectos	Contribuições
<i>Clientes e Fornecedores</i>	<ul style="list-style-type: none"> • Adoção • Relacionamento 	<ul style="list-style-type: none"> • Relacionamento
<i>Produtos e Serviços</i>	<ul style="list-style-type: none"> • Adequação 	<ul style="list-style-type: none"> • Customização • Inovação • Novos Canais • Promoção
<i>Organização</i>	<ul style="list-style-type: none"> • Estratégia • Comprometimento 	<ul style="list-style-type: none"> • Novas Oportunidades • Estratégia Competitiva • Economia Direta
<i>Tecnologia</i>	<ul style="list-style-type: none"> • Privacidade e Segurança • Sistemas Eletrônicos de Pagamento • Aspectos de Implementação • Aspectos Legais 	<ul style="list-style-type: none"> • Infra-estrutura Pública

Tabela 1: Estrutura de Análise do Comércio Eletrônico
 Fonte: ALBERTIN, 2000a, p199

Semelhante a linha de pensamento de ALBERTIN (2000a), LICO (2001) encara os aspectos do comércio eletrônico como fatores de sucesso. Para LICO (2001), em qualquer projeto deve-se identificar quais são os fatores de sucesso, aqueles em que se deve dedicar mais tempo e que será todo o diferencial competitivo do projeto. No caso específico do comércio eletrônico, os fatores de sucesso são importantes da mesma forma. São identificados sete fatores de sucesso de qualquer projeto de comércio eletrônico que são considerados chaves em qualquer implementação e devem ser alinhavados com um excelente Gerenciamento de Projetos. Um bom Gerenciamento de Projetos além de garantir que o projeto seja implementado no tempo correto, também possibilita que o projeto fique dentro do orçamento. Os sete fatores de sucesso no comércio eletrônico são:

1. Estratégia de E-Commerce. A estratégia de negócio deve ser dinâmica, baseada em oportunidades para fornecer vantagens. O desenvolvimento desta estratégia requer uma compreensão da empresa, do segmento do mercado e dos recursos e

atributos exclusivos das tecnologias da Internet, com visão para o futuro. Os objetivos de uma estratégia de comércio eletrônico podem ser:

- Melhor interação e atendimento ao cliente;
- Diferencial tecnológico;
- Maior conhecimento e percepção da marca;
- Expansão do alcance geográfico;
- Maior faturamento e participação no mercado;
- Redução de custos operacionais;
- Processos agilizados, redução de erros e maior produtividade.

LICO (2001) afirma que para se atingir esses objetivos, a estratégia de comércio eletrônico deve levar em conta os seguintes aspectos:

- Ser Focada no Consumidor;
- Possuir Criação de Valor Contínua (a organização não deve perder oportunidades estratégicas);
- Ser ágil (fazer com que a companhia seja ágil internamente para responder as constantes mudanças pedidas pelos consumidores e pelas novas estratégias da concorrência).

Uma estratégia deve ser realista desde o início, ou seja, seus objetivos devem ter como base o orçamento para a tecnologia e *marketing* (IBM, 2000).

2. Customer Experience: aborda a experiência do usuário na Web estudando os mesmos aspectos que ele vivência no mundo físico. Ao “navegar” pela Internet, a maioria das pessoas necessita saber eventos básicos que as orientem , da mesma forma que ocorre quando dirigem uma estrada ou quando percorrem o interior de um edifício.
3. Gerenciamento do Relacionamento com o Consumidor: manter um melhor relacionamento com os consumidores é a mais óbvia vantagem no comércio eletrônico.
4. Segurança: medidas de segurança devem ser tomadas contra as ameaças no comércio eletrônico.
5. Personalização: Como os *sites* da Web não estão "equipados" com vendedores, a Personalização torna-se extremamente importante para reter, direcionar,

incentivar e fazer com que o cliente não desista de fazer suas compras. Há quatro estágios no ciclo de vida do consumidor na Internet (TABELA 2) e a cada estágio um certo tipo de Personalização é requerido para manter o consumidor altamente interessado e envolvido para o relacionamento continuar e o cliente querer ir para o próximo estágio.

Estágio	Ações de Personalização
Chegada	O Sistema é o Alvo - Quando um visitante desconhecido chega pela primeira vez no <i>site</i> , a apresentação do <i>site</i> pode ser adaptada para o seu ambiente (Ex: Compatibilidade do <i>Browser</i> , Sistema Operacional, Linguagem etc.).
Anônimo	O Comportamento é o Alvo - Quando o visitante anônimo começa a explorar o <i>site</i> , seus interesses podem ser inferidos pelo rastreamento de seu comportamento e adaptando o conteúdo do <i>site</i> e navegação de acordo com seu gosto. Isto é feito sem que o <i>site</i> peça nenhuma informação explícita para o Visitante
Interessado	A Necessidade é o Alvo - Uma vez que o visitante tornou-se familiar com o <i>site</i> e fornece informações pessoais e preferências, um conteúdo bem determinado pode ser recomendado baseado em seu profile para manter o cliente engajado.
Conhecido	Espaço de Trabalho Pessoal - O <i>site</i> pode oferecer páginas personalizadas que são explicitamente customizadas para as necessidades do cliente. O <i>site</i> pode oferecer a capacidade do cliente definir páginas personalizadas cada vez que ele retorne ao <i>site</i> .

Tabela 2: Estágios no ciclo de vida do consumidor na Internet

Fonte: LICO, 2001, pg 82

6. Logística: importante em qualquer empresa, especialmente nas empresas que operam seus negócios de forma B2C. Vender é bom, mas entregar a mercadoria para o cliente também é primordial. Deve-se ter uma estrutura inicial antes de iniciar os negócios de forma eletrônica. Um bom exemplo nesse fator de sucesso, são as parcerias entre diversas empresas.
7. Meios de Pagamento: os meios de pagamento são um grande diferencial em qualquer *site* de comércio eletrônico. A desistência de compra por parte dos usuários é porque os tipos de pagamentos disponibilizados pelo *site* não são adequados as certas pessoas. Os *sites* não devem somente se restringir aos

pagamentos eletrônicos com Cartão de Crédito e Débito em Conta Corrente e os tradicionais, como o Boleto Bancário. Devem ser contemplados pagamentos em cheque e, até com dinheiro.

Portanto, nota-se que para se ter uma visão do que é comércio eletrônico deve-se levar em conta seus principais conceitos, aspectos, seus fatores de sucesso e suas contribuições oferecidas às organizações. Isolando a questão da segurança, que é um dos fatores chaves de sucesso que vêm preocupando tanto os responsáveis pelas empresas que operam no comércio eletrônico quanto os usuários. Medidas diversas devem ser providenciadas afim de assegurá-los. Assunto este que será abordado no próximo item.

2.5 SEGURANÇA NO COMÉRCIO ELETRÔNICO

O explosivo crescimento do uso da Internet nas empresas fornecem grandes vantagens de informações, compras e serviços aos consumidores. Embora haja um crescimento do comércio eletrônico, seu potencial ainda não é totalmente utilizado em razão de preocupações e medo com os riscos em realizar negócios eletronicamente. Na verdade, a segurança das transações eletrônicas não acompanham o crescimento do comércio eletrônico.

Para CORREA (2000), há empresas que ainda olham segurança como despesa. Além disso, as empresas encontram-se cada vez mais vulneráveis aos ataques e invasões.

Segundo ALBERTIN (1998), artigos e publicações têm proclamado o valor da Internet, porém o número de compras realizadas ainda é muito baixo. Além disso, os casos reais ou não, referentes a falta de segurança e problemas de desempenho têm contribuído para a renúncia.

TAPSCOTT (1996), define que um dos desafios que a economia digital enfrenta está relacionado com a segurança e a privacidade das informações.

O que os usuários e empresas procuram hoje é maior privacidade de suas informações. Para que o comércio eletrônico se viabilize é necessário que os usuários de Internet confiem nela para realizar suas transações comerciais e financeiras. Existem alguns requisitos funcionais para a segurança que são encarados como aspectos básicos e fundamentais. Esses requisitos são aqui adotados como referência para o estudo de algumas medidas de segurança.

De acordo com ALCOFARADO & SÊMOLA (2000), tais requisitos descritos a seguir garantirão maior confiança na Internet para que os consumidores realizem sem medo transações comerciais e financeiras. Estão divididos em cinco categorias – são elas:

1. Confidencialidade – Ninguém deseja que alguém leia correspondências particulares de negócios ou que veja o valor das compras realizadas sem consentimento próprio. Privacidade e discrição são fundamentais em qualquer negócio.
2. Autenticação – Deseja-se saber qual a real origem das informações que se recebe.
3. Integridade – Mais do que ter certeza da origem das informações, deseja-se confiar que nenhuma delas foi alterada no caminho.
4. Não repúdio – Não se deseja que alguém faça uma encomenda e depois diga que não foi o autor do pedido.
5. Autorização – Certamente o banco em que se tem conta, não deseja movimentação de somas acima do limite de saldo ou realizando operações com produtos que não se possuem.

2.6 PROBLEMAS COM A SEGURANÇA

Segundo BASTOS (1998), o uso da Internet nas empresas trouxe novas vulnerabilidades na rede interna. Se não bastassem as preocupações existentes com espionagem industrial, fraudes, erros e acidentes, as empresas precisam se preocupar agora com os *hackers*, invasões, vírus e outras ameaças que penetram através desta nova porta de acesso.

Essa vulnerabilidade ocorre devido à interatividade com os clientes, parceiros e fornecedores compartilhando as informações, banco de dados, estratégias, tecnologias, etc. A idéia de isolamento está quebrada com o aumento de transações entre empresas na Internet.

Nas operações B2B, apesar de uma das empresas adotar um elevado padrão de segurança com seus dados, por ser compartilhado, esses dados sigilosos poderão não estar seguros se a outra empresa que receber esses dados não adotar o mesmo padrão de segurança ou até mesmo ignorar a questão da segurança da informação, expondo dessa forma informações importantes.

A dependência de empresas parceiras no B2B faz com que a segurança do negócio deverá ser implementado em todos os perímetros da rede, ou seja, clientes, fornecedores e parceiros, devendo no mínimo utilizar o mesmo padrão de segurança internamente. Desta forma haverá um reconhecimento do valor estratégico da informação. Um outro problema desta categoria é a segurança na estação (cliente), ou seja, há a existência de um “browser” o

qual o usuário tem acesso aos recursos e serviços da rede, como por exemplo as bases de dados.

Segundo ALBERTIN (1998), mecanismos de controle de acesso precisam ser estabelecidos para assegurar que os usuários apropriadamente autenticados terão acesso a somente aqueles recursos previamente a eles autorizados. Tais mecanismos incluem proteção de senha, cartões inteligentes criptografados, *biometrics* e *firewall*.

Também relacionado como um problema de segurança está o transporte ou transmissão das informações. Esses devem possuir uma atenção especial. É utilizada a criptografia de dados para garantir a privacidade e a integridade das informações que transitam pela Internet ou Intranet. Além disso, a criptografia vai garantir a autenticidade do conteúdo de uma informação. Para ALBERTIN (1998), o objetivo é invalidar qualquer tentativa de assumir uma outra identidade quando correio eletrônico ou outras formas de comunicação de dados que estão envolvidos.

Sem deixar de mencionar, há também a segurança nos servidores. Nesta categoria os problemas mais ocorrentes são roubos de informações, “pichações de *sites*”, alteração de arquivos por imagens pornográficas ou dados físicos, enfim. Esse tipo de problema compromete muito a imagem e reputação de uma empresa. O que as empresas têm tomado como medida de segurança neste ponto é o uso de *Firewall*, que controla os serviços e acessos, monitorando o uso e as tentativas de violação protegendo contra as invasões.

Foram apresentados apenas alguns dos principais problemas relativos à segurança. Existem outros e novos problemas, os quais não foram citados, mas que merecem igual atenção.

2.7 PRINCIPAIS MEDIDAS DE SEGURANÇA

As medidas de segurança apresentadas a seguir requerem um nível adequado à natureza de cada atividade.

- **Criptografia:** é a ciência de se escrever mensagens de forma que apenas o receptor consiga decifrá-la. A segurança para correio eletrônico tem sido feita através do uso de criptografia, pois, como já foi citado, a criptografia preserva a privacidade e integridade das informações que transitam via Internet ou Intranet. A mensagem original é passada por um processo de codificação gerando uma mensagem cifrada. O receptor que recebe essa mensagem cifrada utiliza um processo chamado decodificação ou decifração para decifrar a mensagem original. Um bom processo de codificação exigiria um longo tempo de

processamento de decodificação para que o conhecimento da mensagem pelo interceptador perca a utilidade ou não comprometa a segurança entre o emissor e o receptor. A criptografia é usada também para autenticar documentos, principalmente aqueles que transitam eletronicamente (FIGURA 2). Segundo ALBERTIN (1998), a criptografia é a mutação de informação em qualquer forma (texto, vídeo ou gráficos) em uma representação não legível por qualquer pessoa sem uma chave de criptografia. Nota-se que só a criptografia não assegura contra as invasões nos sistemas. Para BAPTISTELLA (2000), o uso da criptografia dos dados evita que as informações sejam utilizadas de maneira inadequada. O fator “segurança na transação” utilizando a criptografia, permite que a informação seja “embaralhada” ao ser enviada e “desembaralhada” quando chega ao destino correto.

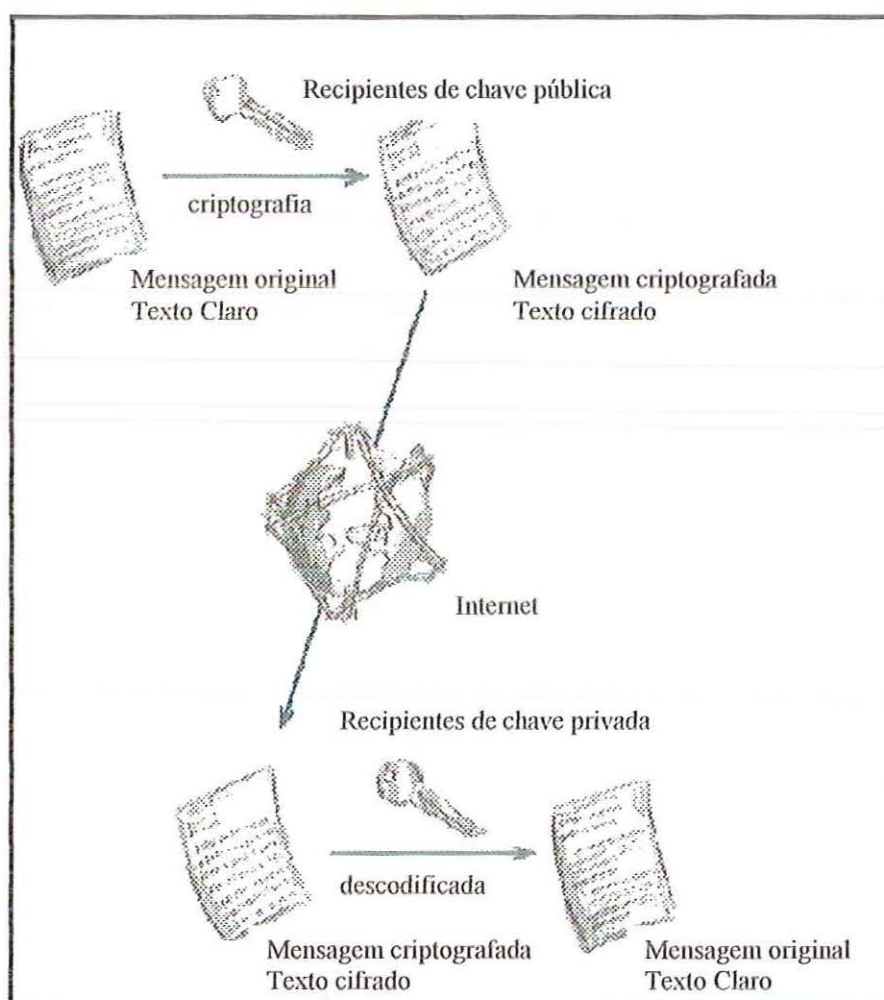


Figura 2: Fluxo de Criptografia
Fonte: LICO, 2001, pg 77

- **PKI – Public Key Infrastructure:** de acordo com ALCOFORADO & SÊMOLA (2000), em uma PKI, certificados digitais são emitidos por uma Autoridade certificadora para cada usuário ou equipamento envolvido. Com estes certificados, são gerados pares de chaves constituídos por uma chave pública e uma privada. Os algoritmos utilizados permitem que aquilo que foi criptografado com a chave pública só possa ser aberto pela chave privada, e vice-versa. Desta forma, tudo que é público é colocado à disposição num diretório onde se tenham acesso às informações de todos, e tudo que é privado é mantido em segredo pelo usuário, quer seja criptografado no disco de seu computador, dentro de um smartcard ou em um servidor da rede. Os certificados digitais contém informações sobre a Autoridade certificadora que os emitiu, e neste ponto se tornam primordiais aspectos de política de confiança.
- **VPN – Virtual Private Network:** a interligação via Internet entre as unidades e escritórios distantes das empresas pode ser feita se o usuário estiver conectado na mesma infra-estrutura de acesso, com um simples programa obtido facilmente na Internet, tendo desta forma, acesso a todos os dados que trafegam no meio. Empresas utilizam VPN pois os dados são criptografados antes de trafegarem por canais inseguros e são apenas decodificados quando chegam a seus destinos. Para BASTOS (1998), uma expansão da VPN é o recurso de acesso remoto seguro que permite que um usuário trabalhando em casa, em trânsito ou em viagem, possa acessar a rede interna da empresa de forma segura através da Internet.
- **SET – Secure Electronic Transaction:** hoje a Internet é vista como uma grande “vitrine eletrônica”, possuindo grandes vantagens do mercado tradicional principalmente por independer de horários comerciais para a realização de compras. Mesmo com *sites* atraentes e cativantes que induzem os consumidores às compras, o comércio eletrônico ainda não chegou no seu potencial. Isto ocorre devido as desconfianças e medo de realizarem transações fornecendo número e senha de cartão de crédito. Não é de se estranhar com tantas invasões e histórias até sensacionalistas que relatam sobre clonagens e sabotagens de cartões de crédito. Uma parceria realizada entre as empresas Visa e a Mastercard deram início ao desenvolvimento de um padrão único, simples e seguro de transação eletrônica. Esse padrão foi denominado SET, tornando-se o padrão oficial para

as transações eletrônicas via Internet. Segundo MORENO (1998), o padrão SET faz uso de criptografia de chave pública e secreta, garantindo a segurança numa transação eletrônica não-presencial através de verificação de assinaturas digitais e confidencialidade de conteúdo. Para o autor, a batalha mais árdua é convencer o usuário a acreditar na segurança do sistema. Neste caso o maior problema é o cultural.

- **Firewall:** ALBERTIN(1998a), definiu que firewall (FIGURA 3) é um método de proteção que visa colocar equipamentos, um computador ou um roteador, entre a rede e a Internet, a fim de controlar e monitorar todo o tráfego entre o mundo externo e a rede local. Tipicamente, o equipamento permite que os usuários internos da organização tenham acesso total a serviços do lado externo, enquanto fornece acesso através de identificação, senha, endereço de IP e outros identificadores. Firewall possui grande facilidade de uso, mas exige ainda avançados conhecimentos técnicos em sua configuração. Protegem o acesso através de um servidor de controle no ponto único de entrada e saída dos dados.

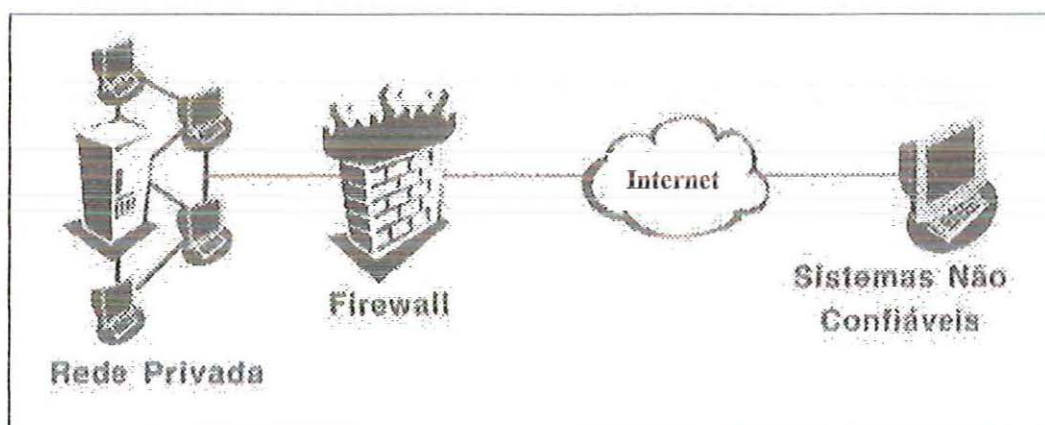


Figura 3: Exemplo de um Firewall

Fonte: <http://www.nai.com>

- **Assinatura Digital:** quando se lida com documentos digitais recomenda-se o uso de assinaturas digitais. As assinaturas digitais são parecidas com as assinaturas manuscritas para documentos impressos. Elas garantem que o documento seja autêntico e íntegro. Previne também que o documento não seja repudiado, ou seja, o assinante não pode alegar posteriormente que não assinou. De acordo com ALBERTIN (1998) a assinatura digital é a prova de autoria e intenção na

forma digital. Há também o uso da criptografia na assinatura digital para que outra pessoa não utilize uma assinatura que não lhe pertence.

Apesar do comércio eletrônico apresentar redução do custo de venda, flexibilidade e velocidade para responder aos estímulos do mercado com promoções instantâneas, os consumidores continuam na “retaguarda”, pois consideram inicialmente que o conteúdo de suas informações serão passadas para um “estranho”. Enquanto houver uma percepção de que os riscos são existentes, o comércio eletrônico não alcançará todo seu potencial.

Portanto se a Internet continuar expandindo o comércio eletrônico, as empresas que buscam realizar seus negócios através dela, devem encarar a segurança com uma óptica diferente da convencional, pois a realidade nos negócios eletrônicos é bem diferente dos negócios tradicionais.

Medidas diversas de segurança são adotadas, para “cativarem” a confiança do consumidor. Mas, embora o cliente esteja seguro em realizar transações comerciais eletrônicas, isso não significa exatamente que esse cliente será fiel com o mesmo *site* de comércio eletrônico. A partir de então surge uma outra preocupação tão importante quanto a questão da segurança dos dados dos clientes: como se relacionar, conhecer, fidelizar e reter clientes “seguros” neste cenário ? Assunto este que será abordado no próximo item.

2.8 RELACIONAMENTO COM O CLIENTE NO COMÉRCIO ELETRÔNICO

O comércio eletrônico trouxe uma fonte rica de informações a respeito dos clientes – informações essas que muitas empresas estão transformando-as em diferencial competitivo.

De acordo com KIMBALL & MERZ (2000), a Web nos presenteia com uma nova e jamais vista fonte de dados: a seqüência de cliques. Essa seqüência de cliques é literalmente um *log* de cada gesto efetuado por cada visitante a cada *site* da Web. A seqüência de cliques é potencialmente um registro muito melhor de comportamento do que outras fontes detalhadas de dados mais tradicionais.

É desta forma que os servidores do comércio eletrônico captam informações que traduzem em comportamento humano. Empresas que atuam nesse cenário estão utilizando arquivos *cookies* para gravar informações a respeito dos clientes quando eles “navegam” nos *sites*.

Para ALBERTIN (1998a), as aplicações de comércio eletrônico podem ter um valor significativo como uma alavanca para novas estratégias de gerenciamento de clientes, pois:

- Conectam diretamente compradores e vendedores;

- Apoiam a troca de informações totalmente digitais entre eles;
- Eliminam os limites de tempo e lugar;
- Apoiam a interatividade e então podem adaptar-se dinamicamente ao comportamento do cliente;
- Podem ser atualizados em tempo real.

De acordo com IBM (1999), para as organizações que operam no comércio eletrônico servirem melhor seus clientes elas devem passar por um ciclo, onde cada ciclo termina e inicia-se outro, fazendo com que a organização nunca pare, estando sempre em movimento, conforme mostrado na **FIGURA 4**.

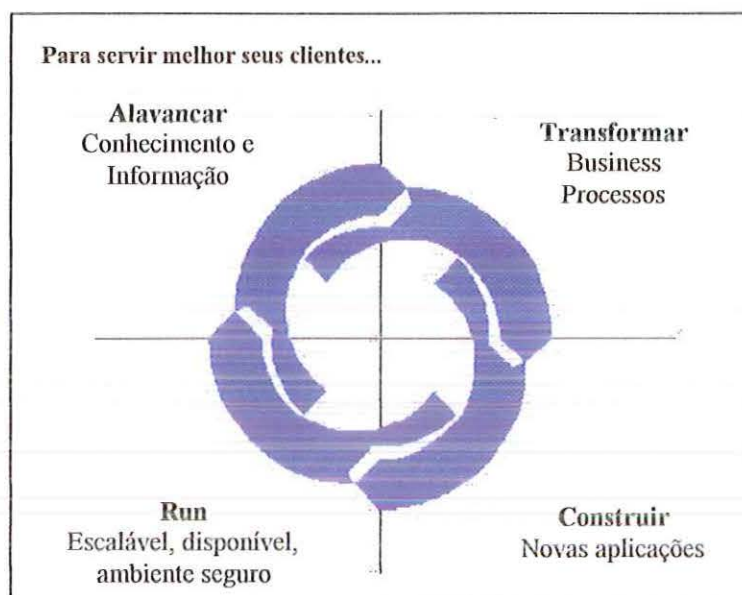


Figura 4: Ciclo de como servir melhor os clientes
Fonte: www.ibm.com

Segundo PEPPERS & ROGERS (2000a), em vez de apenas agir como outro meio para que as empresas possam comunicar-se com os clientes, os *sites* tornar-se-ão cada vez mais a própria empresa, pois manterão os relacionamentos com os clientes. As empresas melhores equipadas para criar relacionamentos de aprendizado (relacionamentos com os clientes), terão a possibilidade de conhecer melhor seus clientes, ou seja, os “mais valiosos na categoria”. Nesta perspectiva, farão o que seja mais conveniente para os clientes continuarem a usar os serviços de sua empresa do que buscar novos relacionamentos.

De acordo com BRETZKE(2000), o ambiente da Web traz contribuições para a viabilização o *marketing* de relacionamento, tais como:

- Canal de relacionamento ampliado: a Internet tem alterado significativamente o relacionamento com clientes e fornecedores, eliminando intermediários e facilitando o acesso de informações, que transforma o processo decisório de compra;
- Inovação e customização de produtos: como a Internet melhora a coleta de informações sobre as necessidades e preferências dos clientes, é possível através da estratégia de CRM desenvolver novos produtos ou adaptar os existentes para que sejam customizados de acordo com as mudanças de necessidades e expectativas de grupos de clientes.

Nota-se, que o relacionamento com o cliente neste cenário tornará estreito se for bem gerenciado e constantemente atualizados, através de uma tecnologia adequada.

2.9 CONSIDERAÇÕES FINAIS DO CAPÍTULO

A Internet tornou-se a base sobre a qual possibilita radicalmente as formas de se negociar. A importância da inserção dos negócios no comércio eletrônico trouxe grandes vantagens para as empresas de pequeno porte – tanto essas como as empresas de médio e grande porte podem divulgar seus produtos e serviços, operando no mesmo mercado, o que era difícil no comércio tradicional.

A inserção do comércio na Internet levou às empresas que tentam neste cenário, rever o conceito de comércio para assegurar a sua competitividade. Muitas das empresas já estão deixando de ser “seduzidas” pela idéia de poder inserir seus produtos e serviços de forma globalizada por um custo baixo, tornando sua marca conhecida de forma virtual, para adotar estratégias que possivelmente garantirão a sobrevivência delas nesse cenário.

Algumas das estratégias adotadas são as medidas de segurança, as quais permitem assegurar aos clientes que realizam suas compras em um *site* de comércio eletrônico. Tais medidas permitem que os clientes realizem transações econômicas seguras, ou seja, forneçam números de cartões de crédito sem que haja futuras complicações.

Não somente a segurança dos dados, mas também o relacionamento com os clientes é de devida importância para garantir o retorno de um cliente ao *site* de comércio eletrônico. O relacionamento com os clientes neste cenário é feito sob o apoio de tecnologia de informação, uma vez que aqui não há contato físico. Todo o “caminho percorrido” pelo cliente dentro de um *site* é armazenado em banco de dados, que “traduzem” o perfil de cada

um. As empresas que conseguirem melhor manipular os dados referentes ao comportamento dos clientes, certamente ganharão uma vantagem estratégica diferencial.

O comércio eletrônico tem trazido contribuições para diversas áreas. Além das inúmeras oportunidades de vendas deste novo canal de vendas, os fornecedores passam a ter contato direto com seus clientes, proporcionando desta forma oportunidades de desafios que eles enfrentam, como:

- As empresas que tiverem excelência no atendimento das necessidades e expectativas de seus clientes, ganharão vantagem competitiva, caso contrário, seus clientes as trocarão com apenas um “clique de mouse”;
- Os clientes poderão comprar diretamente de seus fornecedores, sem intermediários, estreitando o relacionamento entre eles e proporcionando aos fornecedores um melhor conhecimento dos hábitos de seus clientes;
- Conhecer os clientes, seus hábitos, quais são os mais rentáveis, os que compram mais, como compram, etc., será possível se as empresas investirem em tecnologia de informação, ou seja, possuir um banco de dados capaz de armazenar um volume histórico das transações dos clientes, as quais poderão ser processadas na busca de alguma informação.

Portanto isso mostra que a realidade do comércio eletrônico sob a forma de B2C, proporcionará um cenário novo para empregar o *Marketing*.

Desenvolver medidas de segurança e “conhecer” o perfil do cliente faz com que essas contribuições se tornem cada vez mais evidentes. Mas isso não quer dizer que esses mesmos clientes serão fieis às empresas de comércio eletrônico. Para eles “basta um clique” para desistir de uma compra e/ou mudar para outro *site*.

Para solucionar o problema ligado a retenção de clientes no comércio eletrônico, as empresas têm adotado a filosofia de CRM: uma estratégia de negócio da área de *marketing* capaz de propor os caminhos de como aproveitar as informações a respeito dos clientes para converter em informações capazes de indicar o perfil dos mesmos – conhecendo-os, as empresas que operam no comércio eletrônico poderão fornecer serviços e produtos personalizados a seus principais clientes. De acordo com CHEIN & CABRERA (2000), CRM é uma estratégia de negócio cuja idéia central consiste em trabalhar com o cliente e não somente para o cliente. O CRM consiste em garantir que todas as informações existentes sobre os clientes dentro da empresa estejam concentradas numa única base de dados para permitir que a empresa ofereça um tratamento altamente personalizado a cada um de seus consumidores. Assunto este que será explorado no próximo capítulo.

3. CRM – CUSTOMER RELATIONSHIP MANAGEMENT

3.1 CONSIDERAÇÕES INICIAIS

As empresas estão percebendo a importância do relacionamento com seus clientes para garantir a lealdade destes num mercado cada vez mais competitivo. Grandes oportunidades aguardam as empresas que conhecem bem os hábitos dos seus clientes. O conceito CRM é aqui inserido como estratégia de negócio para garantir a fidelidade e retenção dos clientes, principalmente no comércio eletrônico.

O objetivo deste capítulo é descrever como era realizado o relacionamento com os clientes antes do CRM; logo após, o CRM é apresentado como estratégia de negócio - os principais conceitos do CRM e a necessidade de se desenvolver o relacionamento personalizado é apresentado neste item. As quatro estratégias do CRM são mencionadas e finalmente é apresentada a importância do CRM em um novo contexto: o comércio eletrônico. Será dado enfoque maior no CRM analítico devido ao cenário proposto pelo trabalho.

3.2 RELACIONAMENTO COM OS CLIENTES ANTES DO CRM

Inicialmente, ao se falar em “Relacionamento com os Clientes” imagina-se conceitos e estratégias de *marketing* que lidam diretamente com o cliente. Para se entender o que é *marketing* alguns conceitos são abordados por diferentes autores, sendo:

“O marketing é uma orientação da administração que visa proporcionar a satisfação do cliente e bem-estar do consumidor, como forma de satisfazer os objetivos e responsabilidades da organização” (KOTLER & ZALTAMANN, 1971, p.12)

“O conceito de marketing significa que uma organização dirige todos os seus esforços no sentido de satisfazer seus clientes e obter lucro”(MCKARTHY & WILLIAMS, 1982, p.6).

“Marketing é uma filosofia empresarial que consiste na integração e coordenação de todas as funções de marketing, as quais por sua vez, estão ligadas a outras funções da empresa, com o objetivo básico de obter o máximo lucro a longo prazo” (FELTON, 1959, p65).

KOLISH & JAWORSKI (1990)⁶ apud BRETZKE (1992), descrevem que o tema central do conceito de *marketing* é a orientação para o cliente ou consumidor, mas se apoia em alguns paradigmas principais que atuam como suporte para guiar a implementação. São eles:

- *Customer Focus*: envolve o conhecimento profundo das necessidades, desejos e preferências dos clientes e consumidores, envolvendo um processo de *Market Intelligence*, e não simples verbalização desses desejos e necessidades;
- *Marketing Coordenado*: indica que a orientação para o mercado é responsabilidade de toda a empresa, que deve conhecer as necessidades dos clientes e responder à elas;
- *Rentabilidade*: é visto como uma consequência da orientação para o mercado.

Entendendo alguns dos principais conceitos sobre *marketing* torna-se viável abranger a forma de como as organizações se relacionavam com seus clientes antes do surgimento de alguma estratégia de *marketing* específica.

Nos primórdios do comércio, os comerciantes conheciam seus clientes pelos nomes, suas residências, que tipo de produtos e quando eles necessitavam. Além disso, sabiam como seus clientes podiam gastar e como eles pagariam.

Para PEPPERS & ROGERS (2000), antigamente os clientes eram divididos em grupos de maior ou menor valor para um certo negócio. Conscientemente ou não, esses mesmos clientes também eram divididos com base em suas necessidades e desejos.

Desta forma, produtos e serviços eram personalizados para que se adequassem às necessidades e ao valor de cada um dos clientes. Todo esse processo era desenvolvido com base no conhecimento do cliente a fim de torná-los fiéis às empresas.

Os negócios foram feitos dessa forma por muito tempo. A partir da inserção da tecnologia, começa a ser permitido às empresas a produção em massa, para mercados de massa, ou seja, a técnica de vender para clientes de forma individual passou a ceder lugar à participação no mercado.

Conhecer clientes, seus hábitos e necessidades era feito por meio de pesquisas de mercado, segmentações e modelos estatísticos (PEPPERS & ROGERS, 2000a).

Para PEPPERS & ROGERS (2000), a produção em massa criou uma grande riqueza e trouxe níveis nunca antes ocorridos de prosperidade em muitas regiões do mundo. Mas houve consequências: a competição global fez com que muitas empresas baixassem seus

⁶ KOLISH, A. K.; JAWORSKI, B. J. (1990). Market Orientation: The Construct, Research Propositions and Managerial Implications, *Journal of Marketing*. Vol 54 April, p.1-18.

preços para conquistarem clientes. Essa medida poderia beneficiar algumas empresas a curto prazo, mas ao longo do tempo o resultado é a diminuição da lucratividade.

Essa abordagem fez com que as empresas:

- deixassem de vender para os clientes, para vender para as empresas;
- buscassem participação no mercado ao invés dos clientes.

A principal filosofia empresarial era focada no produto e não no cliente. A partir de então surge como alternativa para esse contexto a Tecnologia de Informação - capaz de trazer de volta as estratégias usadas nos primórdios do comércio, ou seja, o foco no cliente.

SILVA (2000) afirma que o papel desempenhado pelo *marketing* vem mudando ultimamente, conforme é ilustrado na TABELA 3.

DE...	PARA...
Conquista de clientes	Manutenção dos clientes
Lucro em cada transação	Gerenciamento do valor do tempo de vida do cliente
Vender produto	Compreender e atender as necessidades dos clientes
Planejamento de cada ferramenta de comunicação	Integração das ferramentas de comunicação

Tabela 3: Mudanças no papel do *Marketing*
Fonte: SILVA, 2000, pg 13-14

Para SILVA (2000), devido as essas mudanças as empresas necessitam cada vez mais:

- Conhecer seus clientes, como forma de tornar seus produtos, condições e serviços mais flexíveis;
- Identificar os clientes mais lucrativos, a fim de estabelecer diferentes níveis de atendimento;
- “Inventar” novas maneiras de criar e transmitir valor a seus clientes alvos, como forma de prestar melhores serviços e corresponder à crescente sofisticação de seus clientes;
- Utilizar melhor os canais de distribuição, tornando o contato com o cliente mais efetivo e econômico.

De acordo com KOTLER (1999) o *marketing* deve evoluir nos próximos, ou seja, as organizações devem alinhar suas atividades de *marketing* para:

- Conhecer os clientes: a maioria das empresas terá bancos de dados privados contendo informações sobre as preferências, necessidades e exigências de seus

clientes. Produtos poderão ter suas especificações determinadas on-line, e encomendado diretamente à empresa eletronicamente;

- Carteira de clientes: mais do que ter participação no mercado, as empresas tenderão a desenvolver a idéia de carteira de clientes, onde o alvo serão aqueles clientes que a empresa pode atender mais eficientemente (superando expectativas) e obter uma maior lucratividade;
- Retenção: O sucesso na superação de expectativas dos clientes a partir do melhor conhecimento dos mesmos, implicará em maior retenção de seus clientes. Conquistar novos clientes se tornará cada vez mais difícil.

Diante dos fatos, somente as empresas que “conhecerem” as reais necessidades e expectativas de seus clientes serão capazes de sobreviver em um mercado cada vez mais competitivo.

Para BRETZKE (1992) a busca da vantagem competitiva exige uma nova abordagem estratégica da forma de administrar os negócios e novas técnicas ou ferramentas, para implementar essa nova estratégia. Vislumbra-se nesse contexto, um novo *marketing* impactado pela tecnologia, mais factível e mais apto para responder aos desafios da era do cliente. A utilização dos produtos e serviços gerados pela Tecnologia de Informação não só impactou a forma de administrar a organização e seus processos internos, mas também, o ambiente na qual ela opera: a atuação da concorrência e o estilo de vida, necessidades e preferências dos seus clientes e consumidores.

Segundo MCKENNA (1991)⁷ *apud* BRETZKE (1992), o significado desse novo *marketing* é analisado por causas intrínsecas apresentadas a seguir:

- crescente diversidade de produtos e serviços;
- crescente concorrência global;
- segmentação de mercados, tecnologias e soluções;
- mudanças organizacionais e estruturais nas empresas, buscando novas maneiras de administrar os negócios;
- canais de distribuição que estão em constante movimento. Ainda que estar próximo do cliente seja preponderante, os canais freqüentemente negligenciam o diálogo com o cliente;
- o ambiente de negócio e o cenário futuro da competitividade são imprevisíveis;
- previsões e pesquisas de mercado não fornecem um rumo claro para as ações.

⁷ MCKENNA, R. (1991). *Relationship Marketing: Successful Strategies for the age of the Customer*. New York, Addison-Wesley Publishing Company, Inc. 1st ed.

“O ambiente competitivo exige que as empresas realmente orientem a sua organização para o mercado, adotando novas estratégias de marketing para a era do cliente. A real solução, sem dúvida, não é mais o marketing, mas o melhor marketing. E isto significa encontrar meios de integrar o cliente à empresa; criar e sustentar o relacionamento entre a empresa e o cliente”(BRETZKE, 1992, pg 10).

Como a tecnologia da informação *per si* não traz os resultados necessários para manter os clientes, há necessidade de se implementar a estratégia de negócio, também conhecida como *marketing 1to1* ou *marketing one-to-one* ou até CRM, assunto que será explorado no próximo item.

3.3 CRM COMO ESTRATÉGIA DE NEGÓCIO

No item anterior, foi abordado que para se obter um resultado necessário para manter os clientes, é necessário unir a estratégia CRM à tecnologia de informação.

É importante mencionar que a tecnologia é um meio de implementar as estratégias para atingir os objetivos pretendidos pelo CRM, mas não é ela que constrói a relação com o cliente e sim os profissionais que estão por trás dela.

“CRM é uma estratégia de negócio voltada ao entendimento e à antecipação das necessidades dos clientes atuais e potenciais de uma empresa. Do ponto de vista tecnológico, CRM envolve capturar os dados dos clientes ao longo de toda a empresa, consolidar todos os dados capturados interna e externamente em um banco de dados central, analisar os dados consolidados, distribuir os resultados dessa análise aos vários pontos de contato com o cliente e usar essa informação ao interagir com o cliente através de qualquer ponto de contato com a empresa” (PEPPERS & ROGERS, 2001, pg 44)

Para TRONCHIN (2000), a essência do CRM é tratar clientes diferentes de forma diferente, simplesmente porque os clientes são diferentes. É devido a esse fato que o CRM envolve a diferenciação do cliente, ou seja, terá como base o valor que o cliente possui para a empresa e suas diferentes necessidades em distintas situações. Para o autor, tratando de forma especial os clientes especiais, cria-se barreiras de mudança e aumenta a fidelidade.

Segundo BRETZKE (2000), a tecnologia de informação influencia profundamente a maneira de pensar e fazer negócios, pois possibilita que a empresa implante uma estratégia de gerenciamento do relacionamento do cliente. Isso permite que as empresas reajam ao mercado competitivo, transformando as informações em conhecimento do comportamento de compra do cliente, para reduzir custos, desenvolver novos produtos e agregar serviços.

A empresa que utiliza a estratégia do CRM deve utilizar intensamente a tecnologia de informação, integrando hardware e software adequados para focar em seu principal fundamento: o cliente.

De acordo com SANTANA (2001), CRM é uma estratégia de negócio voltada ao atendimento e antecipação das necessidades dos clientes atuais e potenciais. Essa estratégia envolve tecnologia, pessoas, processos e ambientes. Para que uma solução CRM dê retorno esperado, é preciso utilizar todas as informações geradas nos contatos cliente/empresa. Para o autor, o CRM pode ser dividido em:

1. Operacional: ocorre através do contato direto da organização com o cliente (mala direta, Internet, *call centers*), ou seja, mantém contato com o cliente.
2. Analítico: é realizado por meio de dados contidos nas bases gerenciais da organização, ou seja, tem a função de analisar os dados minerados dos diversos pontos da empresa, destinados às necessidades da gerência no processo de tomada de decisão. Isso pode envolver complexas consultas que necessitam acessar um grande número de registros – portanto, é importante a existência de muitos índices criados para acessar as informações de maneira mais rápida possível. Um DW armazena informações históricas de muitos anos e por isso deve ter uma grande capacidade de processamento e armazenamento dos dados que se encontram de duas maneiras: detalhados ou resumidos. A função do CRM analítico é determinar quais são os clientes que devem ser tratados de forma personalizada (*one to one*) efetuando a diferenciação ao invés da segmentação e quais devem ser deslocados para níveis de prioridade inferior.
3. Colaborativo: estabelece um valor vital com os consumidores – relação de parceria, serviços colaborativos como edição personalizada, e-mail, comunidades, conferências e centros e interação de cliente capacitado à Web.

De acordo com a FIGURA 5 pode-se ter uma noção melhor do que foi mencionado anteriormente.

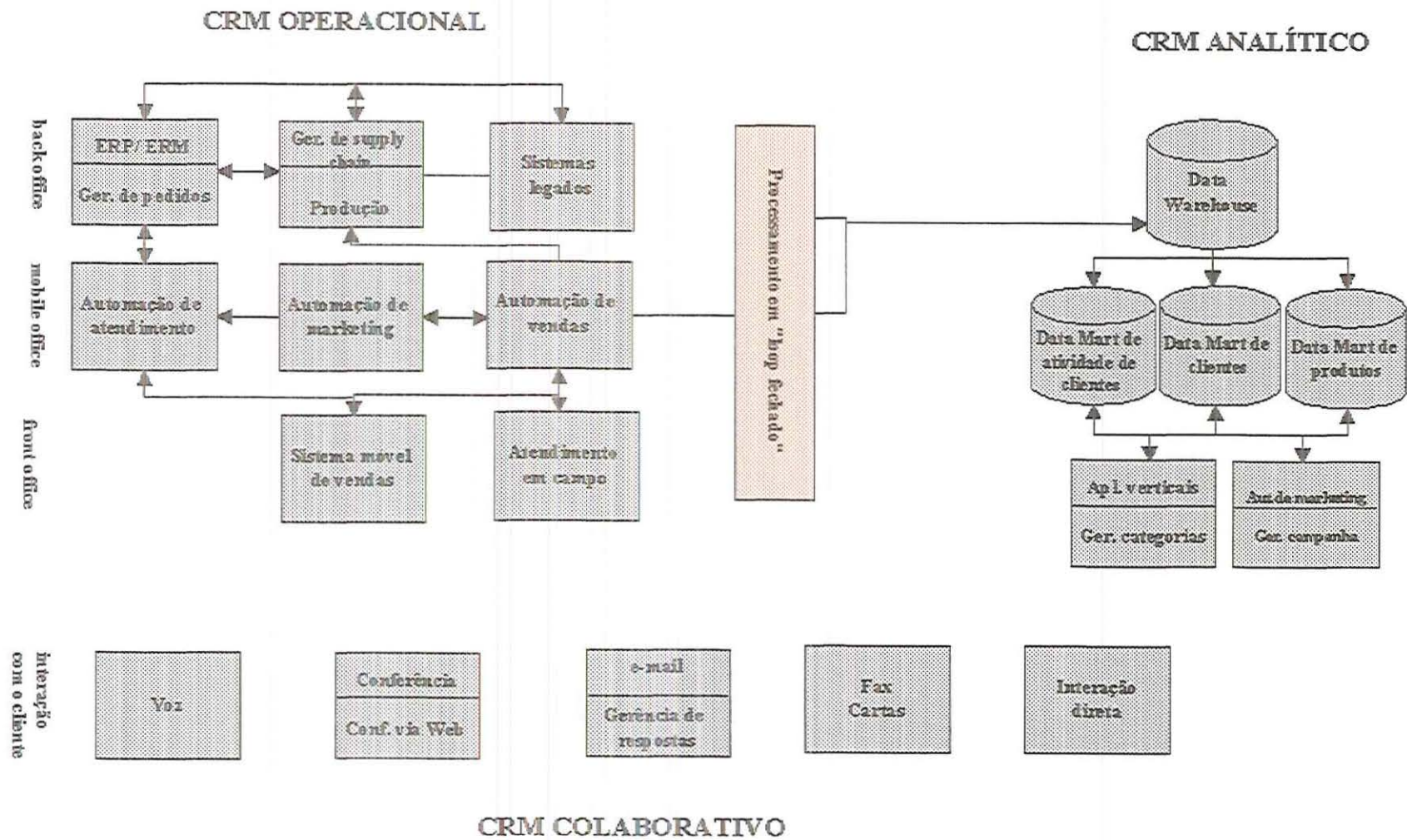


Figura 5: Visão geral do CRM
 Fonte PEPPERS & ROGERS, 2001, pg 51

Enquanto o *marketing* de massa visa aumentar a base de clientes, concentrando-se em um produto ou serviço focado em uma necessidade específica, o *marketing one-to-one* tem como objetivo aumentar o número de necessidades atendidas de um grupo específico de clientes: os clientes de maior valor e maior potencial, conforme ilustra FIGURA 6 (TRONCHIN, 2000a).

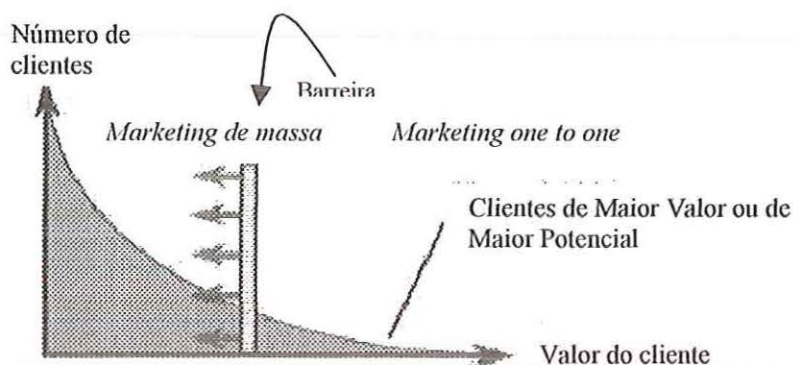


Figura 6: Estratégia da barreira entre *marketing* de massa e *marketing one to one*
 Fonte: PEPPERS & ROGERS, 2001 pg 21

De acordo com BRETZKE (1992), o momento que está se vivendo hoje é a Era do Cliente, pois há uma força vinda do consumidor que está mudando o mercado. Os consumidores querem voltar ao *status* cliente, no sentido de serem tratados como indivíduos, com suas necessidades entendidas e atendidas.

Para TAURION (2001), o CRM é uma evolução do *marketing* de massa para um *marketing* mais personalizado, buscando criar e manter relacionamentos com os clientes.

Personalização se cria com aprendizado, com um processo interativo e mútuo de investimentos na troca de informações. (TAURION, 2000a).

De acordo com IBRAHIM (2000), o CRM é uma estratégia integrada e integral de negócios que possibilita às empresas gerenciar com eficiência os relacionamentos históricos com seus clientes.

Muitos autores descrevem o que de fato é o CRM. Utilizando termos mais ou menos complexos e semelhantes. O que todos dizem em suma é que o CRM é uma estratégia de negócio direcionada ao entendimento antecipado das necessidades dos clientes potenciais de uma empresa. E para que a empresa ofereça um tratamento altamente personalizado a cada um desses clientes, deve-se fazer uso da tecnologia de informação, ou seja, garantir que todas as informações existentes sobre eles estejam concentradas em uma única base de dados - desta forma otimizarão seus investimentos, ganhos e a satisfação de seus consumidores. A

fidelização de clientes é construída pelo relacionamento constante, diferenciado e personalizado com os clientes.

Don Peppers pioneiro e criador do conceito *marketing one-to-one*, juntamente com Martha Rogers, , definem em PEPPERS & ROGERS (2001) quatro estratégias para o CRM. São elas:

1. IDENTIFICAR

O primeiro passo para iniciar uma relação *one-to-one* com os clientes é identificar cada um deles. Deve-se identificar todos para saber qual vale a pena iniciar uma relação *one-to-one*. Para os autores, identificar o cliente significa conhecer sua identidade, forma de contato preferida, todas as transações e interações realizadas com a empresa, todas as reclamações feitas e quais foram as providências tomadas.

O maior desafio é que todos esses dados devem estar disponíveis em todos os pontos de contato do cliente. Isso quer dizer que é a empresa que tem que identificar cada um de seus clientes e não um departamento ou funcionário. Além disso, o cliente tem que ser identificado independentemente do canal escolhido para entrar em contato com a empresa. A identificação do cliente pela empresa é prioridade no estabelecimento de iniciativas *one-to-one*. Sem identificar o cliente é impossível conhecê-lo. A tecnologia pode exercer um papel importante nesse sentido, e as novas ferramentas que estão surgindo tornam cada vez mais fácil incluir e consultar registros; mas para isso deve-se manter as bases de dados atualizadas.

2. DIFERENCIAR

Os clientes são diferentes e como já foi mencionado, ignorar ou desconhecer essas diferenças não os fazem iguais. Os clientes são diferentes em seus valores e em suas necessidades para a empresa. O objetivo da diferenciação de clientes é encontrar os Clientes de Maior Valor (CMV) e os Clientes de Maior Potencial (CMP). Entende-se por CMV aqueles clientes com o valor real mais elevado para a empresa – os que fazem a maior parte dos negócios, geram as mais altas margens de lucro, são mais predispostos a cooperar e tendem a ser mais fiéis. Os CMVs são aqueles nos quais as empresas tem a mais alta participação. Já os CMPs são aqueles nos quais o valor estratégico – que é o valor potencial do cliente – excede acima do valor real. São clientes que possuem o maior potencial de

crescimento, o que pode ser alcançado por meio de *cross selling*⁸, por meio da manutenção do cliente por um longo período de tempo ou talvez pela alteração do comportamento do cliente, fazendo operar de uma maneira que o custo da empresa seja menor. Assim, com esses clientes, pode-se desenvolver a relação de aprendizado.

- Diferenciação por valor: é o valor que o cliente tem em toda sua história transacional com a empresa, as referências que ele fez e que se transformaram em vendas. Aqui o valor é expresso em termos de lucratividade. Há também a diferenciação por potencial.
- Diferenciação por necessidades: forma de diferenciar os clientes através de suas necessidades. Quanto maior o número de necessidades conhecidas, maior a oportunidade de participar no cliente.
- Estratificar os clientes: deve-se implementar programas de retenção ou seja, programas de reconhecimento e de possibilidade de utilização de canais de comunicação exclusivos são formas de recompensar os melhores clientes. Porém, é através da relação de aprendizado e da personalização que torna-se cada vez mais forte a relação com esses clientes. Somente assim garante-se a retenção desses clientes.

3. INTERAGIR

Uma vez descobertos os CMV e os CMP o próximo passo é incentivá-los a interagir com as empresas. A interação é a única forma de conhecer cada vez mais os clientes. Com eles, busca-se como objetivo final desenvolver uma relação de aprendizado, onde cada vez mais personaliza-se serviços e produtos para que o cliente veja conveniência em continuar com uma certa empresa. Há uma relação íntima entre a interação e a diferenciação. Na realidade essas duas estratégias caminham juntas, pois uma depende da outra.

As interações iniciadas pelo cliente têm um potencial muito maior que as interações iniciadas pela empresa. Os autores descrevem que essa é uma excelente oportunidade para conhecer melhor o cliente pois cada nova pergunta respondida é conhecimento acumulado a seu respeito, e os funcionários que mantêm o contato devem ser treinados e incentivados a conhecê-lo melhor e a registrar esse conhecimento nas bases de dados de clientes da empresa.

⁸ Nota da autora: *cross selling* (venda cruzada) consiste em vender mercadorias e serviços relacionados entre si para um determinado cliente. Esse processo é somente uma das formas de aumentar a participação no cliente.

Por outro lado, quando a interação é iniciada pela empresa, deve-se obedecer a duas regras básicas: deve-se utilizar o canal de comunicação preferido do cliente e deve-se ter a permissão do cliente para utilizar esse canal.

4. PERSONALIZAR

A partir do conhecimento das necessidades dos clientes, deve-se personalizar os serviços e produtos para fazer com que a experiência de realizar negócios com a empresa seja única. Quanto mais personaliza-se mais valor entrega-se ao cliente, que vê uma certa conveniência em continuar com a empresa.

A personalização é algo bem simples quando se conhece as necessidades e as preferências do cliente, mas exige muita flexibilidade da empresa e treinamento adequado das pessoas que têm contato com ele. Elas devem estar preparadas para mudar seu comportamento de acordo com cada cliente.

Se a empresa possui um *Web site*, a personalização pode ser levada ao extremo, com a disponibilização de conteúdo que tenha a ver exatamente com os gostos e preferências de cada um de seus clientes. Porém, é importante ressaltar que o cliente espera essa personalização em todos os canais de contato com a empresa. Quanto mais bem-sucedida for a personalização, mais simples e conveniente para o cliente é fazer negócio com a empresa (PEPPERS & ROGERS, 2001). Mas vale a pena ressaltar que a estratégia *one to one* não é para todos os clientes, e sim para os melhores. Para os outros clientes, quando for o caso, aplicam-se as práticas conhecidas de *marketing de massa* e de *database marketing*, pois a utilização do *marketing one to one* não invalida ou dispensa outras formas de *marketing*, conforme ilustra a FIGURA 7.

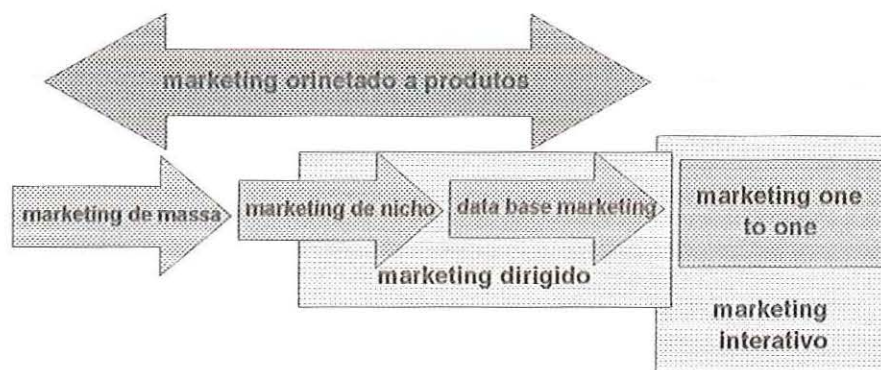


Figura 7: Tipos de Marketing
Fonte: PEPPERS & ROGERS, 2001, pg 59

Deslocando de cenário surgiu a Internet. A aplicação do CRM no comércio eletrônico é uma inovação que engloba todo o relacionamento com o cliente via Internet através de fontes de informação. BRETZKE (1992) afirma que nesta nova era onde o indivíduo ganha nova dimensão, o *marketing* deve buscar novas formas de se comunicar e de se estabelecer um relacionamento duradouro e gratificante com os clientes potenciais, como forma de obter vantagem competitiva sustentável. Assunto que será discutido no próximo item.

3.4 GERENCIAMENTO DO RELACIONAMENTO COM O CLIENTE NO COMÉRCIO ELETRÔNICO

A Web trouxe mudanças em todas as áreas do negócio, principalmente no que se diz respeito ao relacionamento com o cliente. Uma empresa que possui negócios no comércio eletrônico, existe apenas porque o cliente assim o permite.

Com a popularização de *sites* no comércio eletrônico, empresas que possuem clientes insatisfeitos correm sérios riscos de perder negócios para um *site* concorrente. É portanto neste cenário que as empresas devem ser capazes de determinar o público alvo, que na verdade é o consumidor final: o cliente. É devido a ele que tudo se realiza, pois sem ele não há negócio.

Para o cliente serão personalizados os produtos e serviços. No entanto, nem todos os clientes que “navegam” nos *sites* de comércio eletrônico devem ser levados em consideração.

“Os consumidores não são iguais. Há alguns que apenas geram custos para sua empresa. É importante tornar mais leais aqueles clientes que lhe interessam” (ALMEIDA, 1997, pg 61).

As empresas devem ser capazes de identificar o verdadeiro cliente potencial, ou seja, aquele que realmente efetiva compras ou tem potencial para realizá-las. E para se atingir essa meta, inúmeros pesquisadores e executivos da área relatam que a necessidade de se aplicar o CRM no comércio eletrônico é de vital importância para conseguir fidelização e retenção de clientes neste cenário.

De acordo com CHEIN & CASTRO (2000), o imenso volume de informações disponíveis na Internet pode confundir o usuário. Porém, utilizando soluções CRM integrado, pode-se capacitar o acompanhamento à preferência do usuário com base nos cliques e não em respostas e perguntas pré-definidas em um determinado cadastro de algum *site*.

Devido ao desenvolvimento dos Banco de Dados, as empresas podem conhecer o perfil, as necessidades e desejos de seu público alvo, e utilizar estes dados para satisfazer seus clientes de forma individual, num processo denominado “*marketing one-to-one*” que são técnicas para se personalizar totalmente as consultas de clientes potenciais aos *sites* de comércio eletrônico.

Neste cenário altamente competitivo, terão sucesso apenas as empresas que utilizarem as melhores ferramentas disponíveis para se diferenciar, através do CRM e do fornecimento de serviços cada vez mais personalizado aos clientes (WIDEL, 2000).

Para PACE (2000), em época de grande competitividade, deve-se pensar em algumas maneiras eficientes de manter e conquistar novos clientes. E é isto justamente que o CRM propõe: ensinar as pessoas a cultivar seus clientes de modo que elas conquistem aquilo que tanto sonham: a fidelidade de seus compradores. O CRM em um ambiente Web ou CRM eletrônico (e-CRM) engloba todo o relacionamento com o cliente via Internet entre as fontes de informações utilizadas. O que ocorre é o *marketing one-to-one*, ou seja, de um para um.

Segundo IBRAHIM (2000), o violento aumento de concorrência motivado pela abertura dos mercados e pelo “fator Internet” tem mudado e valorizado o conceito de interação com o cliente. As empresas percebem hoje que precisam tratar os clientes de maneira diferenciada, única, com respeito e com memória. A diferenciação por serviços, qualidade e preços vem se tornando menor. Relacionamento, percepção e confiança são os critérios mais relevantes no ato da contratação de uma empresa (para concorrentes em condições iguais de competitividade).

A Internet tem forçado empresas de base e indústrias a se comunicarem com seus clientes diretos e com consumidores finais (antes distantes de sua realidade), valorizando assim, sua marca e produto.

PEPPERS & ROGERS (2000), descrevem que ao se conceber um *site* de comércio eletrônico, deve-se ter em mente que, na realidade, está abrindo um canal extremamente poderoso para a criação e o desenvolvimento de relações com os clientes. Por isso, tem-se chamado essa iniciativa de e-CRM. A convergência de e-CRM e comércio eletrônico parece inevitável. Para os autores, os cinco pontos mais importantes a considerar no desenvolvimento de um *site* de comércio eletrônico competitivo são:

- Não apenas proteger a privacidade dos clientes, mas dizer-lhes como isto será feito: as relações baseiam-se em confiança. Os melhores *sites* da Web esclarecem e expõem suas políticas e privacidades de forma ressaltada na *home page*. Isso assegura que as informações fornecidas pelo cliente somente serão utilizadas para seu benefício incentivando sua colaboração;

- Explicar os motivos e justificar a criação da relação: os melhores *sites* são explícitos e detalham exatamente como as informações fornecidas pelos clientes serão utilizadas e qual será o benefício relacionado. Após os clientes tomarem conhecimento que os dados não serão vendidos ou compartilhados com outras empresas (através da política de privacidade), a próxima etapa é saber por que os mesmos devem fornecer informações pessoais através do *site* e qual será o benefício em fazê-lo;
- Organizar o *site* por necessidades de clientes, não por produto: os melhores *sites* da Web são organizados por necessidades dos clientes e não por produtos. Isso permite que os usuários rapidamente encontrem o que procuram e tenham ofertas consistentes com seus interesses, não com os interesses de venda da empresa;
- Permitir que os clientes tenham controle sobre o conteúdo e sobre seus dados: os melhores *sites* estão bastante avançados em relação ao controle que fornecem a seus usuários com relação a utilização de dados pessoais, conteúdo, escolha de assunto de maior interesse e a respeito dos quais gostariam de receber notificações via e-mail, fax etc;
- Motivar os clientes e incentivar a colaboração: os melhores *sites* incentivam a permanência dos clientes e os desenvolvem ao longo do tempo. A melhor forma de fazê-lo é através da observação do comportamento do cliente no *site*, de suas transações e interações. Isso permite antecipar as necessidades dos clientes e apresentar ofertas, informações e recomendações personalizadas, consequência de uma “relação de aprendizado”. Os melhores *sites* armazenam dados que permitem a antecipação da necessidade.

LICO (2001) observa que o valor do CRM pode ser mostrado através de um gráfico XY, onde o eixo X representa a Lealdade do cliente e o eixo Y a sua Satisfação (FIGURA 8). Quanto maior for a satisfação do cliente, via personalização de produtos e serviços baseado nas necessidades e expectativas do cliente, provendo um excepcional suporte quando algum serviço ou produto não for de encontro a isso e outras ações que façam com que o cliente confie na organização, maior será sua Lealdade.

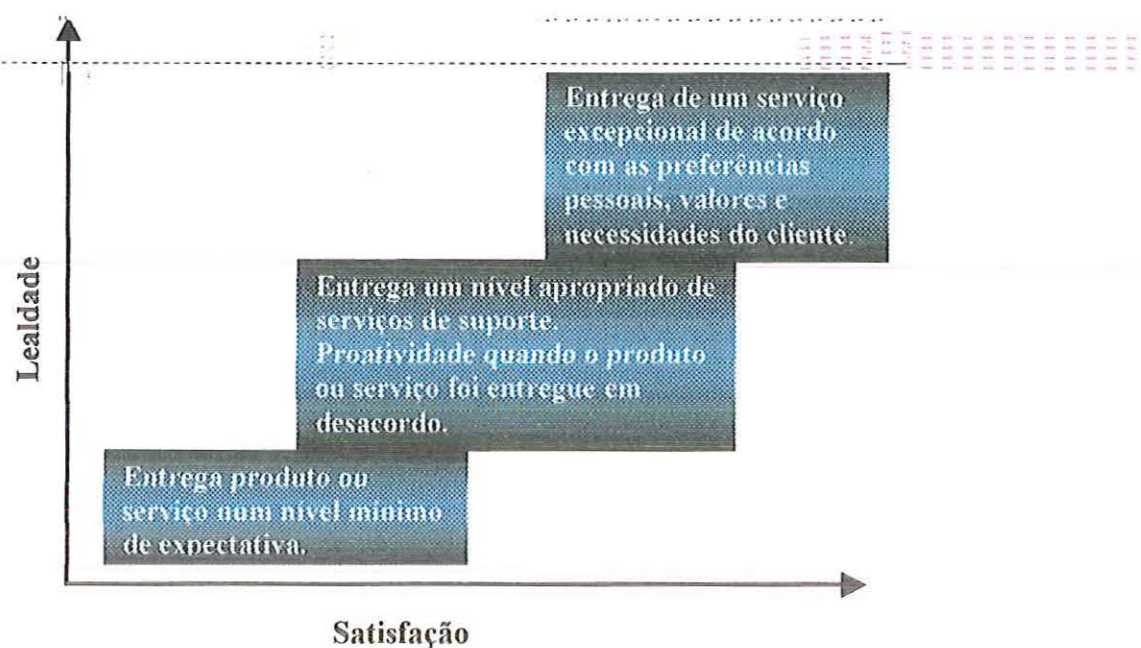


Figura 8: Valor do CRM
 Fonte: LICO, 2001, pg 66

LICO (2001) ainda explana que para manter a Lealdade do cliente via a sua Satisfação em relação ao *site*, deve-se ter um processo. Esse processo é um ciclo dividido em três partes, sendo que o foco é o cliente (FIGURA 9):

Acquisition: Nessa fase tenta-se atrair para o *site* não qualquer cliente, mas o cliente certo.

Development: Entrega ao cliente o que ele necessita, começando aqui uma relação de confiança com o *site*.

Retention: É a fase mais difícil e importante. É aqui que se constrói e mantém a lealdade do cliente, entregando a ele não só mais do que ele necessita num nível básico, mas sim, criando mais facilidades, promoções e antecipando suas necessidades.



Figura 9: Processo CRM
Fonte: LICO, 2001, pg 67

O CRM baseado na Web e, portanto, baseado no cliente, utiliza a Web para integrar e simplificar os processos de negócio de relacionamento com clientes (EDELHERTZ, 2000).

Contudo, as empresas que operam no comércio eletrônico devem estar cientes que o CRM é mais do que uma filosofia, é uma estratégia, devendo tomá-la como um diferencial competitivo.

3.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Inúmeros autores apresentaram neste capítulo os conceitos e opiniões a respeito de CRM. Nota-se que a preocupação de todos os autores é garantir a fidelidade dos principais clientes em uma determinada empresa: tanto empresa que opera no comércio tradicional como a empresa que operam no comércio eletrônico.

O CRM possui estratégias que auxiliam os analistas de *marketing* nas tomadas de decisões a respeito dos clientes. Manter um cliente fiel é uma das metas que toda a empresa deseja obter.

As estratégias do CRM apresentadas evidenciam o valor que representa cada cliente: como atraí-lo e mantê-lo fiel, identificação dos potenciais e não potenciais, suas preferências, a interação deles com a empresa e o melhor momento para oferecer produtos e serviços personalizados.

No comércio eletrônico essa filosofia é de grande importância. Os clientes que realizam compras pela Internet não possuem uma certa “obrigação” de comprar em um único

site. Não havendo deslocamento físico, os clientes estão diante de apenas “um clique” de mouse para não efetuar uma compra ou até mesmo mudar de empresa.

Devido essa problemática o conceito foi adaptado para este novo cenário. Autores mencionados, mostraram em suma que o CRM no comércio eletrônico, ou e-CRM, tem se tornado uma estratégia indispensável para as empresas que desejam fidelizar seus principais clientes.

Mas possuir um cenário e uma estratégia não é o suficiente para as empresas. A necessidade da inserção da tecnologia de informação é de grande importância para que a empresa que opera neste cenário seja capaz de resolver uma de suas problemáticas: a retenção de seus clientes. Como proposta, o DW surge como uma ferramenta que ajuda a solucionar tal problemática no comércio eletrônico. Assunto que será abordado no próximo capítulo.

4. DATA WAREHOUSE: PRINCIPAIS ASPECTOS

4.1 CONSIDERAÇÕES INICIAIS

Com o aumento das bases de dados, verifica-se uma dificuldade notável em obter informações úteis, que auxiliem no processo de tomada de decisões gerenciais. Muitas vezes os dados encontram-se em bases heterogêneas e espalhadas pela empresa, gerando apenas resultados isolados.

A necessidade de integrar estes dados de forma que eles possam auxiliar os administradores de empresas a tomarem decisões estratégicas rapidamente e com segurança, tem se tornado uma necessidade cada vez mais considerável.

A tecnologia de informação tem sido reconhecida como de grande importância para garantir a “sobrevivência” no mercado globalizado. Analistas de negócios utilizam-na para o auxílio da tomada de decisão, a qual deve ser precisa e rápida.

A ferramenta DW surge nesse contexto. Uma tecnologia capaz de auxiliar os tomadores de decisões através da integração de dados distribuídos em banco de dados operacionais. Com o estudo do DW, observa-se que ele vem atender essa necessidade tão grande nos dias atuais – oferece recursos para suprir as informações gerenciais necessárias para estabelecer um maior controle.

DWs mostram-se muito apropriados para empresas que possuem grandes volumes de dados gerados e acumulados durante sua existência. Sua aplicação se dará quando surgir a necessidade de extrair informações para uma análise gerencial, que, por meio de ferramentas adequadas, disponibilizará para o usuário final realizar suas consultas e análises.

O objetivo deste capítulo é apresentar os principais conceitos de DW e as suas principais ferramentas e características. É apresentada a modelagem de seus dados, comparando-o com o modelo relacional. Por possuir grande importância, o conceito de Metadados é explicado. Logo, é apresentada a arquitetura de um DW e o conceito de Data Mart. O capítulo tem também como objetivo explicar como o DW se insere em um novo cenário: o comércio eletrônico e, finalmente apresenta duas ferramentas de mineração e análise de dados.

4.2 DATA WAREHOUSE - O SURGIMENTO DO CONCEITO

Nos primórdios da década de 80 as empresas, de pequeno a grande porte, passaram a investir em sistemas computacionais, direcionando à criação de sistemas para captura, processamento e armazenamento de dados operacionais da empresa. Estes sistemas



operacionais, tais como, os sistemas de contabilidade, de folha de pagamento e de controle de estoque, ajudam as empresas a exercer o seu dia a dia, mas não são projetados para gerar e armazenar informações estratégicas. O tipo de processamento destes sistemas é conhecido com OLTP (*Online Transaction Processing*), onde um tempo de resposta pequeno é essencial.

Já na década de 90, o aumento da competitividade empresarial e a globalização evidenciaram um notável crescimento das bases de dados. As empresas se deparam com uma dificuldade intensa em obter informações úteis, principalmente, que auxiliem no processo de tomada de decisões. Apesar dos sistemas operacionais produzirem diariamente milhares de *bytes* de dados sobre a empresa, seu negócio e seus clientes; estes dados estavam indisponíveis, como informações estratégicas, para os executivos e tomadores de decisão. Muitas vezes os dados encontram-se em bases heterogêneas e espalhados pela empresa, dificultando uma análise integrada destes, prejudicando desta forma o processo de tomada de decisão. A utilização da tecnologia da informação como solução para este tipo de problema tem se tornado cada vez mais solicitada.

Com a evolução da tecnologia da informação e o crescimento do uso de computadores interconectados, empresas de médio a grande porte têm utilizado os sistemas informatizados para efetuar seus processos mais importantes, o que resulta em grandes quantidades de dados referentes ao negócio, dados estes não interligados. Estes dados muitas vezes armazenados em um ou mais bancos de dados operacionais não servem em seu estado original como um recurso estratégico, tornando-os desta forma, vagos e sem valor para o processo de tomada de decisões; decisões estas que normalmente são tomadas com base nas experiências dos administradores, ao invés de serem tomadas com base em fatos históricos que foram armazenados pelos diversos sistemas de informação.

O conceito de DW surgiu, pois, nos anos 90, como uma nova tecnologia para suprir esta demanda. Genericamente, entende-se por DW um banco de dados especializado que integra e gerencia o fluxo de informações a partir de dados armazenados em banco de corporativos. Este novo conceito impulsionou a criação de aplicações e ferramentas direcionadas a um outro tipo de processamento, conhecido como OLAP (*On line Analytical Processing*). /

Segundo TIEZZI (1997) a tecnologia da informação (TI) é o caminho encontrado pelas organizações para viabilizar a execução de novas estratégias de negócios. A TI é capaz de: proporcionar ênfase na dinâmica do mercado; efetuar uma maior aproximação com o cliente; oferecer qualidade nos produtos e nos serviços; e efetuar o crescimento planejado.

“O DW é uma nova tecnologia da informação que as empresas estão utilizando para auxiliar os seus gerentes nos processos decisórios. No cenário atual, onde estes gerentes necessitam tomar decisões rápidas em resposta às adversidades que ocorrem frequentemente em seus negócios, faz-se necessário um sistema informatizado que dê suporte às suas decisões, levantando informações rápidas e confiáveis e apresentando-as em um formato facilmente compreensível” (BISPO & CAZARINI, 1998, CD-ROM).

Antes do surgimento do DW, o processamento dos Sistemas de Apoio à Decisão ocorria em um único banco de dados ou plataforma que servia a todas as finalidades. Segundo INMON (1997), a necessidade de se criar um ambiente próprio para manipular informações vem como resposta a muitos fatores, tanto tecnológicos como econômicos e organizacionais.

Para FONSECA (1997), a introdução do DW se apresenta como uma razão natural para resolver os problemas de obtenção de informações estratégicas consolidadas.

“Um DW é um repositório centralizado de dados do negócio, que consolida as informações de múltiplas fontes operacionais internas ou dados externos para obter uma visão acurada e uniformizada dos clientes e da própria empresa” (BRETZKE, 2000 pg 190).

Willian H. Inmon, um dos precursores da ferramenta, a define da seguinte forma: "Um DW é um conjunto de dados baseado em assuntos, integrado, não volátil, e variável em relação ao tempo, de apoio às decisões gerenciais." (INMON, 1997, p. 33). De acordo com o autor, a ferramenta é:

- **Orientada por Assuntos**

Um DW sempre armazena dados importantes sobre temas específicos da empresa e conforme o interesse das pessoas que irão utilizá-lo. Ex: Uma empresa pode trabalhar com vendas de produtos alimentícios no varejo e o seu maior interesse é o perfil de seus compradores, então o DW será voltado para as pessoas que compram seus produtos e não para os produtos que ela vende.

Os sistemas operacionais clássicos são organizados em torno de aplicações e funções da empresa. No caso de instituições financeiras, as aplicações podem ser empréstimos, poupança, cartões magnéticos e financiamentos. O mundo do DW é organizado em torno dos principais assuntos como clientes, vendedores, produtos e atividades.

O mundo funcional e das aplicações está preocupado tanto com o projeto de banco de dados como com o projeto dos processos. Por outro lado, o desenvolvedor de DW focaliza-se exclusivamente na modelagem de dados e no projeto de banco de dados. O

projeto funcional e o de processos (em sua forma clássica) não fazem parte do processo de desenvolvimento do DW. (FIGURA 10)

O DW não inclui dados que não serão usados para processamento SAD - os dados orientados às aplicações possuem detalhes que satisfazem aos requisitos imediatistas do processamento funcional, que podem ou não ter relevância ou interesse ao analista de SAD.

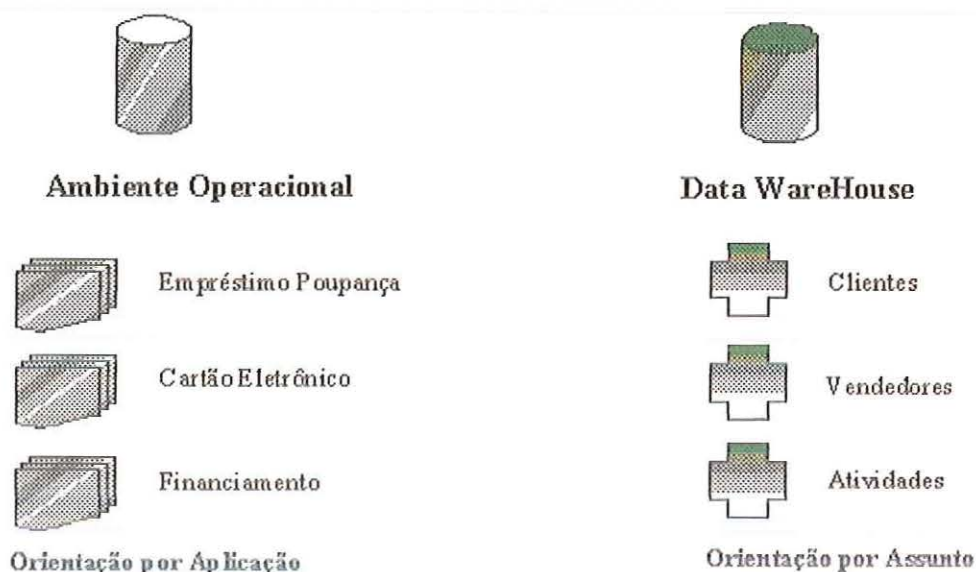


Figura 10: Diferença entre o Ambiente Operacional e Data Warehouse
Fonte: Adaptação INMON, 1997, p34

• Integrado

A segunda característica marcante do DW é o fato de que ele é integrado. De todos os aspectos do DW, esse é o mais importante. As diversas decisões de projeto que os projetistas de aplicações tomam ao longo dos anos, aparecem sob diversas formas diferentes. As aplicações não apresentam coerência em termos de codificações, unidades de medidas de atributos, atributos físicos, convenções de atribuição de nomes e assim por diante. Esta característica refere-se à consistência de nomes, das unidades das variáveis, etc., no sentido de que os dados foram transformados até um estado uniforme. Por exemplo, considere-se sexo como um elemento de dado. Uma aplicação pode codificar sexo como M/F, outra como 1/0 e uma terceira como H/M. Conforme os dados são trazidos para o DW, eles são convertidos para um estado uniforme, ou seja, sexo é codificado apenas de uma forma. Da mesma maneira, se um elemento de dado é medido em centímetros em uma aplicação, em polegadas em outra, ele será convertido para uma representação única ao ser colocado no

DW. Cada projetista de aplicação recebeu “carta branca” para tomar suas próprias decisões. (FIGURA 11)

Seja qual for o problema do projeto, o resultado é o mesmo: Os dados precisam ser armazenados no DW de uma forma única, aceitável globalmente, mesmo quando as aplicações operacionais subjacentes armazenam os dados de modo diferente.

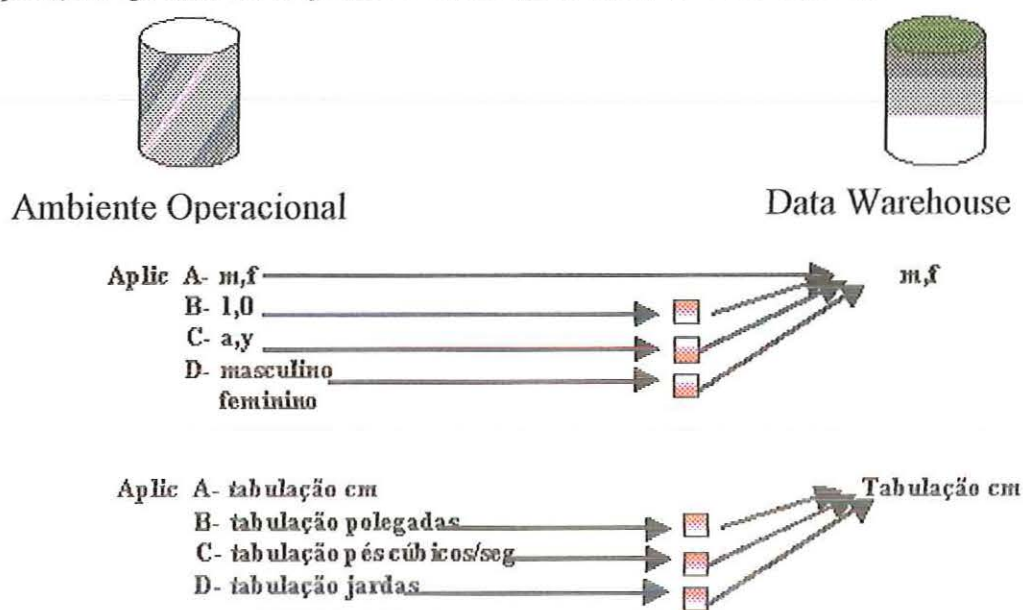


Figura 11: Exemplo de Integração no Data Warehouse
Fonte: Adaptação INMON, 1997, p35

- Não Volátil

Os dados operacionais são regularmente acessados e tratados um registro por vez. No ambiente operacional, os dados sofrem atualizações. Contudo, os dados existentes no DW apresentam um conjunto de características muito diferentes. Os dados do DW são carregados (normalmente em grandes quantidades) e acessados. Mas a atualização dos dados não ocorre no ambiente do DW. (FIGURA 12)

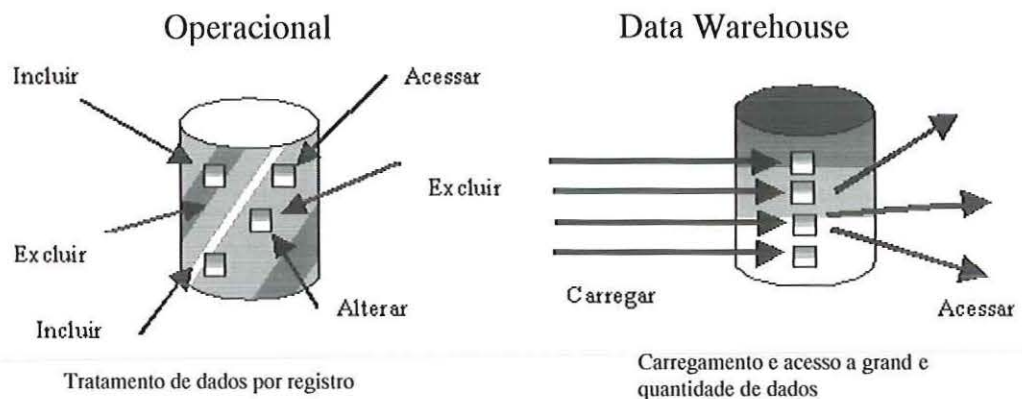


Figura 12: Exemplo de não volatilidade no Data Warehouse
 Fonte: Adaptação INMON, 1997, p36

• **Variável em Relação ao Tempo**

Os diversos modos pelos quais a variação em relação ao tempo se manifesta, são (FIGURA 13):

- Os dados no DW representam informações sobre um horizonte de tempo muito amplo: de cinco a dez anos. O horizonte de tempo representado nos dados operacionais é muito mais curto - de sessenta a noventa dias.
- A estrutura de chave dos dados operacionais pode conter ou não, elementos de tempo, como ano, mês, dia, etc. A estrutura de chave do DW sempre contém algum elemento de tempo.
- Os dados do DW, uma vez armazenados corretamente, não podem ser atualizados. Os dados operacionais, por outro lado, quando corretos no momento do acesso, podem ser atualizados na medida em que for necessário.



Ambiente Operacional

Horizonte de tempo – 60 a 90 dias
 A chave pode ou não conter um elemento de tempo
 Os dados podem ser atualizados



Data Warehouse

Horizonte de tempo – 5 a 10 dias
 A chave contém um elemento de tempo
 Os registros não podem ser atualizados

Figura 13: Exemplo de Variação em relação ao tempo no Data Warehouse
 Fonte: Adaptação INMON, 1997, p37

O DW é um conjunto de técnicas que aplicadas em conjunto geram um sistema de dados que proporcionam informações importantes aos tomadores de decisões (VIEIRA, 2001).

Para AMARAL (1997), DW é conjunto de arquiteturas e/ou sistemas de informação que viabilizam processos de tomada de decisões em diversos níveis organizacionais. Tais processos são baseados em grande volume de dados, principalmente históricos (5-10 anos), que manipulam dados nos níveis analítico e/ou sintético, relacionais ou multidimensionais, entrelaçados ou não, através de consultas (*read-only*) invariavelmente não-previsíveis.

De acordo com SOUZA & SULAIMAN (1997), o DW se refere ao processo de criação de um armazém de dados pela depuração e o armazenamento desses por assunto, tornando-os disponíveis para recuperação on-line. Esse armazém contém dados históricos integrados por assunto, ou domínio de aplicação, para a utilização em ambientes de análise de dados e sistemas de apoio à decisão.

Segundo ALVES (1997), o DW é um grande banco de dados que armazena informações integradas a partir de outros bancos de dados operacionais de uma organização. Uma das principais características da ferramenta é integrar dados. Os arquitetos de DW devem transformar e integrar dados operacionais dentro e fora da empresa e, então colocá-los em um DW, o qual será usado como ferramenta estratégica pelos usuários de primeira linha – os que tomam decisões-chave sobre negócios baseados em informações disponíveis. O autor ainda menciona que os dados armazenados no sistema operacional de uma empresa são um recurso, mas, de modo geral, raramente servem como recurso no seu estado original. É pela extração de dados do sistema operacional e pela respectiva integração deles ao DW que uma organização transforma os dados operacionais em uma ferramenta estratégica. As informações analíticas e integradas, em vez dos dados “crus”, permitem que uma organização tome decisões sobre missões importantes e negócios estratégicos.

TAURION (1997) define DW como um banco de dados voltado para suporte à decisão de usuários finais, derivado de diversos outros bancos de dados operacionais. Para o autor, o DW objetiva satisfazer as necessidades de análise de informação dos usuários, como monitorar e comparar as operações atuais com as passadas e, prever situações futuras, além de realizar análises de tendências e identificar relacionamentos muitas vezes desconhecidos. Ao transformar, consolidar e racionalizar as informações dispersas por diversos bancos de dados e plataformas, permite que sejam feitas análises estratégicas bastante eficazes em informações antes inacessíveis ou subaproveitadas.

Para LAMBERT (1997), um DW é um sistema complexo que integra diversos componentes, tais como software, hardware, redes de computadores, sistema de comunicação de dados, servidores, mainframes, sistemas de administração de banco de dados e diversas pessoas de diferentes unidades organizacionais com objetivos diferentes.

PARENTE (2001) também afirma que o DW é um processo complexo composto por vários itens como metodologias, técnicas, máquinas, bancos de dados, replicação e principalmente recursos humanos.

Segundo BERSON (1997) o DW é um conjunto de tecnologias que permitem converter uma grande quantidade de dados em informações utilizáveis; transformando um banco de dados operacional num ambiente que permite o uso estratégico dos dados. Uma característica importante no DW é que é um banco de dados modelado para tarefas analíticas usando dados de diferentes aplicações.

CIELO (2001) conceitua DW como grandes “*containers*” de dados por assunto, detalhados, não voláteis, históricos, integrados e estão focados sobre o propósito de um departamento ou da empresa toda. Além disso, são vistos como uma fonte de dados para geração dos sistemas de apoio à decisão.

Diversos conceitos são abordados por diferentes autores, mas todos com o mesmo intento, o de considerar um DW como um banco de dados especializado, que gerencia o fluxo de informações a partir de banco de dados corporativos e fontes de dados externas à empresa, a partir de dados históricos, que foram extraídos das fontes de dados operacionais da empresa e, em seguida, transformados, consolidados, agregados e validados, com o intuito de apoiar o processo de tomada de decisão para satisfazer as necessidades dos usuários (normalmente executivos, gerentes e analistas de negócios).

4.3 PRINCIPAIS FERRAMENTAS UTILIZADAS EM UM DATA WAREHOUSE

De acordo com BISPO (1998), existem várias ferramentas utilizadas em um DW:

1. Ferramenta para Armazenamento: são os bancos de dados, considerados o coração do DW e parte imprescindível do projeto.
2. Ferramenta para a extração de dados: busca na base de dados operacionais os dados que vão ser armazenados no DW.
3. Ferramenta para a transformação de dados: ajusta os dados para o formato do DW. Este formato auxilia as futuras pesquisas.

4. Ferramenta para a limpeza de dados: faz os ajustes necessários nos dados, fazendo correções, desmembramento e fusões de dados, quando necessário, visando melhorá-los para facilitar as futuras pesquisas.
5. Repositórios de metadados: estão intimamente relacionados às ferramentas de extração. Metadados são as definições dos dados que permitem saber a origem da informação, bem como todas as vezes em que ela foi alterada. Sua função é manter a consistência dos dados.
6. Transferência de dados e replicação: pode ser considerada um subconjunto das ferramentas de extração. Não faz nenhum tipo de processamento e transformação, apenas transfere um dado de um lugar "A" para "B". Geralmente, é utilizado para facilitar e viabilizar uma resposta mais rápida às consultas ou análises, movendo os dados para um lugar apropriado e fazendo o que for necessário para agilizar o serviço solicitado.
7. Gerenciamento e administração: é a típica ferramenta que só faz sentido depois que o DW está construído. Monitora o dia-a-dia, como o desempenho e a segurança do sistema.
8. *Query* ou ferramentas para gerenciamento de consultas: fazem consultas e/ou geram relatórios retirando os dados do DW, resumindo-os e apresentando-os em um formato apropriado.
9. Ferramentas para gerenciamento de relatórios: são semelhantes às ferramentas do item anterior, porém estão voltadas para a geração de relatórios mais complexos contendo, por exemplo, relatórios sintéticos e analíticos em conjunto, gráficos, e outros tipos de visualização dos dados.
10. OLAP: as ferramentas OLAP são aplicações que os usuários finais têm acesso para extraírem os dados de suas bases com os quais gera relatórios capazes de responder as suas questões gerenciais. Essa ferramenta surge juntamente com os sistemas de apoio à decisão para realizarem a extração e análise dos dados contidos nos DWs (www.datawarehouse.inf.br, 2000).
11. Data mining: ferramenta utilizada para descobrir novas correlações, padrões e tendências entre as informações de uma empresa, através da análise de grandes quantidades de dados armazenados em DW usando técnicas de reconhecimento de padrões, estatísticas e matemáticas.
12. Simulação: projeta cenários respondendo perguntas do tipo "e se", por exemplo: "e se os juros aumentarem, qual será o comportamento de minhas vendas?".

Para WELDON (1998), as ferramentas não devem apenas permitir o acesso aos dados, mas também permitir análises de dados significativas, de tal maneira a transformar dados brutos em informação útil para os processos estratégicos da empresa. O sucesso de um DW pode depender da disponibilidade da ferramenta certa para as necessidades de seus usuários.

A TABELA 4 apresenta um resumo com os principais tipos de ferramentas e suas características, que podem ser utilizadas para extrair informações de um ambiente de DW.

Tipo de ferramenta	Questão básica	Exemplo de resposta	Usuário típico e suas necessidades
Pesquisa e Relatórios	“O que aconteceu?”	Relat. Mensais de vendas, histórico do inventário	Dados históricos, habilidade técnica limitada
OLAP	“O que aconteceu e por que?”	Vendas mensais versus mudança de preço dos competidores	Visões estáticas da informação para uma visão multidimensional; tecnicamente estudo
EIS	“O que eu preciso saber agora?”	Memorandos, centros de comando	Informações de alto nível ou resumidas; pode não ser tecnicamente astuto
<i>Data Mining</i>	“O que é Interessante?” “O que pode acontecer?”	Modelos de previsão	Tendências e relações obscuras entre os dados; tecnicamente astuto

Tabela 4: Ferramentas do Data Warehouse
Fonte: Revista Byte Brasil, Janeiro 1997, pg64

4.4 CARACTERÍSTICAS DO DATA WAREHOUSE

O DW é uma coleção de dados derivados dos dados operacionais (banco de dados tradicional) para sistemas de suporte à decisão. Estes dados derivados são muitas vezes referidos como dados gerenciais, informacionais ou analíticos.

Os bancos de dados operacionais armazenam as informações necessárias para as operações diárias da empresa, são utilizados por todos os funcionários para registrar e executar operações pré definidas, podendo sofrer, portanto, constante mudanças conforme as necessidades atuais da empresa. Por não ocorrer redundância nos dados, e as informações

históricas não ficarem armazenadas por muito tempo, este tipo de banco de dados não exige grande capacidade de armazenamento.

Já um DW armazena dados analíticos, destinados às necessidades da gerência no processo de tomada de decisões, armazenando dados importantes sobre temas específicos da empresa e conforme o interesse das pessoas que irão utilizá-las.

Isto pode envolver consultas complexas que necessitam acessar um grande número de registros, por isso é importante a existência de muitos índices criados para acessar as informações da maneira mais rápida possível. Um DW armazena informações históricas de muitos anos e por isso deve haver uma grande capacidade de processamento e armazenamento dos dados que se encontram de duas maneiras, detalhados e resumidos.

Na TABELA 5 estão relacionados algumas diferenças entre Banco de Dados Operacionais e DW bem como as diferenças dos dados que eles manipulam segundo o DAL'ALBA (1998).

Características	Banco de Dados Operacionais	Data Warehouse
Objetivo	Operações diárias do negócio	Analisar o negócio
Uso	Operacional	Informativo
Tipo de processamento	OLTP	OLAP
Unidade de trabalho	Inclusão, alteração, exclusão	Carga e consulta
Número de usuários	Milhares	Centenas
Tipo de usuário	Operadores	Comunidade gerencial
Interação do usuário	Somente pré-definida	Pré-definida
Condições dos dados	Dados operacionais	Dados analíticos
Volume	Megabytes – Gigabytes	Gigabytes – Terabytes
Histórico	60 a 90 dias	5 a 10 anos
Granularidade	Detalhados	Detalhados e resumidos
Redundância	Não ocorre	Ocorre
Características	BD operacionais	<i>Data Warehouse</i>
Estrutura	Estática	Variável
Manutenção desejada	Mínima	Constante
Acesso a registros	Dezenas	Milhares
Atualização	Contínua (tempo real)	Periódica (em batch)
Integridade	Transação	A cada atualização

Número de índices	Poucos / simples	Muitos / complexos
Intenção dos índices	Localizar um registro	Aperfeiçoar consultas

Tabela 5: Comparação entre Banco de Dados Operacional e Data Warehouse

Fonte: <http://www.geocities.com/SiliconValley/port/5072>

De acordo com DAL'ALBA (1998), o DW é um meio que as empresas dispõem para analisar informações históricas podendo utilizá-la para melhoria dos processos atuais e futuros; são resumos de dados retirados de múltiplos sistemas de computação normalmente utilizados a vários anos e que continuam em operação; são construídos para que tais dados possam ser armazenados e acessados de forma que não sejam limitados por tabelas e linhas estritamente relacionais; seus dados podem ser compostos por um ou mais sistemas distintos e sempre estarão separados de qualquer outro sistema transacional, ou seja, deve existir um local físico onde os dados desses sistemas serão armazenados. O autor ainda menciona que antes deste deslocamento, sempre haverá aplicação de técnicas de filtragem, agrupamento e/ou refinamento dos dados.

4.5 METADADOS

CIELO (2001a), define metadado como sendo “dados dos dados”, ou seja, contém informações sobre todos os dados e processos envolvidos na construção e manutenção de um DW.

Segundo BERSON (1997), metadado é uma abstração dos dados, ou ainda, dados de mais alto nível que descrevem dados de um nível inferior. Para o autor, sem os metadados, os dados não possuem significado. Os metadados contém no mínimo:

- A localização e a descrição de um sistema de Warehouse e os componentes de dados;
- Nomes, definições, estruturas e conteúdo do DW;
- Regras de transformação e integração usadas para povoar um DW;
- Regras de transformação e integração usadas para “entregar” dados às ferramentas analíticas de usuário final;
- Autorização de segurança, lista de controle de acesso.

Para INMON (1997), metadados são dados sobre dados. Permitem que o usuário final/analista de SAD navegue pelas possibilidades. É por meio dos metadados que a utilização mais produtiva do DW é alcançada. Os metadados englobam o Warehouse e

mantém informações sobre o que está onde se encontra no Warehouse. Para o autor, os aspectos sobre os quais os metadados mantêm informação são:

- A estrutura dos dados segundo a visão do programador;
- A estrutura dos dados segundo a visão do analista de SAD;
- A transformação sofrida pelos dados no momento de sua migração para o DW;
- O modelo de dados;
- O relacionamento entre o modelo de dados e o DW;
- O histórico das transações.

MCELREATH (1997)⁹ *apud* BISPO (1998), os metadados são usados para controlar o uso, a segurança, a integridade e o custo do DW. Ajudam os administradores de banco de dados (DBAs) a povoarem e gerenciarem o DW, e os usuários a saber a origem dos dados. Não existem muitas ferramentas disponíveis no mercado para gerenciá-los, e muitas empresas costumam desenvolver as suas próprias ferramentas.

Normalmente os metadados em um DW podem ser apresentados em três camadas diferentes:

- Metadados Operacionais: definem a estrutura dos dados mantidos pelos bancos operacionais, usados pelas aplicações de produção da empresa;
- Metadados centrais do DW: são orientados por assunto e definem como os dados transformados devem ser interpretados, incluem definições de agregação e campos calculados, assim como visões sobre cruzamentos de assuntos.
- Metadados do Nível do Usuário: organizam os metadados do DW para conceitos que sejam familiares e adequados aos usuários finais.

KIMBALL (1998a)¹⁰ *apud* BISPO (1998) apresenta vários exemplos de metadados:

- dados sobre os sistemas de informação que servem de fontes de dados;
- descrições dos negócios envolvidos em cada fonte de dados; estruturas dos bancos de dados fontes;
- descrições das propriedades de cada fonte de dados;
- frequência de atualização das fontes de dados;
- frequência de atualização do DW;

⁹ MCELREATH, J. (1997). Data warehouses: an architectural perspective. Perspectives, Oct. 1995. http://www.csc.com/about/tech_dw_arch.html (29 dez.).

¹⁰ KIMBALL, R. (1998a). Meta meta data data. DBMS Magazine, Mar. 1998. <http://www.dbmsmag.com/9803d05.html> (27/08).

- fidelidade dos dados, por exemplo: se todos os dados de janeiro foram recebidos durante o mês de janeiro;
- relações e restrições de integridade referencial;
- purgação e períodos de retenção;
- regras de replicação e de distribuição;
- métodos e regras de agregação.
- limitações legais do uso de cada fonte de dados;
- horários de trabalhos dos sistemas onde se encontram as fontes de dados;
- métodos e permissões de acesso, privilégios e senhas das fontes de dados;
- dados sobre o conjunto de ferramentas para extração automática;
- extratos ou resultados de cada processo de extração, incluindo o tempo necessário, conteúdos e nível de perfeição da operação;
- agenda de transmissão de dados e os resultados destas operações;
- definições da tabelas de fatos e das dimensões;
- especificações do trabalho de união dos dados de várias fontes, incluir/excluir campos e atributos;
- especificações para a auditoria de dados;
- versões dos diversos softwares envolvidos em todas as fases;
- estatísticas de uso dos dados;
- regras de segurança para a extração, limpeza e transformação dos dados;
- regras de segurança para a transmissão dos dados;
- regras de segurança para o armazenamento dos dados;
- descrição das diversas tabelas utilizadas para o armazenamento;
- descrição dos diversos índices utilizados;
- especificações dos dispositivos de armazenamento; da limpeza dos dados;
- especificações da transformação dos dados;
- especificações das agregações;
- especificações definições dos processos de visualização dos dados;
- definições dos tipos de consultas e tipos de relatórios possíveis;
- especificações da documentação e do treinamento para os técnicos e usuários;
- estatísticas do uso da ferramenta.

Em uma visão geral os metadados são as informações sobre os dados mantidos pela empresa. Dados em um dicionário de dados - são exemplos de metadados. Com uma grande variedade de metadados consegue-se manter a funcionalidade de um ambiente de DW.

Idealmente o usuário poder ter acesso aos dados de um DW sem que tenha que saber onde residem estes dados ou a forma como estão armazenados.

4.6 ARQUITETURA DO DATA WAREHOUSE

Ultimamente o conceito de DW evoluiu rapidamente de um considerável conjunto de idéias relacionadas para uma arquitetura voltada para a extração de informação especializada e derivada a partir dos dados operacionais da empresa. O estudo de uma arquitetura descrevendo o ambiente de DW permite compreender melhor a estrutura geral de armazenamento, integração, comunicação, processamento e apresentação dos dados que servirão para subsidiar o processo de tomada de decisão nas empresas.

Existem diferentes tipos de arquiteturas para DW, dentre elas: a arquitetura genérica, a arquitetura de duas camadas e a arquitetura de três camadas. Este trabalho procurou limitar-se na arquitetura genérica por permitir que as diferentes abordagens encontradas no mercado atual possam se enquadrar nesta descrição genérica.

A arquitetura genérica apresentada por ORR (1996) é apresentada na FIGURA 14 e logo após explicada. Esta arquitetura é composta pela camada de dados operacionais e outra fonte de dados que são acessados pela camada de acesso aos dados. As camadas de gerenciamento de processos, transporte e DW formam o centro da arquitetura e são responsáveis por manter e distribuir os dados. A camada de acesso à informação é formada por ferramentas que possibilitam os usuários extrair informações do DW. Todas as camadas desta arquitetura interagem com os metadados e com o gerenciador de processos.

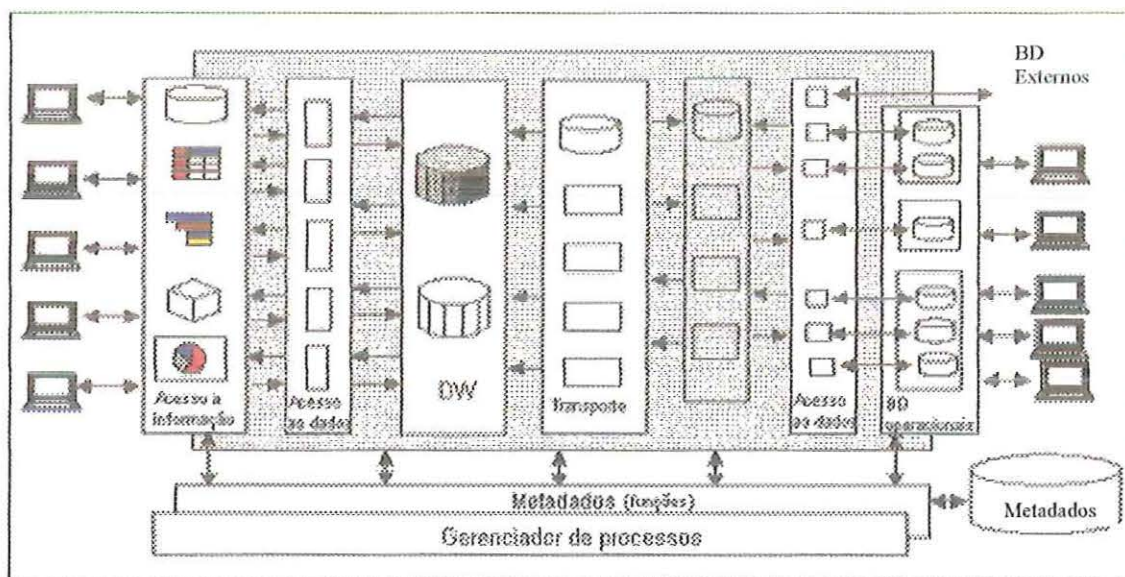


Figura 14: Arquitetura Genérica do Data Warehouse

Fonte: <http://www.kenorrinst.com/warehousing.html>

Camada de Banco de Dados Operacionais e Fontes Externas: Composta pelos dados dos sistemas operacionais das empresas e informações oriundas de fontes externas que serão integradas para compor o DW;

Camada de Acesso aos Dados: Realiza a ligação entre as ferramentas de acesso à informação e os bancos de dados operacionais. Esta camada se comunica com diferentes sistemas de bancos de dados, sistemas de arquivos e fontes, sob diferentes protocolos de comunicação.

Camada de Gerenciamento de Processos: É responsável pelo gerenciamento dos processos que contribuem para manter o DW atualizado e consistente. Está envolvida com o controle das várias tarefas que devem ser realizadas para construir e manter as informações do dicionário de dados e do DW;

Camada de Transporte: Esta camada gerencia o transporte de informações pelo ambiente de rede. Inclui a coleta de mensagens e transações e se encarrega de entregá-las em locais e tempos determinados. Também é usada para isolar aplicações operacionais ou informacionais, do formato real dos dados nas duas extremidades;

Camada do DW: É o DW. Corresponde aos dados utilizados para obter informações. Às vezes, o DW pode ser simplesmente uma visão lógica ou virtual dos dados, podendo não envolver o armazenamento dos mesmos ou armazenar dados operacionais e externos para facilitar seu acesso e manuseio;

Camada de Acesso a Informação: Envolve o hardware e o software utilizado para obtenção de relatórios, planilhas, gráficos e consultas. É nesta camada que os usuários finais interagem com o DW, utilizando ferramentas de manipulação, análise e apresentação dos dados, incluindo-se as ferramentas de *Data Mining* e visualização;

Camada de Metadados (Dicionário de Dados): Metadados são as informações que descrevem os dados utilizados pela empresa, tais como, descrições de registros, comando de criação de tabelas, diagramas Entidade/Relacionamentos (E-R), dados de um dicionário de dados, etc. É necessário que exista uma grande variedade de metadados no ambiente de DW para que ele mantenha sua funcionalidade e os usuários não precisem se preocupar onde residem os dados ou a forma como estão armazenados.

4.7 MODELAGEM DOS DADOS

Apontar no momento, uma metodologia consolidada e amplamente aceita para o desenvolvimento de DWs, é difícil. O que se nota na literatura, são propostas no sentido de construir um modelo dimensional a partir da base dos bancos de dados operacionais da empresa, de forma incremental. De qualquer forma, a metodologia a ser adotada é ainda bastante dependente da abordagem escolhida, em termos de ambiente, distribuição, etc.

O DW objetiva responder, aos usuários, questões estratégicas a respeito do negócio da empresa. Desta forma, o primeiro passo na modelagem do DW é identificar e entender os tipos de questões que serão propostas. No processo de modelagem, as entidades das quais se deseja saber recebem o nome de dimensões enquanto o dado que se quer saber sobre elas recebe o nome de fato.

As dimensões normalmente são dados não-numéricos, enquanto as tabelas de fatos são dados numéricos. Desta forma, as agregações são feitas sobre as medidas, por exemplo, o total ou a média das vendas. Este é motivo pelo qual a análise destes dados é chamada de multidimensional. Por conseguinte, a modelagem destes dados é feita através de estruturas que procuram representar esta multidimensionalidade.

Dados multidimensionais podem ser representados por estruturas relacionais e armazenados em bancos de dados relacionais. Os esquemas em "estrela" e em "flocos de neve" são exemplos de formas de modelagem de dados multidimensionais em tabelas relacionais.

Segundo INMON (1992), a especificação de requisitos do ambiente de apoio à decisão associado a um DW é diferente da especificação de requisitos dos sistemas que sustentam os processos usuais do ambiente operacional de uma empresa - estes são identificáveis a partir das funções a serem executadas pelo sistema e são estáveis ao longo do tempo. Requisitos de sistemas de suporte à decisão são, por sua vez, indeterminados e são instáveis, ou seja, dependem das variações das necessidades de informações daqueles responsáveis pelas tomadas de decisões dentro da empresa. A função do DW é fornecer dados com qualidade. Os requisitos dependem das necessidades de informação individuais de seus usuários.

No entanto, embora as necessidades por informações específicas mudem com frequência, os dados associados não mudam. Imaginando-se que os processos de negócio de uma empresa permaneçam relativamente constantes, existe apenas um número finito de objetos e eventos com as quais uma organização está envolvida. Por esta razão, um modelo de dados é uma base sólida para identificar requisitos para um DW.

INMON (1992) afirma que não se deve pensar que técnicas de projeto que servem para sistemas convencionais serão adequadas para a construção de um DW.

Outra questão interessante, discutida por KIMBALL (1998), é a adequação do modelo E-R ao tipo de transação de sistemas OLTP. O principal objetivo da modelagem, neste caso, é eliminar ao máximo a redundância, de tal forma que uma transação que promova mudanças no estado do banco de dados, atue o mais pontualmente possível. Com isso, nas metodologias de projeto usuais, os dados são "fragmentados" por diversas tabelas, o que traz uma considerável complexidade à formulação de uma consulta por um usuário final. Por isso, salienta KIMBALL (1998), esta abordagem não parece ser a mais adequada para o projeto de um DW, onde estruturas mais simples, com menor grau de normalização devem ser buscadas.

Para KIMBALL (1998), desenvolver um DW significa unir as necessidades dos seus usuários com a realidade dos dados disponíveis. O autor ainda menciona um conjunto de nove pontos de decisão, fundamentais no projeto da estrutura de um DW:

- os processos, e por conseqüência, a identidade das tabelas de fatos;
- a granularidade de cada tabela de fatos;
- as dimensões de cada tabela de fatos;
- aos fatos, incluindo fatos pré-calculados;
- os atributos das dimensões;
- como acompanhar mudanças graduais em dimensões;
- as agregações, dimensões heterogêneas, minidimensões e outras decisões de projeto físico;
- duração histórica do banco de dados;
- a urgência com que se dá a extração e carga para o DW.

4.7.1 MODELO DIMENSIONAL

De acordo com BISPO (1998), o modelo dimensional é assimétrico. Possui uma grande tabela, que é a principal, está localizada no centro do diagrama e possui outras tabelas secundárias ao seu redor, que são menores e que se relacionam com a tabela principal. A tabela central é chamada de tabela de fatos e as demais são chamadas tabelas de dimensão. A tabela de fatos armazena as medidas numéricas do negócio, por exemplo: unidades vendidas em cada transação, unidades produzidas em cada lote, etc. Esta tabela é composta por uma chave composta. As tabelas de dimensão armazenam as descrições textuais das dimensões do negócio, por exemplo: a dimensão produto possui os dados sobre a marca, a categoria, o tipo

de embalagem, o tamanho do produto, etc. As dimensões cliente e fornecedor possuem os seguintes dados: nome, endereço completo e outros dados específicos que auxiliam as análises do negócio. Cada tabela de dimensão tem uma única chave primária, que corresponde aos componentes da chave composta da tabela de fatos. O DW possui várias tabelas de fatos, cada uma representando um assunto ou negócio diferente dentro da organização.

A resposta às perguntas complexas e que envolvam questões de análise dos negócios de uma empresa, normalmente requerem uma visão dos dados de perspectivas diferentes. As respostas a esse tipo de perguntas é que podem levar a tomadas de decisões acertadas ou não. As ferramentas baseadas em SQL¹¹ podem ajudar na pesquisa de dados relacionais a este tipo de consulta - o que ocorre é que normalmente as respostas não são conseguidas em um tempo curto, principalmente pela falta de flexibilidade destas ferramentas.

Para exemplificar, imaginemos um negócio qualquer que esteja querendo melhorar o desempenho de suas vendas ou saber se suas promoções estão trazendo bons resultados. Para isso, é necessário examinar os dados sobre as vendas disponíveis na empresa. Uma avaliação deste tipo requer uma visão histórica do volume de vendas por produto, volume de vendas por marca, volume de vendas por filial, volume de vendas por período de tempo. Chama-se de dimensões as diferentes perspectivas envolvidas, no caso, produto, marca, filial e mês.

Segundo KIMBALL (1998), a modelagem Dimensional é o nome de uma técnica de projeto lógico freqüentemente usada para DW, cujo principal objetivo é apresentar o dado em uma arquitetura padrão e intuitiva, que permita acessos de alta performance.

Dados dimensionais podem ser armazenados e representados em estruturas relacionais, para isso é necessário utilizar formas específicas de modelagem como o modelo “Estrela” descrito a seguir.

4.7.1.1 MODELO ESTRELA

Tradicionalmente, modelos de bases de dados relacionais apresentam tabelas com relacionamentos complexos e com múltiplas uniões circulares entre dois pontos do modelo. Para a maioria dos usuários que utilizam ferramentas para compor suas consultas é necessário que o acesso à base de dados seja simples o suficiente para facilitar o acesso direto. Para acomodar as necessidades de todos os usuários e facilitar a utilização do DW o projetista deve criar um modelo que o usuário final possa facilmente entender em termos de negócio.

¹¹ Nota da autora: SQL é uma linguagem padrão para acessar banco de dados relacionais.

O principal tipo de modelo dimensional, é o chamado Modelo Star (Estrela) , onde existe uma tabela dominante no centro, chamada de tabelas de fatos, com múltiplas junções conectando-a a outras tabelas, sendo estas chamadas de tabelas de dimensão. Cada uma das tabelas secundárias possui apenas uma junção com a tabela central. O modelo Estrela, ilustrado na FIGURA 15, tem a vantagem de ser simples e intuitivo, mas também faz uso de novos enfoques de indexação e união de tabelas.

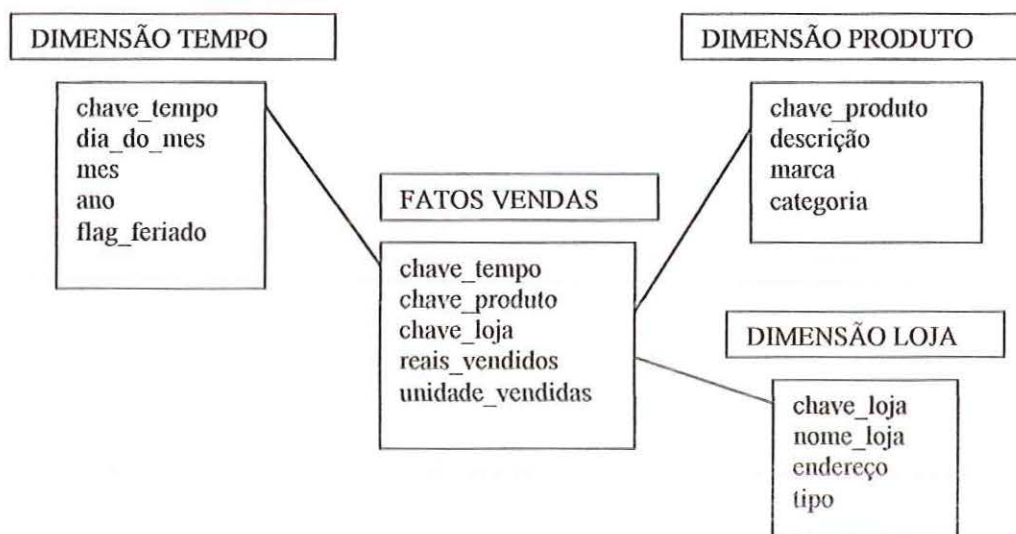


Figura 15: Modelo Estrela

A tabela de fatos contém milhares ou milhões de valores e medidas do negócio da empresa, como transações de vendas ou compras. Cada uma destas medidas é tomada segundo a interseção de todas as dimensões. As medidas numéricas são o número de reais vendidos e o número de unidades vendidas. Os fatos melhores e mais úteis são numéricos , continuamente valorados (diferentes a cada medida) e aditivos, já que estes facilitam a geração do conjunto de respostas. Uma outra característica da tabela de fatos é a esparsidade, ou seja, se não existe um cruzamento para alguns valores das dimensões, a tabela de fatos não armazena valores zeros.

Para DINIZ (2000), as tabelas de dimensão armazenam as descrições textuais das dimensões do negócio. Cada uma dessas descrições textuais ajuda a definir um componente da respectiva dimensão.

Uma das principais funções dos atributos de tabelas de dimensão é servir como fonte para restrições em uma consulta ou como cabeçalhos de linha no conjunto de resposta do usuário.

Um fato importante é que a tabela de fatos na verdade representa os relacionamentos muitos-para-muitos entre as tabelas de dimensões, esta tem como chave primária uma chave composta de todas as chaves estrangeiras das tabelas de dimensão.

Para um bom desempenho do Modelo Estrela é necessário que os projetistas saibam antecipar, na modelagem do DW, as consultas mais frequentes a serem realizadas pelos usuários. Com a redundância seletiva e relacionamentos pré-estabelecidos o projetista pode simplificar os dados facilitando seu sucesso.

4.7.1.2 VANTAGENS DO MODELO ESTRELA

Segundo KIMBALL(1998), o modelo dimensional apresenta várias vantagens no que diz respeito a sua utilização para DW:

- O modelo estrela tem uma arquitetura padrão e previsível. As ferramentas de consulta e interfaces do usuário podem se valer disso para fazer suas interfaces mais amigáveis e fazer um processamento mais eficiente;
- Todas as dimensões do modelo são equivalentes, ou seja, podem ser vistas como pontos de entrada simétricos para a tabela de fatos. As interfaces do usuário são simétricas, as estratégias de consulta são simétricas, e o SQL gerado, baseado no modelo, é simétrico;
- O modelo dimensional é totalmente flexível para suportar a inclusão de novos elementos de dados, bem como mudanças que ocorram no projeto. Essa flexibilidade se expressa de várias formas, dentro as quais temos:
 - Todas as tabelas de fato e dimensão podem ser alteradas simplesmente acrescentando novas colunas a tabelas;
 - Nenhuma ferramenta de consulta ou relatório precisa ser alterada de forma a acomodar mudanças;
 - Todas as aplicações que existiam antes das mudanças continuam “rodando” sem problemas.
- Existe um conjunto de abordagens padrões para tratamento de situações comuns no mundo dos negócios. Cada uma destas tem um conjunto bem definido de alternativas que podem então ser especificamente programadas em geradores de

relatórios , ferramentas de consulta e outras interfaces do usuário. Dentre estas situações temos:

- Mudanças lentas das dimensões : ocorre quando uma determinada dimensão evolui de forma lenta e assíncrona;
- Produtos heterogêneos : quando um negócio , tal como um banco, precisa controlar diferentes linhas de negócio juntas, dentro de um conjunto comum de atributos e fatos, mas ao mesmo tempo esta precisa descrever e medir as linhas individuais de negócio usando medidas incompatíveis.
- Outra vantagem é o fato de um número cada vez maior de utilitários administrativos e processo de software serem capazes de gerenciar e usar agregados, que são de suma importância para a boa performance de respostas em um DW.

4.7.2 COMPARAÇÃO ENTRE O MODELO RELACIONAL E O MODELO DIMENSIONAL

Uma das principais diferenças entre o modelo relacional e o modelo dimensional é a complexidade . Por ser normalizado, o modelo relacional gera dezenas ou até centenas de tabelas conectadas entre si, tornando-o confuso e de difícil compreensão por parte dos usuários. Já o modelo multidimensional do tipo estrela apresenta uma estrutura muito mais simples, onde uma tabela central, a tabela de fatos, é ligada a várias tabelas de dimensões uma única vez, tornando o modelo de fácil compreensão.

Conforme DAL'ALBA (1998) modelos relacionais são bem flexíveis para consultas pré-definidas. O desempenho dos sistemas de banco de dados relacionais tradicionais é melhor para consultas baseadas em chaves do que consultas baseadas em conteúdo, o que acontece utilizando o OLAP.

A construção do modelo relacional se baseia no relacionamento entre os dados, ou seja, a escolha das entidades e seus relacionamentos é feita baseando-se em como as coisas acontecem. Este tipo de modelo divide os dados em diversas tabelas, que se relacionam entre si, formando um complexo diagrama de difícil interpretação e análise.

Já no modelo dimensional, a escolha das dimensões e atributos de fatos são baseados na opinião do usuário, o que faz com que o modelo reflita as suas necessidades. Enxerga também as informações sob uma perspectiva histórica ao invés de transações atômicas, tal como no modelo relacional.

O modelo multidimensional é globalmente consistente para toda a empresa, ao passo que o modelo relacional é consistente somente dentro de seu escopo. A principal razão disso é que o modelo dimensional tem como objetivo fornecer uma visão integrada dos dados da empresa.

Os relacionamentos no modelo relacional são modelados explicitamente, ao passo que no modelo multidimensional estes relacionamentos são representados pela existência de fatos no cruzamento entre as dimensões.

Para DINIZ (2000a), o modelo dimensional irá permitir ao usuário analisar seu negócio e qualquer uma de suas dimensões e visões de negócio; podendo enxergar um conjunto de dados de uma forma concreta e compreensível. De forma ilustrada o autor apresenta um exemplo (FIGURA 16) utilizando a ferramenta OLAP para acesso e análise dos dados, para que fique mais claro como o modelo multidimensional funciona.

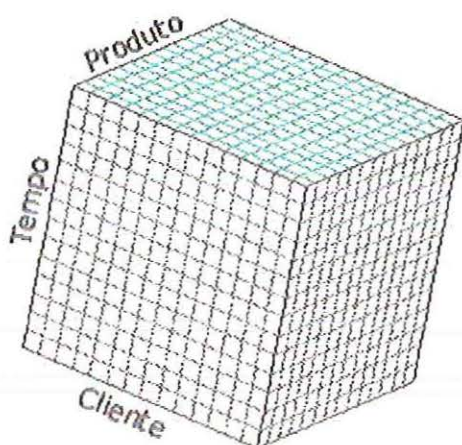


Figura 16: Faces de um cubo
Fonte: <http://www.datawarehouse.inf.br>

Após as informações estarem modeladas o usuário iniciará a análise, necessitando em principio de algumas informações.

Primeiramente é realizada a procura na coluna da tabela, as ocorrências do objeto pedido pelo usuário. Depois de toda a varredura é apresentado ao usuário o seu pedido.

Após a análise acima descrita, as informações obtidas serão cruzadas com as de outras dimensões para uma consulta mais completa e específica. A FIGURA 17 apresenta uma informação em que o período compreende o ano de 1999.

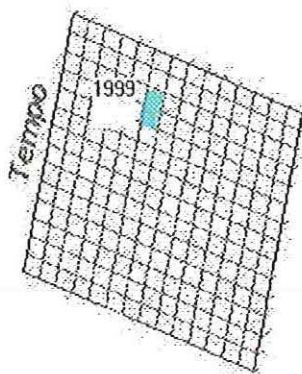
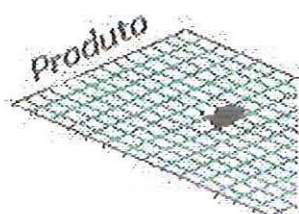


Figura 17: Dimensão Tempo
 Fonte <http://www.datawarehouse.inf.br>

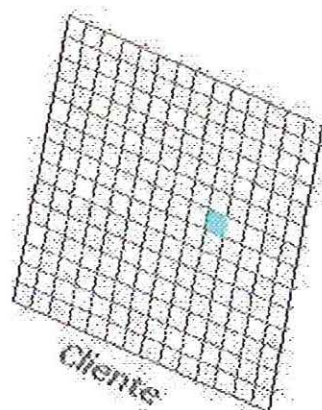
Já na FIGURA 18 o usuário pede a informação sobre um produto específico, como por exemplo, Laranja.



Dentro de uma variedade enorme de produtos só será capturado como informação os produtos que o usuário escolheu. Ex: Laranja

Figura 18: Dimensão Produto
 Fonte <http://www.datawarehouse.inf.br>

Finalizando a consulta, o usuário solicita a informação de um determinado Cliente. Por exemplo: Supermercado Super Super. Ilustrado na FIGURA 19.



Em muitas ocorrências, somente é capturada a informação desejada pelo cliente.

Figura 19: Dimensão Cliente
 Fonte <http://www.datawarehouse.inf.br>

Finalizando , as informações são cruzadas e passadas para o usuário final, conforme ilustrado na **FIGURA 20**.

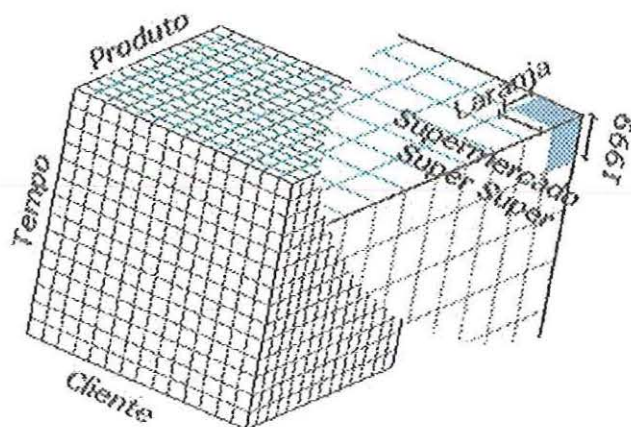


Figura 20: Informações Cruzadas

Fonte <http://www.datawarehouse.inf.br>

Nota-se portanto que o modelo dimensional proporciona um ganho de tempo na consulta, uma melhor organização do sistema e principalmente a sua utilização de forma intuitiva para o usuário. (DINIZ, 2000).

/ 4.8 DATA MART

Os primeiros projetos sobre DW, referiam-se a uma arquitetura centralizada. Embora fosse de interesse fornecer uniformidade, controle e maior segurança, a implementação desta abordagem não é uma tarefa fácil pois requer uma metodologia rigorosa e uma completa compreensão dos negócios da empresa. Esta abordagem pode ser longa e dispendiosa e por isso sua implementação exige um planejamento bem detalhado.

Devido a esse problema, surgiu o DW departamental, ou datamart; uma abordagem descentralizada que passou a ser uma das opções de arquitetura DW.

DEPOMPA (1998)¹² apud BISPO (1998) argumenta que muitas empresas construíram seus datamarts especializados para atender as necessidades imediatas de consultas e análises dos dados, sem antes criar uma infra-estrutura física e lógica para o DW.

¹² DEPOMPA, B. (1998). *Companies that start whit data marts find it difficult to move to a warehouse*. Information Week, <http://techweb.cmp.com/iw/605/05olwa1.htm> (14/01).

Por meio desta ferramenta, atualmente, os usuários podem obter ganhos quantitativos e qualitativos no suporte ao processo decisório, perdendo menos tempo com o acesso e a análise dos dados e podem dedicar mais tempo elaborando estratégias para seus negócios.

Para NIMER (1998) datamarts são subconjuntos de dados, dentro de um DW, projetados para dar suporte a negócios de unidades organizacionais específicas.

De acordo com INMON (1997), a fonte de todos os dados departamentais é o DW. O nível departamental é, por sua vez, chamado de nível do datamart.

Segundo STEVENS (2000), datamarts são tipicamente menores e menos complexos do que os DWs, além disso, são mais fáceis de se construir e de se manter.

Os datamarts possuem a mesma estrutura dos DWs, o que difere é o escopo. Enquanto o DW é voltado para os assuntos da empresa toda, os datamarts são projetados para atender as necessidades de um departamento específico, área da empresa ou grupo de usuários. Podem ser usados independentes, integrados a outros datamarts e até interligados para, juntos construírem um DW. A FIGURA 21 ilustra o datamart

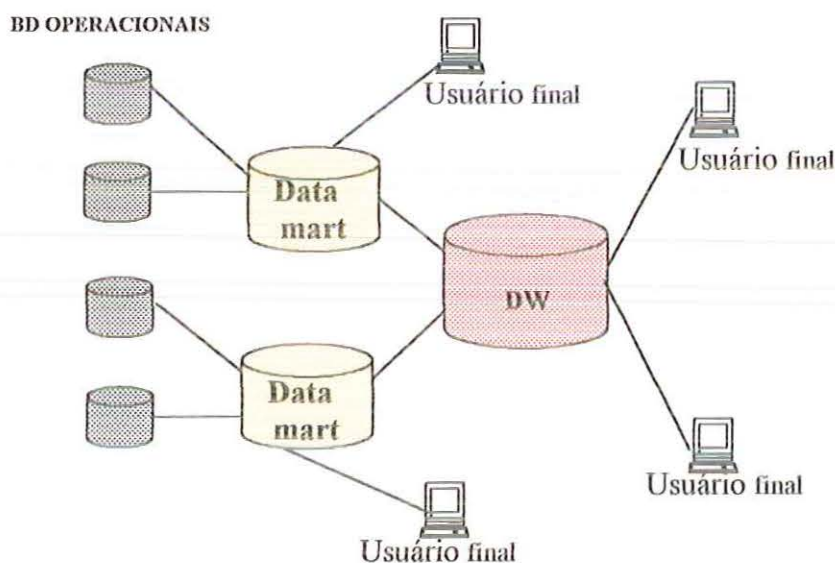


Figura 21: Datamart

Os datamarts podem surgir de duas maneiras:

- Top-down: a empresa cria um DW e depois parte para a segmentação, ou seja, divide esse DW em partes menores, gerando assim pequenos bancos de dados, orientados por assunto departamentalizados;

- Botton-up: ocorre a situação inversa. A empresa primeiro cria um banco de dados somente para um departamento ou grupo de usuários. Com isso os custos são bem inferiores de um projeto completo de DW.

É de responsabilidade da empresa avaliar sua demanda e optar pela melhor solução: implementar um Data Mart a curto prazo ou implementar um DW a longo prazo. /

4.9 A IMPORTÂNCIA DO DATA WAREHOUSE NO COMÉRCIO ELETRÔNICO

De acordo com CABRERA (2000), com o advento do DW, surgiu a necessidade de mapear melhor as questões gerencias dos *sites* da Web. O aumento significativo de páginas dedicadas ao comércio eletrônico, fizeram com que fosse pensada uma maneira de criar um Armazém de Dados que seguisse conceitos e características semelhantes aos do DW criando assim uma espécie de Armazém de Dados para Web, também reconhecido por Data Webhouse. Seguindo as mesmas características do DW, o Data Webhouse é orientado por assunto, não volátil, variável no tempo e integrado.

Para CABRERA (2000), o Data Webhouse permite:

- analisar todo o caminho realizado por um visitante em um *site* da Web;
- mapear cada clique;
- conhecer melhor cada cliente através da informação disponibilizada nos arquivos de *Log* do servidor e *Cookies*.

Com o Data Webhouse é possível realizar a análise de cada evento realizado dentro do *site* pelo visitante.

O motivo de se “trazer” a Web para o Data Webhouse está ligado ao fato de que uma das intenções dos analistas de *marketing* e de negócios é a de registrar a quantidade máxima de “gestos” efetuados por indivíduos navegando por *sites* remotos da Web. Tais registros se traduzem em informações comportamentais, e a compreensão de registros comportamentais podem “guiar” as ações de *marketing*, ou seja, tanto para programas de promoções de vendas quanto para saudações personalizadas. De qualquer forma, para que tais registros comportamentais “guiem” as ações de *marketing*, os mesmos devem estar armazenados em um banco de dados que apresente as mesmas características que o DW, o Data Webhouse. Trazer a Web para o DW significa identificar vários tipos de camadas de comportamento através da seqüência de cliques realizada pelo usuário. Os dados de seqüência de cliques, por sua vez, são colecionados simultaneamente por servidores físicos diferentes através de arquivos de *log*, mesmo se usuário está interagindo com um único *site* da Web. Além dos próprios arquivos de *log*, pode-se obter dados dos clientes através de :

- Parceiros direcionados (*referring*);
- *Internet Service Provider (ISP)* (FIGURA 22);
- Serviços de estatística da Web;
- Sistema de busca a qual dirige o usuário para o *site* da Web.

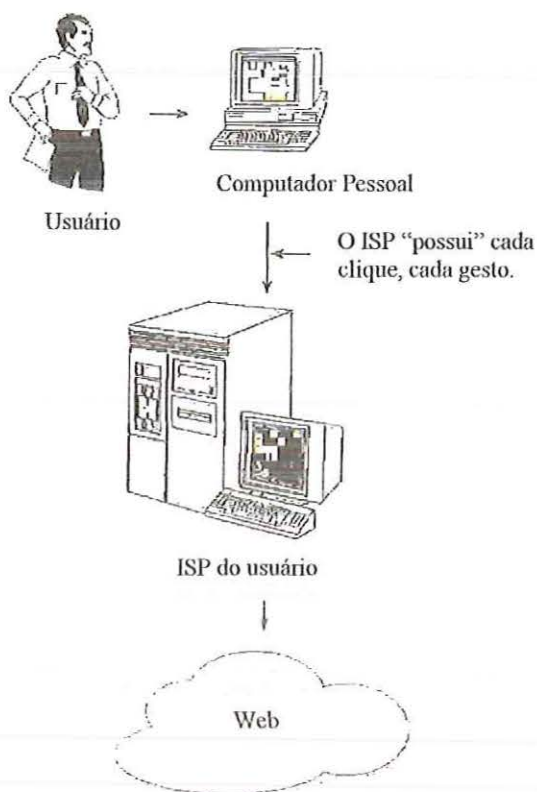


Figura 22: O relacionamento único entre o ISP e o usuário conectado
 Fonte: KIMBALL & MERZ , 2000, pg 23

De acordo KIMBALL (1999), atualmente, o número de aplicações ao redor de páginas da Web vem crescendo. Para se analisar todos os cliques da Rede, construir uma visão compreensível do fluxo dos inúmeros cliques que chegam ao servidor, saber com que tipo de usuário se lida, deve-se criar um Data Mart de Fluxo de Cliques (DM-FC), ou seja, um pequeno DW capaz de detalhar o comportamento de um cliente. O DM – FC através de um modelo dimensional (FIGURA 23) será capaz de identificar:

- Qual o local mais visitado de um determinado *site* ?
- Qual o local que apresenta maior número de vendas em um determinado *site* ?
- Qual o local mais supérfluo ou visitado com menos frequência em um determinado *site* ?

- Que página do *site* é vista como “Sessão Final”, onde os usuários geralmente encerram a Sessão ?
- Onde o novo usuário clica nas primeiras visitas de acordo com seu perfil ?
- Qual é o perfil de navegação de um cliente existente ? O cliente é lucrativo ? O cliente reclama freqüentemente ?
- Como induzir o cliente a se registrar no *site* para que possamos conhecer mais sobre ele ?
- Quantas visitas o cliente faz antes de se registrar ?
- Antes de registrados eles compram algum produto ou serviço ?

Para identificar cada usuário, o modelo dimensional utiliza *cookies*, que são arquivos com informações que o usuário aceita armazenar em sua máquina. Os mesmos, contém um ID para os dados armazenados em seu próprio Banco de Dados. São acessados e enviam informações ao servidor de origem toda vez que uma página for carregada pelo browser do usuário.

Esses dados são consolidados para poderem ser utilizados pelo DW.

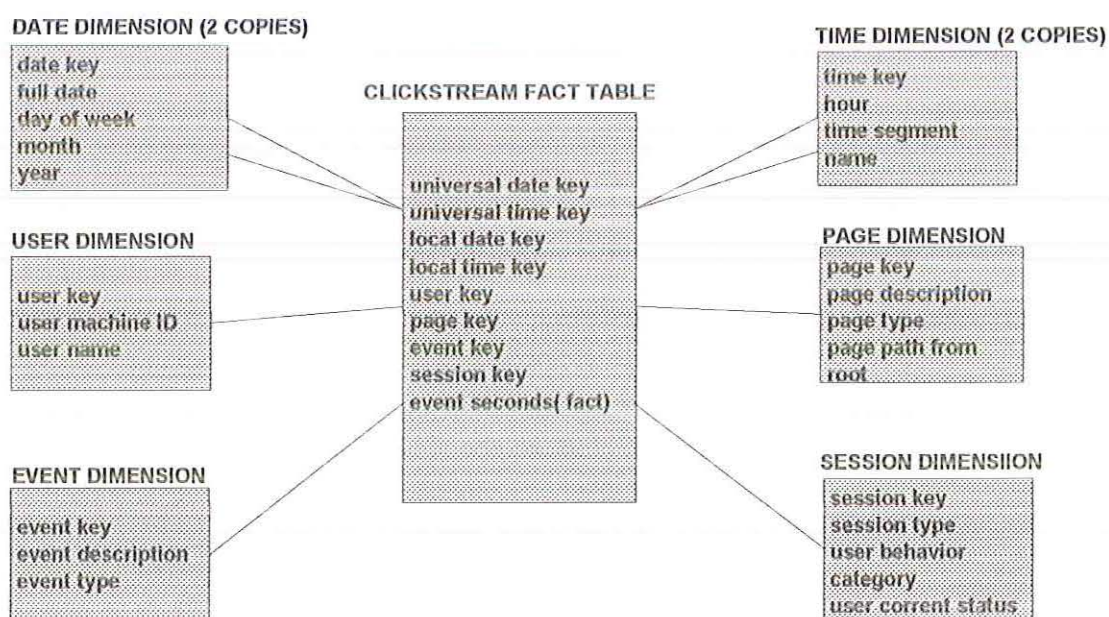


Figura 23: Dimensões e Fato de um DM – FC

Fonte: http://www.intelligententerprise.com/db_area/archives/1999/990501/warehouse.shtml

Portanto, nota-se que com a globalização, os negócios não possuem fronteiras. Informações corretas em um espaço curto de tempo têm se tornado o grande diferencial competitivo para as empresas. DWs têm mostrado que servem como fonte de dados para diversas aplicações, assegurando a consistência, a integração, e a precisão dos dados. Além

disso, podem auxiliar tanto na forma convencional, como em um ambiente Web; transformando um grande número de informações em estratégia de negócio.

Para BRETZKE (1992), quando se incorpora TI como ferramenta, significa que a base de informações sobre a qual são tomadas as decisões está sendo aperfeiçoada. O aperfeiçoamento pode ocorrer de várias formas, como:

- Sistematização de um fluxo constante e atualizado de dados sobre clientes, concorrência, ambiente, etc., para a empresa;
- Sistematização de um fluxo constante de comunicação e informações entre os clientes e a empresa;
- Utilização do computador e a tecnologia de Banco de Dados, com a finalidade de ter um grande número de dados disponíveis, integrar os dados e realizar simulações das atividades da empresa, para projetar situações e relacionamentos, transformando dados “brutos” em acessíveis e poderosas informações mercadológicas;
- Aplicação de técnicas estatísticas em Banco de Dados de Clientes e *Prospects*¹³ para:
 1. analisar o comportamento dos clientes e consumidores;
 2. identificar segmentos de mercado relativamente homogêneos;
 3. identificar individualmente clientes e *prospects* em termos de seus hábitos de consumo, necessidades, desejos e a probabilidade de prever o seu comportamento de resposta ou compra em função da ação de comunicação (no caso, um determinado *site* da Web);
 4. atuar criativamente sobre as oportunidades de mercado identificadas, para estabelecer relacionamento duradouro e rentável com os *prospects* e clientes.
- A integração dos dados internos (vendas produção, custos, etc.) aos dados externos (clientes, concorrência, mercado), criando um completo Sistema de Informação de Suporte ao *Marketing*.

O impacto da TI surge quando se une ao *marketing*, criando uma nova abordagem para o próprio *marketing*. Essa abordagem:

- Integra o cliente à empresa, facilitando o processo de inovação no desenvolvimento de personalização de produtos;

¹³ Nota da autora: *Prospects* deve ser entendido como cliente potencial, ou seja, pessoas que possuem potencial financeiro e o poder de decisão para compra.

- Permite às empresas através de customizações e personalizações de produtos e serviços a conquista de mercado, pois responde de forma individualizada as necessidades do clientes;
- Mantém um Canal Permanente de Comunicação para criar e sustentar um relacionamento efetivo com os seus clientes.

Portanto a Tecnologia de Informação de Data Warehouse juntamente com suas ferramentas, permitirá às organizações “entender” o comportamento de seus consumidores na Internet bem como realizar simulações, além de outras simulações relevantes ao processo decisório.

4.10 EXTRAINDO AS INFORMAÇÕES DE UM DATA WAREHOUSE

As informações referentes ao perfil do cliente de uma determinada empresa que opera suas vendas no comércio eletrônico, encontra-se armazenada nos bancos de dados que pode chegar a um volume de até *terabytes* de dados de *marketing*. Para INMON (1992) ainda pode existir um longo caminho a ser percorrido até que estas informações estejam de fato disponíveis. A extração da informação de forma eficaz, para subsidiar decisões, depende da existência de ferramentas especializadas que permitam a captura de dados relevantes mais rapidamente e a sua visualização através de várias dimensões.

As ferramentas segundo WELDON (1998) não devem apenas permitir o acesso aos dados, mas também permitir análises de dados significativas, de tal maneira a transformar “dados brutos” em informação útil para os processos estratégicos da empresa. O sucesso de um DW pode depender da disponibilidade da ferramenta certa para as necessidades de seus usuários.

Há várias maneiras de extrair informações de um DW, as formas de extração mais comuns no mercado são:

- Ferramentas de consulta e emissão de relatórios;
- *Executive Information Systems* (EIS);
- Ferramentas Data Mining;
- Ferramentas OLAP.

Neste trabalho serão apresentadas as ferramentas *Data Mining* e OLAP. A diferença básica entre elas está na maneira como a exploração dos dados é abordada. Com ferramentas

OLAP a exploração é feita na base da verificação, isto é, o analista conhece a questão, elabora uma hipótese e utiliza a ferramenta para confirmá-la. Com Data Mining, a

questão é total ou parcialmente desconhecida e a ferramenta é utilizada para a busca de conhecimento.

Nos próximos itens serão apresentadas essas duas ferramentas que realizam o acesso aos dados para descobrirem novos padrões e análises destes.

4.10.1 DATA MINING

Ultimamente, as empresas que operam seus negócios sob a forma B2C têm armazenado uma quantidade incrível de dados, principalmente no que se diz respeito ao perfil de seus consumidores. Manipular essas informações torna-se cada vez mais crítico ao próprio negócio. É necessário converter toda essas informações em estratégias de *marketing*. Como o DW possui bases de dados bem organizadas e consolidadas, os dados devem ser extraídos para a análise em busca de descobrir novas tendências, padrões e correlações. É neste cenário que o Data Mining está cada vez mais sendo utilizado nas grandes empresas com o intuito de explorar grandes bases de dados e descobrir relações e padrões existentes nessas informações.

O DW mantém em um único lugar uma quantidade enorme de informações históricas de uma empresa e proporciona um acesso para análise. Uma empresa que opera no comércio eletrônico pode usar essas informações para melhorar o atendimento do cliente e explorar novas oportunidades de negócios. A tecnologia de Data Mining (“garimpagem de dados”) utiliza ferramentas automatizadas para extração de dados do DW e análise dos padrões, tendências e relacionamentos.

Para HODEL (1998), Data Mining é o processo de extrair informação significativamente e previamente desconhecida que pertence a um grande banco de dados.

BRETZKE (2000), afirma que os *softwares* de Data Mining pesquisam grandes volumes de dados, procurando as características interessantes ou padrões que permitam prever acuradamente o comportamento de clientes e *prospects* em relação a certos eventos como receber uma campanha de mala direta ou propensão a comprar certos produtos. Eles utilizam técnicas sofisticadas como redes neurais, *rule-induction*, algoritmos genéricos, clusterização, entre outras para descobrir relacionamentos ocultos entre os dados. Normalmente estes softwares já estão sendo oferecidos juntos com o Data Base *Marketing* (DBM), com o DW, e no CRM.

“O termo Data Mining trata-se da procura por relações e padrões de dados em grandes bancos de dados. Como o próprio termo sugere, Data Mining possui uma orientação

exploratória para a busca do conhecimento obscurecido pelos complexos padrões de associação e pelas grandes massas de dados” (SILVA, 2000, pg 24).

De acordo com IBM (1996), *Data Mining* (ou mineração de dados) é o processo de extrair a informação válida, previamente desconhecida e de máxima abrangência a partir de grandes bases de dados, usando-as para efetuar decisões cruciais. *Data Mining* vai muito além da simples consulta a uma banco de dados, no sentido de que permite aos usuários explorar e inferir informação útil a partir dos dados, descobrindo relacionamentos “escondidos” no banco de dados. Pode ser considerada uma forma de descobrimento de conhecimento em bancos de dados (KDD - *Knowledge Discovery in Databases*), área de pesquisa de bastante evidência no momento, envolvendo Inteligência Artificial e Banco de Dados como citam KRISHNAMURTHY & IMEILINSKI (1991) e FAYAD & SHAPIRO (1996).

Para FAYAD & SHAPIRO (1996), *Data Mining* é utilizado com os seguintes objetivos:

- Explanatório: explicar algum evento ou medida observada, por exemplo como e porque a venda de um determinado produto reduziu em uma determinada região ou cidade;
- Confirmatório: confirmar uma hipótese. Uma companhia de seguros, por exemplo, pode querer examinar os registros de seus clientes para determinar se famílias de duas rendas tem mais probabilidade de adquirir um plano de saúde do que famílias de uma renda;
- Exploratório: analisar os dados buscando relacionamento novos e não previstos. Uma companhia de cartão de crédito pode analisar seus registros históricos para determinar que fatores estão associados a pessoas que representam risco para créditos.

Resumidamente, *Data Mining* é o processo de encontrar correlações ou padrões entre milhares de campos em grandes bases de dados, permitindo agilidade na tomada de decisões.

No comércio eletrônico, assim como no tradicional, a ferramenta de *Data Mining* é capaz de:

- Criar parâmetros para entender o comportamento do consumidor;
- Identificar afinidades entre as escolhas de produtos e serviços;
- Prever hábitos de compras;
- Analisar comportamentos habituais para se detectar fraudes.

Nota-se que grandes quantidades de dados são “entregues” pelos usuários para a ferramenta de *Data Minings* em busca de tendências ou agrupamentos dos dados.

Ferramentas de Data Mining utilizam-se das mais modernas técnicas de computação, como redes neurais, descoberta por regra, detecção de desvio, programação genética, para extrair padrões e associações de dados.

Para SANTOS & HENRIQUES & REIS (2000), a análise automatizada e antecipada oferecida por Data Mining, vai muito além da simples análise de eventos passados, que é fornecida pelas ferramentas de retrospectiva típicas de sistemas de apoio à decisão como consultas SQL do universo transacional (OLTP).

De acordo com DW (2000), as três grandes áreas de uso de um Data Mining, são:

- **Cross-selling:** Identificar associação entre produtos, procurando tirar proveito dessas associações;
- **Up-selling:** Identificar entre os clientes atuais aqueles que tem potencial para adquirir produtos que proporcionem maior lucro para a companhia;
- **Fidelização:** Combinar e descobrir fatores que causem a perda de clientes, e como forma economia, manter estes clientes, o que é mais barato que conquistar novos;

De acordo com NETO (2001), as tarefas do *Data Mining* são classificadas em seis categorias:

1. **Estimativa e predição:** estimativa consiste em examinar atributos de um conjunto de entidades e, baseado nos valores destes atributos, assinalar valores a atributos de uma nova entidade que se quer caracterizar. O termo predição é usado quando a estimativa é usada para prever o futuro valor de um atributo;
2. **Classificação:** consiste em examinar os atributos de uma determinada entidade;
3. **Descoberta de associações:** consiste em identificar quais atributos estão associados com outros em um determinado ambiente;
4. **Descoberta de aglomerações (clustering):** consiste em segmentar uma população heterogênea em subgrupos homogêneos de entidades. Para INMON (1997) os resultados de uma operação de clusterização podem ser usados de duas diferentes maneiras: para produzir um sumário da base de dados ou como dados de entrada para outras técnicas;
5. **Visualização de dados:** é a tarefa de descrever informações complexas através de telas visuais (diagramas). Para INMON (1992) As técnicas de visualização não são propriamente técnicas de Data Mining, são meios de analisar e observar os dados de uma determinada base de dados de forma gráfica, fornecendo, desta forma, meios de obter sumários visuais dos dados de uma base de dados;

6. **Exploração Interativa de dados:** é o processo de inspecionar grande volume de dados através de controles interativos que permitem rapidamente inspecionar conjunto de dados de interesse a um perito.

O Data Mining faz uso da simulação, que representa a ação de construir um modelo, ou cenário, de uma situação onde se sabe a resposta e aplicá-lo a outra situação onde a resposta é desconhecida. Aplicando-se técnicas para executar a simulação é possível achar padrões relevantes e de interesse do usuário. Dentre as técnicas utilizadas, WELDON (1998) explana algumas, dentre elas:

- **Lógica baseada em casos:** consiste na derivação de regras a partir de estudos de caso.
- **Descoberta de regras:** consiste na execução de algoritmos de análise de dados em grandes massas de dados na busca de padrões e correlações que possam subsidiar a formulação de regras. A busca pode ser dirigida (procura por dados para apoiar uma determinada regra) ou não dirigida (permite que padrões de dados sugiram possíveis regras).
- **Pontuação:** utilizando um esquema de pontuação, dados históricos podem ser analisados e uma árvore de decisão construída baseada em um conjunto de valores;
- **Processamento de sinais:** técnicas de processamento de sinais, tal como filtragem digital, podem identificar ilustres de observações com características similares.
- **Fractais:** consiste da utilização de fractais para comprimir grandes bancos de dados sem perda de informação - permite a análise sobre todo o universo do banco de dados com um tempo de resposta surpreendente.
- **Redes Neurais:** são modelos de previsão baseados e, princípios similares àqueles do pensamento humano. Em uma rede de nós, cada nó recebe entrada e envia uma saída para nós subsequentes baseado no que recebeu. A rede é "treinada" usando uma amostra de dados para determinar "pesos" apropriados para cada nó. Ela então produz valores específicos para dados subsequentes, conforme **FIGURA 24**. Uma rede neural é organizada em camadas. Usualmente as camadas são classificadas em três grupos:
 - **Camada de Entrada:** onde os padrões são apresentados à rede;
 - **Camadas Intermediárias ou Escondidas:** onde é feita a maior parte do processamento, através das conexões ponderadas; podem ser consideradas como extratoras de características;

- Camada de Saída: onde o resultado final é concluído e apresentado.

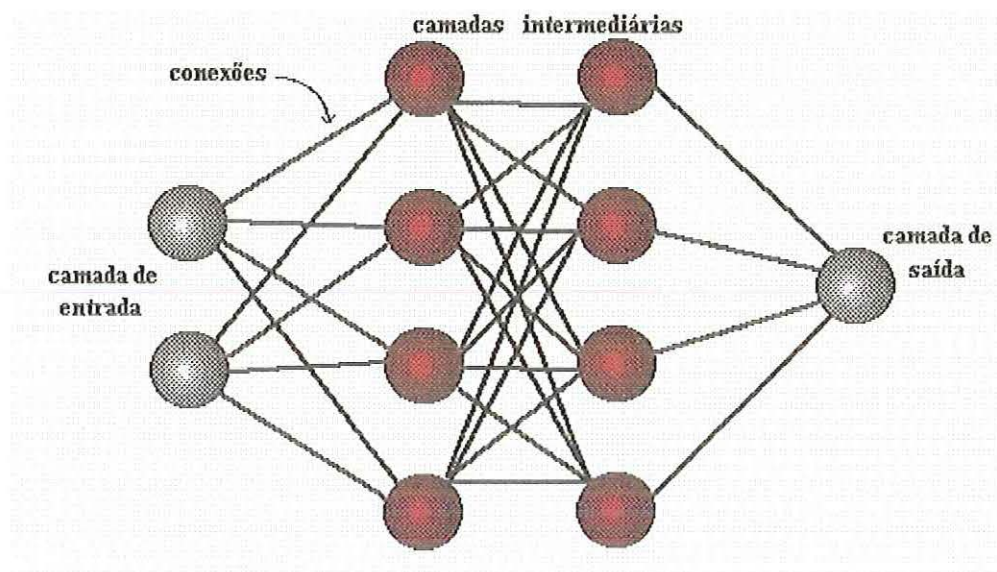


Figura 24: Exemplo de uma Rede neural
 Fonte : <http://www.br-business.com.br>

Segundo DW (2000) as técnicas de Data Mining têm sido aplicadas com sucesso para a solução de problemas em diversas áreas, como:

Vendas:

- Identificar padrões de comportamento dos consumidores;
- Encontrar características dos consumidores de acordo com a região demográfica ;
- Prever quais consumidores serão atingidos nas campanhas de *marketing*.

Finanças:

- Detectar padrões de fraudes no uso dos cartões de crédito;
- Identificar os consumidores que estão tendendo a mudar a companhia do cartão de crédito;
- Identificar regras de estocagem a partir dos dados do mercado;
- Encontrar correlações escondidas nas bases de dados.

Seguros e Planos de Saúde:

- Determinar quais procedimentos médicos são requisitados ao mesmo tempo;
- Prever quais consumidores comprarão novas apólices;
- Identificar comportamentos fraudulentos.

Transporte:

- Determinar a distribuição dos horários entre os vários caminhos;

- Analisar padrões de sobrecarga.

Medicina:

- Caracterizar o comportamento dos pacientes para prever novas consultas;
- Identificar terapias de sucesso para diferentes doenças.

Finalizando, SILVA (2000) afirma que sob certos aspectos, o comércio eletrônico se assemelha a venda por catálogos e varejo. Portanto é de se esperar que o sucesso obtido na aplicação de *Data Mining* nestes casos possa ser replicado na Internet. Certas características da Internet a fazem um meio muito mais eficiente e interativo que a venda por catálogos, pois na Internet é possível acompanhar a ordem de compra dos produtos em um *site*, e sugerir ao final outros produtos que clientes do mesmo perfil geralmente também compram. Na Internet também, a aquisição dos dados é imediata, gerando uma massa de dados crítica e já estruturada para processamento.

4.10.2 OLAP - Processamento Analítico On-line

Segundo INMON (1997), OLAP refere-se ao tipo de processamento e ferramentas voltados para análise de dados típica do suporte à decisão, onde os dados são apresentados através de uma visão multidimensional. Esta visão é independente de como os dados estão armazenados.

OLAP é uma ferramenta utilizada pelos usuários finais para a análise dos dados extraídos do DW. A análise pode ser feita de forma personalizada para cada tipo de negócio e o processo de consulta é interativo, possibilitando ao usuário analisar o “porquê” dos resultados obtidos e até mesmo descobrir padrões “escondidos”.

WELDON (1998) afirma que OLAP sempre envolve consultas interativas aos dados, seguindo um caminho de análise através de múltiplos passos, por exemplo: aprofundar-se sucessivamente por níveis mais baixos de detalhe de um quesito de informação específico. OLAP envolve capacidades analíticas, incluindo a derivação de taxas, variâncias, etc., e envolvendo medidas ou dados numéricos através de muitas dimensões, devendo suportar modelos para previsões, análises estatísticas e de tendências. A FIGURA 25 exemplifica um OLAP.

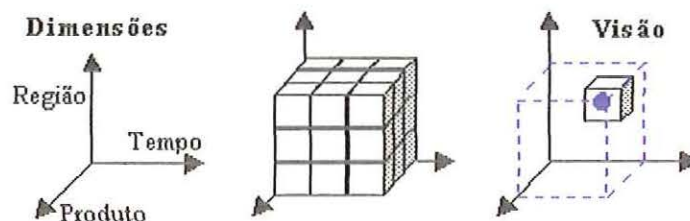


Figura 25: Exemplo de um OLAP

Fonte: <http://www.datawarehouse.da.ru>

Segundo BISPO (1999), a característica principal dos sistemas OLAP é permitir uma visão conceitual multidimensional dos dados de uma empresa. A visão multidimensional é muito mais útil para os analistas do que a tradicional visão tabular utilizada nos sistemas de processamento de transação. Ela é mais natural, fácil e intuitiva, permitindo a visão em diferentes perspectivas dos negócios da empresa e desta maneira tornando o analista um explorador da informação.

A modelagem dimensional é a técnica utilizada para se ter uma visão multidimensional dos dados. Nesta técnica os dados são modelados em uma estrutura dimensional conhecida por cubo. As dimensões do cubo representam os componentes dos negócios da empresa tais como "cliente", "produto", "fornecedor" e "tempo". A célula resultante da interseção das dimensões é chamada de medida e geralmente representa dados numéricos tais como "unidades vendidas", "lucro" e "total de venda". Além dos componentes dimensão e medida outro importante aspecto do modelo multidimensional é a consolidação dos dados uma vez que para a tarefa de análise são mais úteis e significativas as agregações (ou sumarização) dos valores indicativas dos negócios. Além da visão multidimensional dos dados da empresa, outras importantes características dos sistemas OLAP são (TEC99):

- Análise de tendências. A tecnologia OLAP é mais do que uma forma de visualizar a história dos dados, ela auxilia os usuários a tomar decisões sobre o futuro, permitindo a construção de cenários ("e se...") a partir de suposições e fórmulas aplicadas, pelos analistas, aos dados históricos disponíveis;
- Busca automática de dados mais detalhados que não estão disponíveis no servidor OLAP. Detalhes não são normalmente importantes na tarefa de análise, mas quando necessários, o servidor OLAP deve ser capaz de buscá-los;
- Dimensionalidade genérica;
- Operação trans-dimensional. Possibilidade de fazer cálculos e manipulação de dados através diferentes dimensões;

- Possibilidade de ver os dados de diferentes pontos de vista (*slice and dice*)¹⁴, mediante a rotação do cubo e a navegação (*drill-up/drill-down*)¹⁵ entre os níveis de agregação;
- Conjunto de funções de análise e cálculos não triviais com os dados.

Para ORR (1996), OLAP apresenta características distintas do OLTP, conforme mostra a TABELA 6.

OLTP	OLAP
Relacional	Multidimensional
Individualizados	Sumarizados
Presente	Histórico
Um registro de cada vez	Muitos registros por vez
Orientados ao processo	Orientados ao negócio

Tabela 6: OLAP versus OLTP
 Fonte: <http://www.kenorinst.com/>

BISPO (1999) apresenta uma lista de 12 regras para avaliar a eficácia de uma ferramenta de OLAP com relação às necessidades deste tipo de processamento:

1. Visão conceitual multidimensional: enfatiza a forma como o usuário "vê" os dados sem impor que eles sejam armazenados em formato multidimensional;
2. Transparência: localização da funcionalidade OLAP deve ser transparente para o usuário, assim como a localização e a forma dos dados;
3. Facilidade de Acesso: acesso a fontes de dados homogêneas e heterogêneas devem ser transparente;
4. Desempenho de consultas consistente: não deve ser dependente do número de dimensões;
5. Arquitetura cliente/servidor: produtos devem ser capazes de operar em arquiteturas cliente/servidor;
6. Dimensionalidade genérica: todas as dimensões são iguais;
7. Manipulação dinâmica de matrizes esparsas: produtos devem lidar com matrizes

¹⁴ Nota da autora: uma das características da ferramenta OLAP. Proporciona analisar as informações por diferentes prismas. Utilizando essa tecnologia as informações podem ser vistas por diferentes ângulos.

¹⁵ Nota da autora: consiste em realizar uma exploração em diferentes níveis de detalhes das informações. Pode-se "subir ou descer" dentro do detalhamento da informação, como por exemplo analisar uma informação específica de uma fotografia do planeta Terra, como a visualização de um determinado ponto da floresta Amazônica, no Brasil.

- esparsas eficientemente;
- 8. Suporte multi-usuário;
- 9. Operações entre dimensões sem restrições;
- 10. Manipulação de dados intuitiva;
- 11. Relatórios/consultas flexíveis;
- 12. Níveis de agregação e dimensões ilimitados: ferramentas devem ser capazes de acomodar de quinze a vinte dimensões.

Para DAL'ALBA (1998), a arquitetura OLAP possui três componentes principais:

- um modelo de negócios para análises interativas, implementado numa linguagem gráfica que permita diversas visões e níveis de detalhes dos dados;
- um motor OLAP para processar consultas multidimensionais contra o dado-alvo;
- um mecanismo para armazenar os dados a serem analisados.

A base de dados usada define se o pacote é um ROLAP, que realiza a interface com um banco de dados relacional de mercado, ou um MOLAP, que se liga a um servidor OLAP, através de um banco de dados multidimensional e dedicado.

✓ De acordo com DW (2000), as categorias de ferramentas OLAP são:

- **MOLAP:** Multidimensional OLAP (MOLAP) é uma classe de sistemas que permite a execução de análises sofisticadas usando como gerenciador de dados um banco de dados multidimensional. Em um banco de dados MOLAP os dados são mantidos em arranjos e indexados de maneira a fornecer uma melhor performance no acesso a qualquer elemento. Além de serem rápidos, outra grande vantagem destes sistemas é o rico e complexo conjunto de funções de análise que oferecem. A maneira de se implementar os arranjos de dados pode variar entre fornecedores de soluções MOLAP. É utilizado, tradicionalmente para organizar, navegar e analisar dados. Basicamente a arquitetura MOLAP processa-se da seguinte forma: com um servidor multidimensional, o acesso aos dados ocorre diretamente no banco, ou seja, o usuário trabalha, monta e manipula os dados do cubo diretamente no servidor. Isto traz grandes benefícios aos usuários em relação à performance, mas apresenta custos elevados para sua aquisição.
- **ROLAP:** *Relational* OLAP, permite que múltiplas consultas multidimensionais de tabelas bidimensionais relacionais sejam criadas sem a necessidade de estrutura de dados normalmente requerida neste tipo de consulta. Neste caso, a consulta é enviada ao servidor de banco de dados relacional e processada no mesmo, mantendo o cubo no servidor. Neste caso o processamento OLAP se dará somente no servidor. A principal vantagem dessa arquitetura é que ela permite

analisar grandes volumes de dados, em contra partida se uma grande quantidade de usuários acessar simultaneamente poderá causar sérios problemas de performance no servidor causando, inclusive, o travamento do mesmo.

- **DOLAP:** essa arquitetura, Desktop OLAP, é composta de Sistema Gerenciador de Banco de Dados (SGDB) com a ferramenta DOLAP. O processamento OLAP acontece na máquina cliente, sem tráfego de rede nem problemas de escalabilidade. Contudo pode trazer problemas em alguns relatórios, quando o volume de dados fica extenso.
- **HOLAP:** recentemente surgiu essa outra arquitetura, denominada *Hybrid OLAP*, ou simplesmente processamento híbrido. Essa nova forma de acessar os dados consiste na mistura de tecnologias onde há uma combinação entre ROLAP e MOLAP. A vantagem é que com essa “mistura” de tecnologias pode-se extrair o que há de melhor de cada uma, ou seja, a alta performance do MOLAP com a escalabilidade do ROLAP.

Portanto nota-se que apesar de outras ferramentas possibilitarem a visualização dos dados, como o Excel ou o Access, as mesmas não apresentam nem a performance adequada nem a capacidade de ver os dados multidimensionalmente, como *slice-and-dice* e *drill* que são os recursos mais importante do OLAP para analisar as informações. Por outro lado, OLAP não é uma solução imediata, configurar o programa de OLAP e ter acesso aos dados requer uma clara compreensão dos modelos de dados da empresa e das funções analíticas necessárias aos executivos e à outros analistas de dados.

4.10 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Os bancos de dados são de vital importância às empresas, bem como a análise desses dados. As empresas detêm um volume enorme de dados e esses estão em diversos sistemas diferentes espalhados por ela. A tomada de decisões embasadas num histórico de dados era impossibilitada. Só a partir do surgimento da ferramenta DW que essa problemática começou a ser “eliminada”.

Foi apresentado por diferentes autores que o DW surgiu como uma ferramenta capaz de consolidar dados distribuídos em bases de dados diferentes. É abordado também que o DW é um processo e não um produto, surgindo para apoiar a tomada de decisão em uma empresa.

A ferramenta DW não surgiu para substituir os sistemas operacionais existentes, mas para complementá-los. Para melhor compreensão, as ferramentas do DW foram explicadas com o intuito de evidenciar que os dados em um DW são dados consolidados historicamente, que devem ser “trabalhados” e transformados em informações proveitosas para o responsável pela tomada de decisão. Ferramentas essas que devem ser escolhidas de acordo com a necessidade da informação desejada.

A modelagem dos dados de um DW (modelo dimensional) difere do modelo relacional. Por apresentar uma modelagem simples e de fácil compreensão, os analistas projetam de acordo com a necessidade do negócio e não de acordo com o que acontece, que é realizado no modelo relacional.

O domínio sobre essa tecnologia permite aos administradores descobrir novas maneiras de diferenciar sua empresa numa economia globalizada, deixando-os mais seguros para definirem as metas e adotarem diferentes estratégias em sua organização, conseguindo assim visualizar antes de seus concorrentes, novos mercados e oportunidades atuando de maneiras diferentes conforme o perfil de seus consumidores.

É desta forma que a utilização desta ferramenta, têm mostrado alto interesse por administradores que operam no comércio eletrônico.

A ferramenta além de seguir suas mesmas características do comércio tradicional, é capaz de integrar dados comportamentais de clientes que “navegam” durante o processo de compra de produtos ou serviços em um determinado *site* de comércio eletrônico.

Portanto a utilização de ferramenta como o DW no comércio eletrônico é capaz dentre outras coisas de “interpretar” o comportamento de um cliente que está “navegando” em um determinado *site* do comércio eletrônico. Neste foco, a ferramenta será capaz de contemplar as análises de campanhas de *marketing*, perfil do cliente, análise de vendas, lealdade, entre outros.

É de extrema importância que a ferramenta seja bem estruturada de acordo com a necessidade de cada empresa: tanto a empresa que opera no comércio tradicional, como a empresa que opera no comércio eletrônico. Só assim o DW pode ser usado como uma poderosa ferramenta que fornece informações essenciais aos tomadores de decisões. Informações essas que são consideradas fontes de dados para as ferramentas de extração de dados.

Muitas organizações que operam de forma B2C têm investido em tecnologias de informação para ajudar na gerência dos negócios. É nesse cenário que Data Mining está cada vez mais sendo utilizado pois é capaz de selecionar, explorar e modelar grandes conjuntos de dados para detectar padrões de comportamento dos consumidores. Este tipo de ferramenta de

extração é capaz de revelar padrões e relacionamentos “escondidos”, gerando hipóteses e testando-as. *Data Mining* é tido como um conceito de análise, pois através de informações históricas utilizando algoritmos estatísticos descobre o comportamento de um cliente tornando-se então, uma poderosa ferramenta de auxílio nas estratégias de *marketing*. A análise de dados deve ser:

- Inserida dentro de um DW (a fim de reduzir erros e custos);
- Entendida e utilizável para um profissional de *marketing*.

Já a ferramenta OLAP, compara os dados e auxilia o entendimento dos fatos. É capaz de responder perguntas e verificar hipóteses. A ferramenta possui recursos para a disponibilização do suporte à decisão e estratégia de negócio, obtendo de forma mais rápida, consistente e interativa uma variedade de visualizações da informação, facilitando a análise multidimensional.

Portanto, o valor de um DW está na qualidade das decisões que ele suporta. Ferramentas poderosas de análise de informações como OLAP e *Data Mining*, quando incorporadas a uma arquitetura de DW, aumentam a competitividade das empresas.

No entanto, a utilização destas ferramentas, ainda requerem um bom nível de conhecimentos - do domínio da aplicação, de estatística e da própria ferramenta.

5. MODELO INTEGRADO

5.1 CONSIDERAÇÕES INICIAIS

A proposta de um modelo integrado surge a partir da necessidade que os responsáveis pelas organizações que atuam no comércio eletrônico vêm se deparando. Tal necessidade é a de garantir a fidelidade dos principais clientes em um determinado *site* de comércio eletrônico. E para garantir fidelidade, a ferramenta DW juntamente com a estratégia de CRM surgem como uma “facilitadora” ou até mesmo uma “solucionadora” para essa determinada necessidade. A partir da mineração e análise dos dados através de ferramentas específicas como ferramentas OLAP e Data Mining, os dados integrados existentes no DW fornecerão informações consolidadas historicamente. Essas informações se forem bem “trabalhadas” tornarão diferencial competitivo para a empresa.

O objetivo deste capítulo é apresentar a funcionalidade e o gerenciamento do DW capacitado para a Web proposto por KIMBALL & MERZ (2000), ou seja, a utilização do DW como uma ferramenta capaz de integrar dados comportamentais de usuários do comércio eletrônico, aplicando a estratégia de CRM.

5.2 INTRODUÇÃO

Este capítulo inicia-se com um breve comentário referente aos autores pioneiros na implementação de um DW para a Web.

Ralph Kimball, tem sido um líder visionário na indústria de DW desde 1982 e, atualmente, é um dos mais conhecidos conferencistas, consultores e professores no assunto. Ele escreve a coluna “Webhouse Architect” para a revista *Intelligent Enterprise* e é o autor dos *best-sellers* “The Data Warehouse Toolkit” e “The Data Warehouse Lifecycle Toolkit”.

Richard Merz é diretor de engenharia de divisão *WebCom* da *Verio*, a maior companhia de *Web hosting* do mundo. Possui uma formação prática muito extensa em arquitetura e aplicativos de DW. Vem gerenciando o desenvolvimento de software de Web e de comércio eletrônico nos últimos cinco anos.

A idéia de se integrar TI ao *marketing* também foi profundamente abordada por Miriam Bretzke em BRETZKE (1999). A autora descreve que o maior potencial da integração entre *marketing* e TI está em :

- Aumentar o conhecimento sobre o cliente;

- Registrar, manter e tratar um grande volume de informações sobre o cliente e recuperá-lo *on line* e em tempo real para estabelecer um relacionamento diferenciado;
- Aumentar a retenção e fidelidade do cliente pela experiência de marca positiva;
- Aperfeiçoar o processo decisório;
- Aperfeiçoar os serviços para clientes/consumidores;
- Diferenciar seus serviços/produtos no mercado;
- Estabelecer uma comunicação próxima e sistemática com os clientes potenciais e atuais.

BRETZKE (1999) ainda argumenta que a integração entre TI e *marketing* permite realizar o planejamento e o controle das estratégias de prospecção e fidelização partindo de análises mais acuradas do comportamento dos clientes. Pode-se inclusive identificar de forma individual clientes em termos de seus hábitos de consumo, necessidades, desejos e a probabilidade de prever seu comportamento de resposta ou compra em função de alguma ação, etc. O impacto causado pela “união” de TI ao *marketing* pode ser transformador, pois integra o cliente à empresa como um participante do processo de desenvolvimento e adaptação de serviços e produtos, facilitando o processo de inovação no desenvolvimento de novos produtos; permite à empresa conquistar o mercado pois possibilita personalização, ou seja, responder de forma individualizada às necessidades dos clientes; mantém um canal permanente de comunicação para criar e sustentar um relacionamento efetivo com os clientes.

Portanto a função do DW aliado ao CRM no comércio eletrônico será inicialmente de consolidar e analisar dados de negócio brutos e então transformá-los em informações claras e conclusivas; capacitar empresas a ter acesso em fontes diversas e consolidadas de cliente, dados de mercado e operacionais e, então, usar esta informação para obter uma vantagem competitiva; fornecer às empresas o conhecimento necessário para descobrir novas tendências mercadológicas; ampliar relacionamentos, reduzir riscos financeiros e criar novas oportunidades de venda. A forma como isso é feito é mostrado nos próximos itens desse capítulo.

5.3 DIMENSÕES DA SEQUÊNCIA DE CLIQUES

Antes de se projetar Data Marts específicos de sequência de cliques, deve-se coletar o maior número de dimensões que possam ter relevância no ambiente de sequência de cliques. Qualquer “esquema estrela” específico não usará todas as dimensões descritas a

seguir, mas segundo KIMBALL & MERZ (2000), é de grande necessidade ter-se uma “pasta” de dimensões prontas esperando para ser utilizadas. Para os autores, esse levantamento de tabelas dimensionais faz com que a equipe desenvolvedora de Data Mart realize um processo de simulação.

5.3.1 DIMENSÃO DATA DO CALENDÁRIO

Nesta dimensão existe um registro para cada dia no calendário. É utilizada sempre que há um registro automático de data ou de data/hora. No caso do registro data/hora, o dia do calendário é separado do horário do dia e, a partir daí as dimensões são separadas. O motivo da separação e conseqüentemente da criação de duas tabelas dimensionais é devido ao fato de que não é prático construir uma única dimensão para todos os minutos ou segundos ou milésimos de segundos em um período de vários anos. Um outro motivo é que mesmo se é tentado construir tal dimensão única, não há muito sentido do ponto de vista dos aplicativos, pois há nomes e rótulos para os dias e para períodos de tempo dentro dos dias, e esses nomes e rótulos estão, de certa forma, correlatos. São por esses motivos que datas/horas são divididas em duas dimensões. A TABELA 7 ilustra a dimensão de data do calendário.

ATRIBUTO	DESCRIÇÃO
Chave da data	Valores substitutos, I..N
Tipo de data	Regular, desconhecido, não medido, corrompido, inaplicável, não acontecido
Tipo de calendário	Gregoriano, japonês, chinês, islâmico, talmúdico
Data completa de SQL	Valor legítimo se tipo = regular, nulo caso contrário
Dia da semana	Segunda-feira, Terça-feira etc.
Número do dia da semana	1..7
Número do dia no mês	1..31
Número do dia no ano	1..366
Número do dia na época	Um número juliano começado em algum ponto poderia ser negativo
Dia útil	Dia útil, feriado – deve-se utilizar as palavras ao invés de sim/não para efeitos de geração

	de relatórios
Feriado	Nenhum ou o nome do feriado
Dia da semana	Dia da semana, fim-de-semana – deve-se utilizar as palavras ao invés de sim/não
Ultimo dia do mês	Dia normal, último dia, utilizado como uma conveniência de restrição
Número da semana no ano	1..53, a semana nº1 é a primeira semana inteira do ano
Número da semana na época	Alinha com o número da semana do ano
Mês	Janeiro, Fevereiro etc
Abreviação	Jan, Fev etc
Número do mês no ano	1..12
Número do mês na época	Alinha com o número do mês
Trimestre no ano	4Q1999, 1Q2000 etc
Ano	Não empregar a notação com dois dígitos nesse campo
Período fiscal	A designação de período fiscal poderia conter vários campos
Estação do ano	Estação do ano, período de férias, poderia conter vários campos
Evento	Indicador especial para um período único e digno de menção no tempo

Tabela 7: Dimensão de Data do Calendário
 Fonte: KIMBALL & MERZ, 2000, pg 131

A chave para a dimensão da data do calendário e para todas as outras dimensões de outros projetos é uma chave substituta, ou seja, a chave é um inteiro anônimo. O primeiro registro nessa dimensão pode assumir o valor de “chave=1”, o segundo terá o valor de “chave=2”, e assim por diante. De acordo com KIMBALL & MERZ (2000), o ponto a favor de uma chave substituta é que não há nenhuma semântica no valor da chave, pois é somente um veículo de unir a tabela de dimensão a uma tabela de fatos.

As chaves substitutas são criadas por situações administrativas importantes que só acontecem no DW. A primeira situação que surge no DW é se o objeto sendo referido é desconhecido, não medido, inaplicável, corrompido ou ainda não acontecido. Por exemplo, se uma data é desconhecida, então não se deve atribuir alguma data real futura que será

interpretada pelos aplicativos como “desconhecida”, em vez da data identificada. Deve-se separar os registros regulares dos registros de casos de exceção referindo-se a uma chave de *tipo* (tipo de data, nesta situação) que será localizada em todas as tabelas de dimensão.

Uma segunda situação para a existência de chaves substitutas é que elas isolam o DW de alterações na administração de chaves nos dados de produção. Essa pode não ser uma questão importante com datas, mas é com a maioria das outras dimensões.

Uma situação final para as chaves substitutas é a necessidade de tratar dimensões que se alterem lentamente. Por exemplo, deseja-se revisar a descrição de um cliente. Talvez o cliente tenha se mudado ou agora esteja casado. Reluta-se em extinguir o perfil antigo, pois o mesmo era verdadeiro e válido em um determinado momento no tempo. Mas se é utilizado algum tipo de ID de cliente de produção para a chave da dimensão do cliente, os analistas se defrontam com o dilema de sobrescrever o registro de dimensão. Para solucionar tal problema, KIMBALL & MERZ (2000) propõem a utilização de uma chave substituta e a publicação de um novo registro de cliente quando for preciso.

5.3.2 DIMENSÃO DO HORÁRIO DO DIA

A dimensão horário do dia acompanha a dimensão da data do calendário. Essa é uma dimensão simples que permite referenciar períodos identificados de tempo durante o dia. Horas individuais são o exemplo mais comum. Dependendo do negócio em que a dimensão se insere, pode-se ter nomes para outras faixas de tempo, como: turno da noite, horário de almoço, períodos de pico etc. Para KIMBALL & MERZ (2000) a dimensão horário do dia pode ser construída no nível do minuto ou mesmo do segundo. Há 1.440 minutos e 86.400 segundos em um dia. A TABELA 8 ilustra a dimensão horário do dia.

ATRIBUTO	DESCRIÇÃO
Chave de data/hora	Valores substitutos, 1..N
Tipo de data/hora	Por exemplo, regular, desconhecido, não medido, corrompido, inaplicável, não acontecido
Segundos desde a meia-noite	0..86399
Minutos desde a meia-noite	0..1439
Registro de data de SQL	Por exemplo, 13:30:58
Hora	0..23
Minuto	0..59

Segundo	0..59
Período de Tempo	Turno da noite, horário de almoço etc. possivelmente vários campos para propósitos diferentes

Tabela 8: Dimensão de Horário do Dia
Fonte: KIMBALL & MERZ, 2000, pg 134

A dimensão do horário do dia também é bem pequena. Nota-se que não é preciso especificar o fuso horário para essa dimensão e nem para a dimensão data do calendário. O fuso horário é muito importante mas é especificado em outras dimensões.

5.3.3 DIMENSÃO CLIENTE

Usuário, visitante ou até mesmo navegador é considerado cliente, pelo simples fato de que a “figura” humana em um determinado *site* de comércio eletrônico acaba sofrendo de certa forma uma influência de valor de compra para a empresa. É por esse motivo que o termo “cliente” é um termo genérico que engloba todo o tipo de “figura” humana. Esta dimensão é a mais importante no Data Webhouse. Um dos desafios enfrentados num ambiente de Webhouse é o de conhecer exatamente quem é que está clicando em um determinado *site* da Web. A dimensão cliente é construída de modo que ela possa ser estendida conforme o que o cliente representará para a empresa com o decorrer do tempo. Os clientes são agrupados de acordo com as situações. A [TABELA 9](#) até a [TABELA 12](#) indicam a dimensão cliente.

Campos do grupo 1 – clientes sempre presentes em um hit de site da Web:

ATRIBUTO	DESCRIÇÃO
Chave de cliente	Valores substitutos, 1..N
Tipo de cliente	Por exemplo, regular, desconhecido, não medido, corrompido, inaplicável, endereço de IP efêmero, endereço de IP constante, <i>cookie</i> de máquina efêmero, <i>cookie</i> constante de máquina, não-cliente identificado, cliente identificado
Endereço de ISP	De diversos valores, visto que o cliente pode conectar do trabalho, de casa ou quando

	estiver em viagem: necessita de tabela associativa entre tabela de endereço de ISP e a dimensão do cliente
ID de <i>cookie</i>	O mesmo ID de <i>cookie</i> para conexão do trabalho, de casa ou em viagem
Registro de data da última alteração	Data de SQL de última variável para esse registro
Última razão de alteração	Conjunto pré definido de razões da alteração

Tabela 9: Dimensão do Cliente – grupo 1
Fonte: KIMBALL & MERZ, 2000, pg 135

Campos do grupo 2 – tem-se um nome útil de cliente e atribui-se um ID de cliente:

ATRIBUTO	DESCRIÇÃO
ID de cliente	ID atribuído
Identificador	Nome real, pseudônimo
Tipo de nome	Pseudônimo, nome real não verificado, nome real verificado
Saudação	Sr., Sra., Srta.
Primeiro nome	Primeiro nome
Nomes intermediários	Nomes intermediários
Sobrenome	Sobrenome
Estilo cultural de saudação	Norte-americano, europeu, latino, islâmico, japonês, chinês etc.
Sexo	Masculino, feminino, não quis declarar, desconhecido
Cidade de IP	Cidade derivada do endereço de IP do cliente
Estado de IP	Estado ou área dentro do país derivados do endereço de IP
País de IP	País derivado do endereço de IP

Tabela 10: Dimensão do Cliente – grupo 2
Fonte: KIMBALL & MERZ, 2000, pg 135-136

Campos do grupo 3 – se o cliente compartilhou o endereço básico e demográfico

ATRIBUTO	DESCRIÇÃO
Tipo de cliente	Residencial, comercial
Formação	Por exemplo, Mestre, Doutros e outros título honorários
Etnia	-
Nome da empresa	Esse e os cinco campos seguintes nulos se não for um cliente comercial
Departamento	Departamento
Cargo	Cargo
Telefone primário	Telefone primário
FAX primário	FAX primário
Endereço eletrônico	Endereço eletrônico
Site da Web	Site da Web
Construção	Construção
Andar	Andar
Mail Stop	Mail Stop
Tipo de endereço	Tipo de endereço
Utilização pretendida	Utilização pretendida
Número da rua	Número da rua
Nome da rua	Nome da rua
Direção da rua	Direção da rua
Caixa de correio	Caixa de correio
Localidade	Localidade
Cidade	Cidade
Estado	Estado
Código postal	Código postal
Região	Região
País	País
Bloco montado de endereço	Um endereço postal completo e válido na seqüência correta com quebras de linhas
Moeda	Tipo de moeda utilizada por esse indivíduo nessa localização
Estado civil	Casado, solteiro, não quis declarar,



	desconhecido etc.
Faixa etária	Um conjunto de <i>brackets</i> de faixas etárias pré definidas
Profissão	Um conjunto de perfis profissionais pré definidos, também opcionalmente fornecido
Interesses	Isso pode ocupar muitos formulários e consistir em vários campos
Saudação escrita	Saudação escrita montada no estilo cultural correto
Saudação verbal	Saudação verbal montada no estilo cultural correto
ID individual pessoal	Identificador particular e único amarrando registros que se referem a um mesmo indivíduo
ID comercial de entidade	Identificador particular e único amarrando registros que se referem a uma mesma entidade comercial

Tabela 11: Dimensão do Cliente – grupo 3
 Fonte: KIMBALL & MERZ, 2000, pg 136-137

Campos do grupo 4 – se o cliente jamais fez uma compra

ATRIBUTOS	DESCRIÇÃO
Recentidade	Data da última compra
Frequência	Número de compras, tempo de vida
Intensidade	Valor total de compras, tempo de vida
Valor projetado de estimativa de vida	Valor total projetado de estimativa de vida
Grupo	Um ou mais campos rotulando o grupo demográfico total do cliente
Perfil de compra	Um ou mais campos que descrevem o perfil da compra do cliente
Perfil de crédito	Um ou mais campos que descrevem o perfil de crédito do cliente
Perfil de devolução	Um ou mais campos que descrevem a propensão do cliente para devolver produtos
Perfil de suporte on-line	Um ou mais campos que descrevem a

	utilização do suporte on-line pelo cliente
Perfil de suporte pelo telefone	Um ou mais campos que descrevem a utilização do suporte pelo telefone pelo cliente

Tabela 12: Dimensão do Cliente – grupo 4
 Fonte: KIMBALL & MERZ, 2000, pg 137

5.3.4 DIMENSÃO DA PÁGINA

A dimensão da página tem por objetivo descrever o contexto de página para um evento de página da Web. O grão (um único registro em uma tabela de fatos) dessa dimensão é o *tipo* de página individual. A definição de página deve ser suficientemente flexível para tratar da evolução das páginas da Web, partindo da atual entrega de página principalmente estática para alcançar a entrega de página altamente dinâmica na qual a página exata que o cliente vê é única naquele momento do tempo. Isso não quer dizer que é criado um registro de página para cada instância de uma página dinâmica, pois isso geraria uma dimensão com um número elevado de registros e, estes, não difeririam de maneira interessante. O que se pretende é um registro nessa dimensão para distinguir cada tipo de página interessante. Páginas estáticas possuem seu próprio registro, mas as dinâmicas devem ser agrupadas por função e tipo semelhantes. Se a definição de uma página estática mudar é porque o webmaster a alterou – o registro na dimensão pode ser tanto sobrescrito como duplicado. De acordo com KIMBALL & MERZ (2000), essa decisão é uma questão de política para o Data Webhouse e, depende das descrições antigas e novas da página deferirem materialmente ou não e, da definição antiga ser ou não mantida por razões de análise histórica. A TABELA 13 descreve o projeto recomendado para a dimensão da página.

ATRIBUTOS	DESCRIÇÃO
Chave de página	Valores substitutos, 1..N
Fonte da página	Estática, dinâmica, desconhecida, corrompida, inaplicável
Função de página	Portal, pesquisa, descrição de produto, informações corporativas, etc.
Modelo de página	Esparso, denso e outros
Tipo de item	SKU do produto, número de ISBN do livro, tipo de tarifa da companhia telefônica

Tipo de imagens gráficas	GIF, JPG, revelação progressiva, tamanho pré-declarado ou uma combinação desses tipos
Tipo de animação	Semelhante a Tipo de imagens gráficas
Tipo de som	Semelhante a Tipo de imagens gráficas
Nome de arquivo de página	Nome de arquivo de página

Tabela 13: Dimensão da Página
Fonte: KIMBALL & MERZ, 2000, pg 139

5.3.5 DIMENSÃO DO EVENTO

A dimensão do evento tem como objetivo descrever o que aconteceu em uma página particular em um momento específico de tempo. Os principais eventos interessantes em uma página são:

- Abrir Página;
- Atualizar Página;
- Clicar em Link;
- Inserir Dados.

Segundo KIMBALL & MERZ (2000), as páginas dinâmicas baseadas em XML estão se tornando mais comuns, a dimensão do evento torna-se muito mais interessante pois a semântica da página será muito mais “compreensível” para o servidores da Web. Cada campo em um documento de XML pode ser rotulado com um *tag* definido pelo usuário. O projeto para dimensão evento é ilustrado na TABELA 14.

ATRIBUTO	DESCRIÇÃO
Chave do evento	Valores substitutos, 1..N
Tipo de evento	Exemplo: Abrir Página, Atualizar Página, Clicar em Link, desconhecido, corrompido, inaplicável,
Conteúdo do evento	Campos dependente de aplicativo cujo conteúdo será extraído dos tags de XML

Tabela 14: Dimensão do Evento
Fonte: KIMBALL & MERZ, 2000, pg 139

5.3.6 DIMENSÃO DA SESSÃO

A dimensão da sessão fornece um ou mais níveis de diagnóstico para a sessão do usuário como um todo. O contexto local da sessão pode ser: “Solicitar Informações de Produto”; mas o contexto total da sessão pode ser: “Fazer o Pedido de um Produto”. O contexto local pode ser passível de decisão a partir da identidade da página atual, mas o contexto total da sessão pode ser julgado somente através do processamento da sessão completa do usuário em tempo de extração de dados.

O atributo *Status* do Cliente é um campo conveniente para “rotular” o cliente durante períodos de tempo, com “rótulos” que não são imediatamente claros tanto para a página quanto para a sessão imediata. *Status* úteis de clientes são definidos como cliente de “Alto Valor Confiável” ou “Novo Cliente”, ou “Prestes a Cancelar”, ou “em Padrão”. Se esses “rótulos” de clientes forem colocados dentro da seqüência de cliques, serão capazes de estudar diretamente o comportamento de determinados clientes. O motivo pelo qual tais “rótulos” de clientes não sejam colocados diretamente na dimensão cliente é pelo fato de que eles podem ser alterados em períodos muito curtos de tempo.

A dimensão da sessão é de grande importância pois fornece uma maneira de agrupar sessões para análises perspicazes – pode ser usada para solicitar:

- Quantidade de clientes que consultaram as informações de produtos antes de fazer o pedido;
- Quantidade de clientes que acessaram as informações referentes aos produtos mas nunca fizeram um pedido;
- Quantidade de clientes que iniciaram o processo de fazer o pedido mas não o terminaram.

O projeto recomendado para a dimensão da sessão é mostrado na TABELA 15:

ATRIBUTO	DESCRIÇÃO
Chave de sessão	Valores substitutos, 1..N
Tipo de sessão	Classificada, não classificada, corrompida, inaplicável
Contexto local	Principalmente o contexto derivado da página, como “Solicitar Informações de Produto”
Contexto total de sessão	Principalmente o contexto derivado da trajetória, como “Fazer o Pedido de um

	Produto”
Seqüência de ação	Rótulo de resumo para uma seqüência total de ações observadas durante a sessão
Status do sucesso	Se a missão total da sessão foi alcançada
Status do cliente	Rótulos como “Alto Valor Confiável” ou “No Padrão”

Tabela 15: Dimensão da Sessão
Fonte: KIMBALL & MERZ, 2000, pg 141

5.3.7 DIMENSÃO DE REFERÊNCIA

Tal dimensão descreve como o cliente chegou na página atual. Os *logs* de servidor da Web geralmente fornecem essas informações. O *Universal Resource Locator* (URL) que é o endereço de texto de um objeto específico na Web, é previamente acessada e identificada e, em alguns casos, informações adicionais são apresentadas. Se a origem do link (*referrer*) foi um sistema de pesquisa, então a cadeia de pesquisa é especificada. No campo Especificação é colocado algum tipo de especificação “limpa” e simplificada, pois, se for colocar a especificação bruta da pesquisa no banco de dados, os analistas não seriam capazes de consultá-las de forma útil pelo fato das especificações serem muito complicadas e idiossincráticas.

A TABELA 16 ilustra o projeto recomendado para a dimensão de referência.

ATRIBUTO	DESCRIÇÃO
Chave de referência/ Referral Key	Valores substitutos, 1..N
Tipo de referência/ Referring type	Por exemplo: <i>site</i> remoto, sistema de pesquisa, corrompido, inaplicável
URL de referência/ Referring URL	Por exemplo: www.organization.site.com/linkspace
Site de referência/ Referring site	Por exemplo: www.organization.site.com
Domínio de referência/ Referring domain	Por exemplo: <i>site.com</i>
Tipo de pesquisa	Por exemplo: correspondência simples de texto, correspondência lógica complexa
Especificação	A especificação real utilizada: útil se texto simples, questionável caso contrário
Alvo	Onde a pesquisa encontrou sua

	correspondência, por exemplo, em metatags, corpo de texto ou título
--	---

Tabela 16: Dimensão da Referência
Fonte: KIMBALL & MERZ, 2000, pg 142

5.3.8 DIMENSÃO DO PRODUTO (OU SERVIÇO)

A dimensão do produto descreve o produto ou serviço de uma página ou alvo de um evento. Essa dimensão será aplicável somente em certas situações – tal dimensão é utilizada somente por empresas que atuam no comércio eletrônico, pois se anexa totalmente à seqüência de cliques e assume valores significativos em muitos contextos.

De acordo com KIMBALL & MERZ (2000) as dimensões de produto são muito bem entendidas pela comunidade de projetos de DW. Elas variam desde a muito explícita unidade de manutenção de estoque, a *Stock Keeping Unit* (SKU), que lista centenas de milhares de produtos varejistas, até bem menores e menos elaboradas listas de serviços fornecidos, por exemplo, por uma empresa de consultoria. Uma boa dimensão do produto possui um grande número de atributos descrevendo cada produto. Pode ser criada uma hierarquia de mercadoria que permite a grupos de produtos serem reunidos em “blocos” menores. Para os autores há uma grande variabilidade de possíveis dimensões de produto/serviço. Dois projetos são mostrados para ilustrar a Dimensão do Produto do varejo eletrônico (TABELA 17) e a dimensão do serviço de uma empresa de consultoria de Data Warehouse (TABELA 18).

ATRIBUTO	DESCRIÇÃO
Chave do produto	Valores substitutos, 1..N
Tipo de produto	Por exemplo: regular, não medido, corrompido, inaplicável
Número de SKU	Código de barras
Descrição	Descrição possivelmente única de produto individual com versão e liberação
Marca	Nome da marca
Fabricante	Nome do fabricante
Categoria	Rollup do revendedor
Departamento	Rollup do revendedor

Tipo de sistema	Componente, software, hardware, sistema completo
Embalagem	Peça, sacola plástica, caixa de prateleira, caixa de chão, outra
Largura da embalagem	Largura da embalagem
Profundidade da embalagem	Profundidade da embalagem
Altura da embalagem	Altura da embalagem
Empilhamento máximo	Empilhamento máximo
Peso da embalagem	Peso da embalagem
Custo padrão	Nesta tabela somente será constante se for por um período longo; caso contrário, em outra tabela de fatos
Preço padrão	Nesta tabela somente será constante se for por um período longo; caso contrário em outra tabela de fatos
Comprador	Nessa tabela somente será constante se for por um período longo; caso contrário em uma dimensão de comprador

Tabela 17: Dimensão do Produto
 Fonte: KIMBALL & MERZ, 2000, pg 143

ATRIBUTO	DESCRIÇÃO
Chave do serviço	Valores substitutos, 1..N
Tipo de serviço	Por exemplo: regular, não medido, corrompido, inaplicável
Código de serviço	O código interno utilizado por sistemas de cobrança e monitoramento, se visível por gerentes
Descrição	Por exemplo: recolhimento de requisitos, treinamento de usuário final
Categoria	Por exemplo: desenvolvimento de Warehouse, consultoria de gerenciamento
Setor	Por exemplo: telecomunicações, operações bancárias, saúde, manufatura, governo

Tabela 18: Dimensão do Serviço
 Fonte: KIMBALL & MERZ, 2000, pg 144

5.3.9 DIMENSÃO CAUSAL

Essa dimensão é importante pois descreve as condições do mercado no momento em que a medida é feita na tabela de fatos. A dimensão causal “tenta” fornecer “dicas” ou fatores causais que possam explicar as razões pelas quais o cliente está interessado em uma determinada empresa ou em um produto ou serviço específico.

O projeto recomendado por KIMBALL & MERZ (2000) para a dimensão causal é mostrada na TABELA 19.

ATRIBUTO	DESCRIÇÃO
Chave causal	Valores substitutos, I..N
Tipo causal	Por exemplo: fator causal específico, nenhum fator causal, corrompido, inaplicável
Tratamento de preço	Por exemplo: regular, 2-por-1, redução de 10%, desconto de R\$2,00
Tipo de anúncio de jornal	Por exemplo: grande e diário, pequeno e diário, grande e de fim de semana, pequeno e de fim de semana
Tipo de anúncio de Web	Por exemplo: sistema de pesquisa constante, fração de tempo de portal de notícias
Tipo de anúncio de rádio	Por exemplo: 24 vezes diariamente em estações de notícias, 10 vezes nos fins de semana em estações de música clássica
Tipo de exibição na loja	Por exemplo: exibição de fim de corredor, etiqueta de prateleira
Tipo de promoção de fabricante	Por exemplo: cupom de desconto, produto extra de bonificação
Outro evento causal	Por exemplo: guerra, furacão, desconto do concorrente

Tabela 19: Dimensão Causal
Fonte: KIMBALL & MERZ, 2000, pg 144

5.3.10 DIMENSÃO DA ENTIDADE DE NEGÓCIO

Tal dimensão descreve uma entidade de negócio que é associada com um registro de fato. Uma entidade de negócio pode reproduzir vários papéis diferentes e até reproduzir mais de um papel simultaneamente. Uma entidade de negócio pode ser fornecedor, parceiro de negócio, uma origem de link (*referrer*), um provedor de serviço, um cliente, enfim.

Para KIMBALL & MERZ (2000), nesta dimensão é explorado o fato de que, independentemente do papel desempenhado, a entidade de negócio tem o mesmo tipo de descrição básica. O projeto recomendado pelos autores para a dimensão da entidade de negócio é ilustrada na TABELA 20.

ATRIBUTO	DESCRIÇÃO
Chave de entidade	Valores substitutos, I..N
Tipo de entidade	Por exemplo: comercial, governamental, individual, desconhecido, corrompido, inaplicável
Nome da entidade	Nome da empresa
Categoria de indústria	Por exemplo: códigos de Classificação Padrão da Indústria (<i>Standart Industry Classification</i>) com equivalentes em texto
Nome do contato primário	Nome do contato primário
Telefone primário	Telefone primário
FAX primário	FAX primário
Contato por correio eletrônico	Contato por correio eletrônico
Site da Web	Site da Web
Construção	Construção
Andar	Andar
Mail stop	Mail stop
Tipo de endereço	Por exemplo: comercial, governamental, individual, desconhecido, corrompido, inaplicável
Utilização pretendida	Utilização pretendida
Número da rua	Número da rua
Nome da rua	Nome da rua
Direção da rua	Direção da rua

Caixa de correio	Caixa de correio
Localidade	Localidade
Cidade	Cidade
Estado	Estado
Código postal	Código postal
Região	Região
País	País
Bloco montado de endereço	Um endereço postal completo e válido na seqüência correta com quebras de linhas
Moeda	Tipo de moeda utilizado por essa empresa nessa localização

Tabela 20: Dimensão da Entidade do Negócio

Fonte: KIMBALL & MERZ, 2000, pg 146

As chaves de monitoramento de seqüência de cliques são explicadas no próximo item antes de entrar diretamente na construção do Data Mart de Seqüência de Cliques a partir do modelo de projeto de cada Tabela Dimensional proposto por KIMBALL & MERZ (2000),

5.4 CHAVES DE MONITORAMENTO DE SEQÜÊNCIA DE CLIQUES

Com as dimensões definidas há a necessidade de se definir também os IDs de produção os quais surgem em várias situações (TABELA 21). O ID de host é um URL que identifica uma fonte ou um alvo. Os IDs de sessão podem ter relevância na construção da dimensão da sessão durante a fase de extrair-transformar-carregar, e os IDs de usuário serão cruciais durante a fase de extrair-transformar-carregar, mas não serão de interesse direto para uma dimensão final de visitante.

<i>Elemento de dados</i>	<i>Descrição</i>	<i>Persistência</i>	<i>Fonte</i>
Host	Endereço de IP ou nome de host completamente qualificado	Sessão	Endereço de IP do computador em que o navegador está sendo executado
ID temporário de sessão	Identificador único para uma sessão de http consistindo em um usuário	Sessão	Servidor de Web – possivelmente como um <i>cookie</i> de sessão

	a partir de um cliente para um host servidor		
ID de sessão	Identificador único atribuído para uma sessão de http consistindo em um usuário de um cliente para um host servidor	Ilimitado	Pós-processador de seqüência de cliques
<i>Cookie</i> de sessão	<i>Cookie</i> colocado em um espaço de armazenamento (da memória) temporário de um navegador	A partir do primeiro contato com um servidor da Web, até que o navegador do usuário seja desativado	Servidor da Web
<i>Cookie</i> persistente	<i>Cookie</i> colocado em um arquivo de <i>cookie</i> do navegador por um servidor da Web	Até a data/hora de expiração do <i>cookie</i>	Servidor da Web ou aplicativo baseado em Web
ID de usuário anônimo	Identificador único para um usuário cuja identidade real não tenha sido validada	Sessão	Servidor de aplicativo de sessão, como um <i>cookie</i> de sessão
ID de usuário persistente	Identificador único para um computador	Entre sessões	Servidor de aplicativo de sessão, via <i>cookies</i> persistentes, ou a partir de um serviço de <i>profiling</i>
ID de membro familiar válido	Identificador único para o membro familiar ou para o cenário de negócios em que um computador específico está localizado	Ilimitada	Link a partir do <i>cookie</i> persistente para o banco de dados de membros familiares

ID de usuário validado	Identificador único para registro no banco de dados do usuário	Ilimitado	Link a partir do aplicativo
------------------------	--	-----------	-----------------------------

Tabela 21: Chaves de Monitoramento
 Fonte: KIMBALL & MERZ, 2000, pg 148

5.5 DATA MART DE SEQUÊNCIA DE CLIQUES

Com as tabelas dimensionais definidas fica possibilitado para os analistas modelarem de acordo com a necessidade do negócio da empresa. O Data Mart de Sequência de Cliques permitira uma análise tanto de sessões de usuário completada, como de utilização de uma página individual.

Este item ilustra através da **FIGURA 26**, o registro mais detalhado de comportamento de clientes no comércio eletrônico através de uma tabela de fatos de sequência de cliques em que o grão (registro único em uma tabela de fatos) é cada evento da página.

A tabela de Fatos de Eventos de Páginas de Sequência de Cliques, é aqui abordada pois a mesma tem como objetivo registrar cada gesto realizado por cada cliente no comércio eletrônico – o evento individual de página, foco deste trabalho.

Antes que qualquer versão de Data Mart de Sequência de Cliques seja implementada, KIMBALL & MERZ (2000), sugerem que os seguintes itens sejam identificados na ordem descrita a seguir:

- As questões de negócio a serem endereçadas por este Data Mart;
- A fonte dos dados de medição numérica (fato);
- O grão de cada tabela de fatos proposta;
- Todas as dimensões necessárias por todas as tabelas de fatos propostas;
- Um plano e um compromisso para adaptar tais dimensões por toda a empresa;
- Todos os fatos numéricos a serem incluídos em cada tabela de fatos;
- Um plano e um compromisso para adaptar quaisquer fatos que apareçam em mais de uma tabela de fatos por toda a empresa.

Com a construção dos primeiros modelos dimensionais, a equipe de Data Mart será capaz de expandi-lo para que o mesmo inclua mais fontes de dados, e para conectá-lo a outros Data Marts abrangidos pelo DW da empresa.

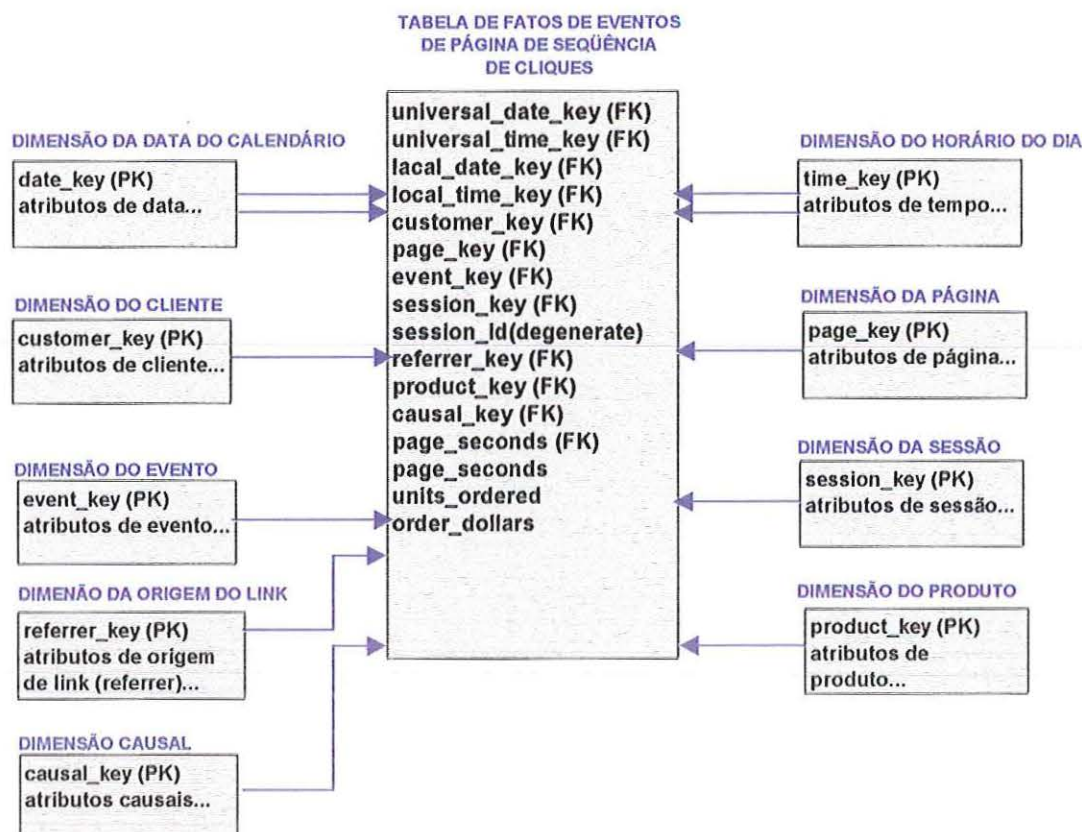


Figura 26: Uma tabela de fatos de seqüência de cliques
Fonte: KIMBALL&MERZ, 2000, pg 154

Interpretando o modelo dimensional mostrado na FIGURA 26, as dimensões necessárias são: data do calendário, cliente, evento, origem do link (*referrer*), causal, horário do dia, página, sessão e produto.

Deve-se dividir as tabelas dimensionais de data do calendário e horário do dia porque ambas possuem elementos descritivos diferentes. A data relaciona-se ao calendário, dia da semana, estações. Logo, o tempo do dia ou horário do dia se relaciona com o traço que é deixado marcado naquele momento. A dimensão data tende a se unir com atributos textuais. A dimensão Horário do Dia pode ser bastante complexa ao menos que se tenha alguns intervalos específicos durante o dia aos quais se possa classificar. Um dos “papéis” representados pelas dimensões data de calendário e horário do dia é sincronizar os horários de diferentes países e expressar o tempo local de “relógio de parede” medido pelo cliente. De acordo com KIMBALL & MERZ (2000), para se alcançar a sincronização das datas e

dos horários internacionais, deve-se registrar todas as datas e horários de sessões uniformemente em um único fuso horário, como o do Meridiano de Greenwich (GMT). A interpretação de início de sessão é feita através das combinações de datas e horários de sessão, sendo que o tempo de permanência da sessão consiste em um fato numérico.

A dimensão da Página refere-se à página individual cujos eventos é registrado. Através dessa dimensão é capaz de se verificar todas as páginas acessadas pelos clientes. A dimensão Página é importante por conter informações significantes para a análise da página no Servidor da Web. Cada página da rede possui alguma descrição simples que identifique o tipo de página e de sua localização. Um local na rede possui uma descrição hierárquica associada a cada página. Esta informação deve ser armazenada na dimensão Página e mantida em constante atualização, pois pode haver modificações nas estruturas da “Árvore” na rede. Em outras palavras, deve-se atualizar o Banco de Dados do Data Mart para que as necessidades de análise do Data Webhouse sejam atendidas.

A dimensão do Evento descreve o que de fato aconteceu na página. Cada evento da página registra o número de segundos que acredita-se ter passado antes do próximo evento da página.

A dimensão da Sessão descreve o resultado da sessão. Pretende descrever classes e categorias de sessões, não as características de cada sessão individual. Serve para mostrar todos os eventos de uma página que constituem uma sessão. Nesta dimensão classifica-se uma sessão e localiza-se suas atividades. As sessões poderiam ser caracterizadas como: “Procurando Informações”, “Olhando por acaso”, “Preço e Características”, “Fazendo Pedido”. Elas devem ser caracterizadas conforme aquilo que o usuário possa vir a fazer dentro do *site*. Deve-se montar pacotes de eventos que formam uma sessão.

Já a dimensão do Produto ou Serviço é acompanhada com a dimensão causal, para que se possa anexar interpretações úteis de mercado às alterações na demanda para determinados produtos.

A dimensão do Cliente deve conter informações úteis sobre quem é ele, e não apenas constar o IP. Porém isso dependerá da maneira com que interagirmos com o mesmo, solicitando ou não suas informações.

Na tabela de Fatos ilustrada na **FIGURA 26**, as combinações de data e horário referem-se ao evento individual de página que está sendo descrito pelo registro que se está construindo. Os fatos finais são reais unidades encomendadas do pedido. Esses campos serão iguais a zero ou nulos para muitos dos registros nessa tabela de fatos, pois o evento específico de página não é o evento que efetua o pedido. Portanto é conveniente fornecer esses campos porque eles “amarram” toda a renda importante proveniente da Web

diretamente com o comportamento do cliente. Segundo KIMBALL & MERZ (2000), se as unidades encomendadas e os reais do pedido estavam somente disponíveis pelo sistema de entrada do pedido de produção em outra parte do Data Webhouse, seria ineficiente realizar a análise de renda com relação ao comportamento através de inúmeras tabelas extensas.

O grão da tabela de Fatos é todo o evento significativo em uma sessão individual para cada usuário. O Data Mart dimensional apresentado na TABELA 26 por KIMBALL & MERZ (2000) permite responder questões bastante poderosas, sendo elas:

- Como direcionar as atividades de *marketing* pelo reconhecimento da identidade do cliente agrupando o cliente por atributos de seqüência de cliques?
- Quando os clientes utilizam um determinado *site* da Web, qual é o horário e local ?
- Quando ocorre a carga simultânea máxima em um determinado *site* da Web de usuários em diversos fusos horários diferentes ?
- Quanto tempo os clientes gastam em um determinado *site* da Web e quantas páginas eles visitam ?
- Como os clientes chegam a um determinado *site* da Web?
- Qual o volume total de vendas de produtos feitas pela Web ?
- Como saber se um cliente está prestes a deixar um determinado *site* da Web ?
- Como saber se um anúncio particular da Web está funcionando ?
- Como saber se as saudações personalizadas estão funcionando?
- Como saber (parcialmente) se uma promoção é lucrativa ?

Não há muito esforço para mapear as partes mais visitadas do *site* e identificar os usuários mais freqüentes. A análise do histórico de cliques e a trajetória dos usuários no *site* depende em como se constrói a “Dimensão Sessão”. Se possuir boas descrições para cada sessão, pode-se executar as análises secundárias necessárias ao negócio.

Observa-se no entanto que o objetivo de se construir um modelo dimensional para analisar a seqüência de cliques que toda empresa deve seguir é o de fazer desses dados, informações úteis para conhecer melhor cada consumidor que transita por um determinado *site* e através disto personalizar produtos e serviços de acordo com o perfil de seus melhores clientes. Atingindo esse objetivo, a empresa terá grandes chances de garantir a fidelidade de seus consumidores, que é um dos pontos que elas mais almejam.

Segundo KIMBALL & MERZ (2000), se o CRM estivesse preocupado apenas em fornecer o *status* do pedido mais recente, não haveria a necessidade do Data Webhouse. Porém o CRM busca uma perspectiva mais ampla no relacionamento com o cliente. Várias decisões tomadas em tempo real pelo cliente durante os processos de negócio precisam de

algo do Webhouse, quer seja a exibição do ritmo de crescimento das compras, quer sejam dicas sobre o que os outros clientes semelhantes necessitam.

A importância de se integrar TI ao *marketing* no comércio eletrônico, vêm sendo discutida por diferentes autores. Todos concordando que a aplicação de um DW no comércio eletrônico trará benefícios às empresas quando se questiona a fidelidade de clientes através de programas de CRM.

De acordo com PEPPERS & ROGERS (2000), o CRM analítico, fonte de toda a inteligência do processo, serve para o ajuste das estratégias de diferenciação de clientes, bem como para o acompanhamento de seus hábitos, com o objetivo de identificar suas necessidades e os eventos que possam ocorrer na vida de cada um deles. Tudo com o objetivo de tomar, para os melhores clientes, a relação com a empresa algo fácil e conveniente, buscando sua satisfação e fidelidade. Além disso, é através das ferramentas analíticas que identifica-se os clientes de menor valor para a empresa. Idealmente, as ferramentas de análise, mineração de dados (Data Mining) e relatórios analíticos, devem ter como única fonte de dados o DW, base de dados histórica e integrada que tem como objetivo fornecer uma visão única, orientada a assuntos, dos dados da empresa.

Para TAURION (2000), o estágio analítico envolve o tratamento das informações sobre clientes para gerar um maior conhecimento sobre seus hábitos e interesses. Com base neste conhecimento, a empresa pode refletir suas ações de *marketing*, refinando a segmentação de seus clientes. Os componentes tecnológicos deste nível envolvem um DW e as ferramentas de análise, como OLAP e Data Mining.

BRETZKE (2000), ainda afirma que a integração entre *marketing* e tecnologia deve proporcionar às empresas meios de se relacionar com o cliente em tempo real.

Portanto aplicando-se o CRM analítico ao DW em um ambiente Web, as empresas “acompanham” seus clientes em toda sua trajetória, pois os dados são históricos e armazenados durante anos. Ocorre assim, um melhor entendimento de suas expectativas e necessidades para melhor servi-los.

Para FLETCHER (1990), a tecnologia está mudando a natureza do *marketing* pelo impacto causado no mercado e no sistema de *marketing*. Novos produtos, novas mídias, novos serviços, *feedback* das informações sobre o comportamento do consumidor, integração com a cadeia de suprimentos e muitos outros desenvolvimentos, indicam uma necessidade crescente dos executivos de *marketing* conscientizarem-se de como isso afetará a sua atuação sobre o mercado e o próprio mercado. Através da habilidade e capacidade de gerenciar informações através de Sistemas de Informação e *Softwares*, aplicativos especializados, aumentam a eficiência com a qual as atividades de *marketing* podem ser desenvolvidas.

“O impacto da tecnologia sobre a filosofia do *marketing* é tão marcante que a resposta mais efetiva a nível tático para que uma empresa consiga orientar suas atividades para o mercado, encontra-se numa disciplina que integra à prática do marketing as inovações tecnológicas(...) portanto, o ambiente altamente mutante e competitivo, aliado às inovações tecnológicas, está propiciando o momento adequado, para que além do reconhecimento da necessidade de mudança na filosofia da empresa, efetivamente os executivos busquem os meios e as ferramentas para promover as mudanças, que permitam atuar sob um novo conceito de *marketing* que resulta na orientação para o mercado” (BRETZKE, 1992, p15).

Não há dúvidas que um DW bem estruturado trará benefícios para a área de *marketing* no comércio eletrônico, mas ainda precisa atingir um certo grau de maturidade no que se refere a identificação de clientes pois, algumas questões, muitas delas referentes ao fator humano estão sendo desprezadas. Como saber o perfil de um cliente, ou seja, suas necessidades, expectativas em relação a um determinado *site* se esse mesmo cliente acessa o mesmo *site* por diferentes computadores ? Como saber se é a mesma pessoa que está “por trás” do computador, uma vez que possivelmente o uso de um mesmo computador não seja exclusivo ? Se a estratégia de negócio do CRM é feita apenas para os melhores clientes (CMV e CMP), isso não significa discriminar os outros clientes restantes? Qual a veracidade dos dados cadastrais de clientes na Tabela de Clientes ? Os *cookies* fornecidos pelos servidores estão identificando um computador ou um indivíduo ? Mesmo criando ID de membro familiar um determinado cliente pode abandonar sua “navegação” no *site* e, sem fechar a página da Web outra pessoa pode iniciar outra navegação no mesmo *site*, será que esse ID de membro familiar exerce realmente sua função ?

Portanto ao tratar a identificação do usuário a equipe desenvolvedora do Data Webhouse certamente passará por esses desafios: anonimato de usuários, identidades falsas, membros familiares utilizando o mesmo computador, o mesmo usuário utilizando diversos computadores, enfim. A análise comportamental dos usuários pode ser feita também utilizando *login* com dois identificadores logo no início da “navegação” do *site* de comércio eletrônico, além dos *cookies* já utilizados. Essa forma de identificação poderia ser uma tentativa de solução no que se refere a identificação do cliente, mesmo que o nome do *login* não tenha relação real com o nome do usuário. Passaria então a ser monitorado as ações de um usuário identificado pelo seu *login* e seu *cookie* e não as ações de um computador.

De um modo geral a proposta de se construir um Data Webhouse apresentada por KIMBALL & MERZ (2000) apresenta uma promissora evolução no comércio eletrônico, no que se diz respeito ao gerenciamento do relacionamento com o cliente. Dados de seqüência de cliques são armazenados historicamente em tabelas dimensionais apresentadas neste

capítulo. Através da análise desses dados pode-se agrupar níveis de clientes e aplicar a estratégia do CRM para garantir a fidelidade destes. Se a equipe desenvolvedora atentar-se mais para questões referentes ao comportamento humano e não da máquina, o Data Webhouse atingirá sua maturidade e provavelmente será uma ferramenta favorável para o mercado eletrônico focado no cliente. Desta forma, a integração de TI com *marketing* cumprirá o seu papel genuíno. As funções da equipe desenvolvedora de um Data Webhouse são abordadas no próximo item.

5.6 GERENCIANDO O DATA WEBHOUSE

“Um bom gerente de Webhouse equilibra todas as tarefas individuais a serem realizadas na construção e suporte de um completo sistema de Webhouse. Um bom gerente de projeto também compreende as pessoas e os papéis envolvidos em um projeto. Acima de tudo, um bom gerente de projeto deve ter perspectiva, discernimento e flexibilidade, pois nenhum projeto jamais seguirá completamente o plano original” (KIMBALL & MERZ, 2000, pg 300).

De acordo com as afirmativas dos autores anteriormente citados, este item tem como objetivo evidenciar que para se ter êxito no gerenciamento de um Data Webhouse, o gerente de projeto deve além de executar uma grande quantidade de tarefas diferentes, gerenciar pessoas com diferentes habilidades e diferentes perspectivas, se comunicar com a equipe, com seus parceiros de negócio e com a gerência sênior sobre as necessidades e perspectivas do projeto. Um dos grandes riscos encontrados em um projeto de Data Webhouse é a falta de integração entre o Webmaster e sua equipe com as pessoas responsáveis pelo Data Webhouse. Os interesses do CRM pode estar em risco se não houver a integração entre esses dois grupos.

A precisão de um projeto de Data Webhouse é difícil de ser executada conforme seu planejamento devido as inúmeras variáveis que interferem o projeto desde seu início até o término. Tais variáveis podem ser:

- Os dados se encontram mais difíceis de serem extraídos do que foi inicialmente planejado;
- Os dados se encontram mais complexos que tornam aos programas de extração uma difícil depuração dos mesmos;
- Usuários finais podem requisitar novos dados adicionais que não estão no plano original;
- Fornecedores podem atrasar a entrega de hardware, software, recursos prometidos ou correção de defeitos;

- A equipe de TI pode distrair-se com um projeto de emergência etc.

Desta forma o líder de projeto de Data Webhouse deve ser capaz de controlar mudanças bruscas de escopo e ter flexibilidade para reconhecer que estas “variáveis” são de extrema importância no projeto, sem poder descartá-las. A comunicação e o envolvimento de toda equipe é vital no planejamento e no gerenciamento da implementação de um projeto de Data Webhouse.

KIMBALL & MERZ (2000) recomendam que sejam examinados cinco indicadores-chave para a agilidade de um projeto de Data Webhouse:

1. Forte envolvimento da gerência;
2. Uma grande motivação para negócios, fazendo com que os usuários finais vejam os *insights* que os dados provavelmente conterão;
3. Boa parceria entre TI e o negócio;
4. Possuir uma cultura analítica de suporte, onde exista uma tradição de gerenciamento baseado em números e confiança em dados para revelar tendências e desenvolvimentos importantes no negócio;
5. A existência de dados reais em um sistema operacional real, contendo conteúdo de qualidade suficiente para servir como base para o Webhouse.

Confirmados positivamente os cinco indicadores-chave do Webhouse, o próximo passo é identificar um patrocinador comprometido e um patrocinador de TI comprometido. O patrocinador comprometido é o indivíduo mais importante de todo o projeto, pois fornecerá a justificativa, a motivação e finalmente os usuário finais para o projeto. Já o patrocinador de TI comprometido, também desempenha um papel crucial em vários pontos, ou seja estará presente quando as expectativas em relação a entrega do Webhouse precisarem ser estabelecidas e quando as discussões sobre as dimensões adaptadas e os fatos adaptados tornarem-se realmente necessárias.

Como qualquer projeto, o Data Webhouse requer o envolvimento de vários papéis separados. Alguns papéis são permanentes e centrais para o negócio e outros somente são necessários durante um determinado período de tempo. Como o projeto de Data Webhouse é de grande porte, há um envolvimento de toda da empresa – os papéis-chave de toda a gerência serão preenchidos por pessoas diferentes e, algumas partes do desenvolvimento e do suporte serão realizadas pelo grupo todo.

KIMBALL & MERZ (2000) sugerem uma equipe de projeto (ilustrado na FIGURA 27), incluindo Webmasters, gerentes de projeto de Webhouse, modeladores de dados de

seqüência de cliques, programadores de extração, administradores de banco de dados, analistas de aplicativos, patrocinadores comerciais e, naturalmente, usuários finais. No caso dos usuários finais, após serem identificadas as funções para cada membro da equipe desenvolvedora do Data Webhouse, serão feitas entrevistas com os mesmos para então selecionar os dados corretos para atender corretamente às necessidades do negócio e das decisões a serem tomadas em seguida. Os papéis (ou funções) identificados, são descritos logo a seguir.

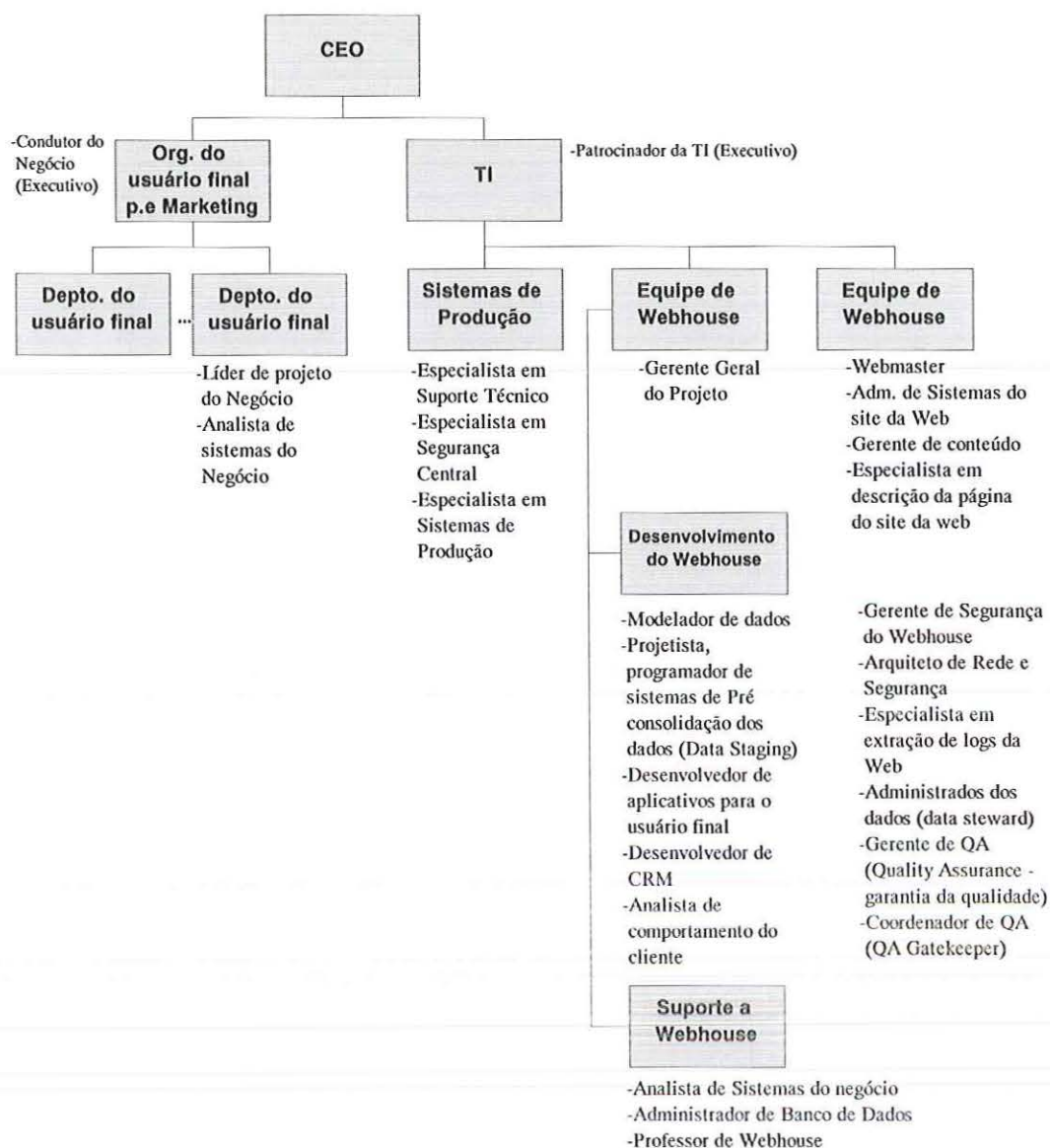


Figura 27: Papéis e Organizações de Webhouse

Fonte: KIMBALL & MERZ, 2000, pg 304

Linha de Frente: patrocinadores e condutores

- *Patrocinador do negócio:* é o papel mais importante de um projeto de Data Webhouse. O patrocinador do negócio é o cliente final. Um bom patrocinador do negócio fornece a demanda e os recursos para o projeto. Deve estar comprometido

em realizar negócios na Web, seja observando o comportamento do cliente pela seqüência de cliques ou se ele depende da Web para conseguir acesso ao DW. Em ambos os casos esse executivo deve ver a Web como a principal maneira de fazer negócios.

- *Patrocinador de TI:* esse executivo disponibiliza os recursos para finalizar o projeto de Webhouse. O patrocinador de TI precisa compreender o “ritmo” do Webhouse e o que torna o projeto mais fácil ou mais difícil. Ele é responsável por “configurar” as expectativas de diferentes gerências sênior na organização para a entrega do projeto de Webhouse. O patrocinador de TI também necessita entender que a extração de dados é difícil e sempre leva mais tempo do que o estimado.
- *Condutor do negócio:* quando o principal executivo patrocinador do negócio não estiver permanentemente disponível, várias organizações designarão um gerente de nível médio para servir como um “elo” com o patrocinador do negócio. O objetivo é ter ao lado do cliente, um gerente dedicado e disponível e com autoridade para tomar decisões e fornecer motivação.

Treinadores: gerentes de projeto e líderes

- *Gerente-geral de projeto:* o principal papel no projeto de Webhouse é o gerente do projeto. Essa é uma posição em tempo integral na TI. O gerente do projeto deve possuir diferentes habilidades – precisa ser bom em lidar com as pessoas, bom conhecimento em tecnologia, bom em gerenciamento de projetos e adquirir recursos em diferentes ocasiões e organizações. Não necessariamente, o gerente-geral de projeto precisa estar incumbido da maior parte da estrutura do gerenciamento de projetos, alguns gerentes de projeto, principalmente em projetos menores, sentem-se mais à vontade apenas supervisionando o trabalho. Nesses projetos menores, eles utilizam a intuição para conhecer as necessidades do projeto e gastam o seu tempo comunicando-as. Já em grandes projetos, os gerentes utilizam ferramentas de monitoração para documentar dependências e *status*. Acima de tudo o gerente-geral de projeto precisa manter uma expectativa de todo o projeto. Se o projeto atrasa, o gerente do projeto deve reconhecer o fato e comunicar tanto ao gerente de TI como ao patrocinador do negócio.

- *Líder de projeto do negócio:* a contraparte do gerente geral é o líder de projeto do negócio, o qual deve estar disponível diariamente para coordenar os trabalhos com o gerente do projeto. No início do projeto, o líder de projeto do negócio, será o instrumento para organizar as entrevistas com o usuário final. Terá como obrigação interpretar as perguntas sobre o conteúdo dos dados, bem como recursos para ferramentas de consultas e de informação. Será de seu encargo também, guiar o desenvolvimento das ferramentas do usuário final. Quando o projeto estiver preparado para ser lançado, o líder de projeto do negócio se certificará de que os usuários finais estão dispostos para o treinamento. Não é necessário que esse profissional seja o responsável por tempo integral, mas ele precisa estar atento às necessidades do projeto.

A equipe básica do projeto

- *Analista do negócio:* é o profissional técnico que entende em detalhes os requisitos do negócio. Um analista do negócio bem sucedido pode ser uma pessoa de TI com uma grande credibilidade em ambientes do negócio para o usuário final ou um analista/usuário final com boas habilidades técnicas. O analista do negócio conduzirá no início do projeto as definições dos requisitos do negócio e, posteriormente, será responsável pela arquitetura de consultas e relatórios padrão criados para a comunidade dos usuários finais. O analista do negócio é o usuário final avançado, pois sua maior responsabilidade deve ser o sucesso total do Webhouse, não somente uma análise particular do negócio.
- *Modelador de dados:* este tipo de profissional precisa entender os dois tipos de modelagem: E-R e a dimensional. Os modelos de E-R freqüentemente vêm do ambiente de processamento de transação, podendo também ser utilizados para limpeza de dados nos bastidores do Webhouse. Modelos dimensionais são necessários para a efetiva entrega de dados aos usuários finais. Os modeladores de dados necessitam ter um bom conhecimento da diferença entre os relacionamentos da modelagem de dados reais e da modelagem de dados abstratos. Na análise final, os dados reais são aqueles entregues ao usuário final e um grande esforço deve ser dedicado nas preocupações práticas com o que realmente existe. Os modeladores de dados sêniores podem ser intitulados como arquitetos de dados.

- *Administrador de banco de dados (DBA)*: é o profissional que é responsável pela implementação do banco de dados. É responsável também pelo desempenho, utilização de espaço, *backup* e recuperação de informações. Um DBA sênior pode ser responsável por determinar as plataformas de *software* nas quais os bancos de dados OLTP e OLAP são executados. Depois da implementação, esse profissional freqüentemente mantém um papel operacional, observando e zelando pelas operações rotineiras do banco de dados corporativo para assegurar que o desempenho e objetivos estão sendo cumpridos. Um DBA deve ser extremamente sensível às questões de capacidade e desempenho, deve ser capaz também de montar sistemas que cumpram os objetivos iniciais e ter também a capacidade de dimensionar tanto o tamanho como a carga.
- *Administrador de Sistemas do site da Web (AS)*: o AS administra todos os aspectos operacionais da hospedagem na Web. É responsável pela seleção do hardware, do software e pela arquitetura total da hospedagem. Esse tipo de profissional administra a complexidade da conexão do *site* da Web com a Internet. O AS pode ser responsável pela segurança do *site*, pela manutenção de *firewalls* e pelo monitoramento de atividades do sistema em busca de evidências de invasão. É essencial que o gerente de projeto do Webhouse empenhe o AS no esboço do projeto pois, muitas das questões sobre capacidade de hardware e segurança de software podem deter tempo para serem resolvidas com o grupo de administração de sistema.
- *Projetista de sistema de pré-consolidação de dados*: esse profissional é responsável pelos sistemas ETL (extrair, transformar e carregar) que alimentam o Webhouse. Tal profissional, deve ter uma posição técnica mais antiga entre o pessoal do departamento de TI e estará encarregado do projeto desde sua concepção inicial até sua instalação. O projetista de sistema deve possuir uma ampla experiência prática com várias tecnologias, incluindo servidores da Web, banco de dados, arquitetura cliente/servidor e UNIX¹⁶. Também é de responsabilidade desse membro da equipe escrever especificações para vários componentes de ETL e deverá atender as variações técnicas do desafio do pós-processador de seqüência de cliques.

¹⁶ Nota da autora: UNIX consiste em um sistema operacional multitarefa. É o sistema operacional para servidores de Internet mais comumente utilizado.

- *Desenvolvedor de aplicativos de usuário final:* é o profissional que possui interesse especial tanto por questões do negócio como por questões técnicas. O desenvolvedor de aplicativos precisa ocupar-se com os usuários finais, compreendendo o que estão tentando fazer, bem como a identificação das dificuldades dos usuários quando estes utilizam o sistema. Devem também, dedicar algum tempo a atividades de suporte, incluindo ajudar os usuários em suas consultas e relatórios e conduzir treinamento. Apenas com esse estreito envolvimento com as questões do usuário final, o desenvolvedor de aplicativos terá condições necessárias para conduzir uma “biblioteca” de relatórios parametrizados.
- *Desenvolvedor de CRM:* é o membro da equipe que implementa o próprio sistema de CRM – é o maior responsável tanto pelos aspectos de produção de sistema de CRM como pelos aspectos do Data Warehouse. Um desenvolvedor de CRM está preocupado em entregar em tempo real a mais eficiente interface com o usuário ao cliente que acessa o *site* da Web. Esse profissional utiliza os recursos do DW para entregar uma experiência personalizada pois, o *site* da Web permite ao cliente realizar consultas que irão para o próprio DW (esse é um exemplo de trazer o DW para a Web). O desenvolvedor de CRM está muito interessado nos resultados da análise da seqüência de cliques, porque isso afeta as decisões do projeto no sistema de CRM. Ele trabalha bem próximo dos desenvolvedores de página da Web.
- *Analista de comportamento de cliente:* esse profissional é outro membro da equipe de CRM. É o profissional que está diretamente responsável por compreender a seqüência de cliques e o relacionamento da seqüência de cliques com as transações capturadas por outros sistemas, como o sistema de entrada de pedidos. Trabalha juntamente com os projetistas de *site* da Web propondo rótulos para páginas da Web e fazendo com que a análise da seqüência de cliques seja entendida pelos projetistas. O analista de comportamento de cliente trabalha com especialistas em extração de *logs* da Web, descritos rapidamente, para diagnosticar sessões de usuário, de modo que a dimensão da sessão seja mais útil e expressiva.
- *Professor de Webhouse:* este profissional também trabalha no desenvolvimento de aplicativos para o usuário final. Ele necessita entender e ensinar o conteúdo de dados, bem como, a utilização da ferramenta. Mais do que o desenvolvedor de aplicativos, esse papel, embora familiar na era do DW pré-Web, foi profundamente afetado pela

Web. Quase todos os detalhes sobre o treinamento em interface com o usuário deverão ser revistos sob o ângulo da interface com o usuário da Web.

- *Webmaster:* é o membro responsável pela aparência e comportamento globais do *site* da Web, para assegurar que ferramentas comuns como recursos de pesquisa estejam disponíveis dentro da arquitetura da hospedagem, e para que toda a atividade da empresa na Internet seja por links comuns e um *front-end* unificado para o usuário da Web. Frequentemente, o webmaster irá selecionar o desenvolvimento da Web e ferramentas de manutenção a serem utilizadas para o desenvolvimento do *site* da Web por toda a empresa. Uma das responsabilidades do webmaster é fornecer estatísticas completas do *site* em termos do sucesso e tempo de resposta, mas esse profissional não estará preocupado em analisar em detalhes qual será o serviço pretendido ao Webhouse.
- *Gerente de conteúdo:* esse profissional é a chave para o projeto de Webhouse e pode, de fato, ser um consumidor primário da produção do Webhouse. O gerente de conteúdo normalmente será um gerente de departamento ou de seção com a responsabilidade de expor a mensagem da organização no *site* da Web corporativo. Este membro frequentemente estará dentro do departamento de *marketing* e poderá ser intitulado de “editor de gerenciamento”. Um departamento de conteúdo possui um ou mais desenvolvedores de página da Web, criando em HTML¹⁷ o conteúdo real da página.
- *Gerente de segurança do Webhouse:* a principal responsabilidade desse profissional é definir e administrar as definições de papéis do usuário. Um ambiente de Webhouse de grande porte pode possuir dezenas ou centenas de definições separadas de papéis. Clientes ou parceiros de negócio confiáveis que possuem acesso ao DW podem, cada um, possuir seu próprio papel na segurança. Metade da responsabilidade dos gerentes de segurança está na definição dos papéis, a outra metade está na aplicação apropriada desses papéis para todas as informações do Webhouse. Além das definições de papéis, o gerente de segurança do Webhouse também precisa monitorá-lo contra invasões e “brechas” de segurança. Para essa capacidade, o gerente de segurança de Webhouse pode trabalhar com a equipe de segurança central de TI.

¹⁷ Nota da autora: HTML ou *Hypertext Markup Language* é uma linguagem de marcação padrão para definir as características de apresentação de documentos da Web.

- *Especialista em descrição da página de site da Web*: este membro da equipe auxilia o Webhouse definindo os atributos para páginas da Web, o que as torna compreensíveis em um contexto de banco de dados. Cada página precisa de um ou mais atributos descritos que pode servir como base para as restrições ou agrupamentos. Exemplificando tem-se: uma página para navegação, outra para preenchimento de pedidos ou uma página para informações de produto. O especialista de descrição de página também necessita trabalhar com a equipe de desenvolvimento de CRM para que páginas dinâmicas geradas em servidores de aplicativo possuem atributos úteis.

Equipes Especiais

- *Arquiteto de rede e segurança*: a função deste profissional é estratégica. É responsável por toda a arquitetura de segurança do Webhouse. Suas decisões incluem escolha de tecnologias para a implementação da segurança, ou seja, local no sistema onde se aplicará a segurança, dispositivos de isolamento da segurança e o projeto do sistema de recuperação após desastre. O gerente de segurança também pode ser o arquiteto de segurança, se houver tempo extra. Não é de responsabilidade do arquiteto de rede e segurança tomar as decisões táticas sobre quais usuários verão quais dados.
- *Especialista de suporte técnico*: são os profissionais da área de TI. São responsáveis por várias partes da infra-estrutura que o Webhouse depende. A infra-estrutura inclui desde computadores e seus inventários até a planta física (incluindo a rede).
- *Programador de pré-consolidação de dados*: a pré-consolidação de dados é uma das fases mais longas e imprevisíveis de qualquer projeto de DW. Um bom programador de pré-consolidação de dados dedica igual tempo ao desenvolvimento lógico primário, testes de unidade, testes de sistema e documentação. Esses profissionais podem programar diretamente em uma linguagem de programação convencional ou podem utilizar uma ferramenta orientada e graficamente integrada, permitindo que a lógica de ETL seja montada e reorganizada na tela.
- *Especialista em extração de logs da Web*: esse profissional é um tipo especial de programador da pré-consolidação de dados, exclusivo ao ambiente de Webhouse. O especialista em extração de logs da Web trabalha com o webmaster para fornecer logs

da Web que são mais excessivos e significativos. Em alguns casos os servidores da Web podem ser configurados para fornecer informações opcionais e mais úteis. Esse profissional irá trabalhar com a equipe de CRM e com as equipes de sistemas operacionais para definir e disponibilizar *tags*¹⁸ de XML¹⁹ que forneça controle e monitoração de comportamento no nível subpágina.

- *Administrador de dados*: é o profissional que é responsável pela definição no Webhouse das dimensões adaptadas e dos fatos adaptados, possuindo uma responsabilidade central na definição dos metadados orientados para negócios, tornando esses metadados úteis para os desenvolvedores de aplicativos e para usuários finais.
- *Especialista de suporte do sistema de produção*: são profissionais que normalmente são atribuído os sistemas de produção de legado da empresa. Eles são “recursos” inestimáveis para as equipes de Webhouse, quando as auditorias dos dados originais estiverem sendo conduzidas nas primeiras etapas de planejamento e quando os sistemas ETL do Webhouse estiverem sendo construídos.
- *Gerente de garantia de qualidade*: este membro da equipe possui um papel estratégico na definição de uma qualidade aceitável de dados para o Webhouse. É de responsabilidade deste profissional definir procedimentos utilizados pelo coordenador de garantia de qualidade a cada “manhã” que os dados são carregados no Webhouse. O gerente de garantia de qualidade alerta a gerência sênior sobre questões significativas relacionadas com a qualidade dos dados, pois a maioria dos problemas de qualidade dos dados são considerados “sintomas patológicos” na organização e não somente problemas técnicos.
- *Coordenador de garantia de qualidade*: esse profissional realiza o julgamento humano final de quando os dados estarão aptos para a publicação. A decisão final para a publicação é humana. Ele utiliza um número de ferramentas de auditoria, incluindo relatórios especiais de usuário final, para determinar se os dados estão completos e sensatos. O coordenador de garantia de qualidade comunica-se

¹⁸ Nota da autora: tag é um HTML. Consiste em um par simétrico para delimitar um determinado formulário. Por exemplo: <nomedotag></nomedotag>

diretamente com a comunidade de usuários provavelmente via correio eletrônico, com o *status* de cada liberação/lançamento de dados do Webhouse.

Os papéis apresentados por KIMBALL & MERZ (2000) evidenciam que a implantação de um Data Webhouse é uma tarefa complexa envolvendo diversas tecnologias e profissionais diferentes de uma organização. Desde o início do projeto, até sua implantação, vários papéis e responsabilidades devem ser desempenhados. O envolvimento dos usuários é essencial. Sem a integração e delegação de responsabilidade desde o início, o projeto certamente passará por dificuldades.

Um fator importante é que como o DW, o Data Webhouse está em constante estado de desenvolvimento, e mesmo após sua disponibilização para uso, necessita da retaguarda dos profissionais que o desenvolveram. O número de membros envolvidos na implementação de um Data Webhouse é muito variável pois, dependerá do tamanho do projeto a ser desenvolvido. As principais funções dos profissionais da equipe de desenvolvimento foram citadas anteriormente, alguns denotam imediatamente um cargo ou função, outros não, e alguns podem ser desempenhados pelos mesmos profissionais. Uma consideração importante a ser feita é que, para que a implementação de um Data Webhouse seja bem sucedida, deve haver uma total integração dos profissionais da área de negócios e de *marketing*.

No entanto, pode-se numerar algumas das principais variáveis que afetam as organizações no que se refere a implantação de um Data Webhouse, são elas:

1. Custo elevado;
2. Complexidade do projeto;
3. Ausência de um patrocinador;
4. Ausência de comprometimento explícito da alta administração;
5. Afastamento de algum membro da equipe por algum motivo;
6. O projeto deve ser guiado pelo negócio e não pela tecnologia;
7. Alteração no escopo;
8. Falha na definição dos dados que devem ser incluídos no Data Webhouse;
9. Falha na adequação de hardware e software para atender as definições do sistema;
10. Falha na integração entre TI e *marketing*;

¹⁹ Nota da autora: XML ou Extended Markup Language é uma linguagem para a marcação de documentos.

11. Falha de flexibilidade no sistema (deve ser extensível para abranger novos dados e novas áreas);
12. Falha na extração de dados;
13. Excesso de dados irrelevantes (deve-se definir a granularidade dos dados) etc.

O Data Webhouse é uma tecnologia muito recente, ainda em fase de amadurecimento. Mas de acordo com a bibliografia pesquisada, os responsáveis pelas organizações que operam seus negócios de forma B2C perceberam que necessitam de ferramentas como esta para gerenciar melhor suas informações a respeito dos clientes a fim de torná-los fiéis aos seus produtos e serviços.

5.7 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Nota-se atualmente que os negócios realizados de forma B2C fornecem às empresas uma rica fonte de dados comportamentais oriundos de uma seqüência de cliques cronológicos.

É neste cenário que a TI está progredindo cada vez mais com o processo de mudança, principalmente mudando as organizações e sua maneira de administrar. Neste contexto o conceito de *marketing* está sendo “repensado” e adaptado vindo como resposta em termos de estratégia para administrar as organizações que operam no comércio eletrônico.

O modelo integrado apresentado neste capítulo evidencia que o Data Webhouse, apesar de ser uma tecnologia recente, se apresenta como uma tentativa de solução a ser adotada pela empresa focada no cliente. Inicialmente sua implementação pode vir com resposta às necessidades de *marketing* - conhecer bem o cliente é um passo estratégico para a maior eficiência dos negócios

Estudar e conhecer a integração da TI com *marketing* pode ajudar os responsáveis pelas empresas que operam no comércio eletrônico a descobrirem novas formas de competir neste mercado globalizado. A adoção desta tecnologia pode ser um grande salto para as organizações que operam de forma B2C.

Esta base de dados mostra que é capaz de armazenar importantes informações dos clientes que posteriormente podem ser processadas em busca de informações mais rentáveis. Tabelas dimensionais são criadas para “traduzirem” a seqüência de cliques em comportamento humano. A partir de então, o Data Mart de Seqüência de Cliques é modelado para suportar a análise feita por ferramentas específicas. Outros Data Marts poderão ser

modelados a partir das tabelas dimensionais apresentadas, o que vai diferir é o grão da tabela Fato.

Este capítulo abordou também que se um líder de projeto for muito bom em lidar com pessoas, dominar a tecnologia e for bom em gerenciar projetos, terá grandes chances de obter sucesso em um projeto de Data Webhouse. Somente as empresas que tiverem excelência no entendimento das necessidades dos clientes e atendimento dessas, ganharão vantagem competitiva, caso contrário os clientes trocarão de empresa sem maiores problemas.

Mesmo que o Data Webhouse ainda se apresente em fase de “amadurecimento” a adoção deste pode apresentar insegurança - mas confia-se, e espera-se, bons resultados dentro do período de quatro anos.

Portanto, a integração do Data Webhouse com o CRM, justifica-se por fornecer dados estratégicos para os negócios da empresa, auxiliando na tomada de decisão, principalmente no que se refere as informações comportamentais.

6. CONCLUSÃO

Hoje o comércio eletrônico não é só uma questão de “moda”, como muitos acreditavam. Nem mesmo uma questão de se salvar a economia. Realizar negócios na Internet significa assumir uma nova postura para responder as tendências e exigências atuais. Para as empresas, comércio eletrônico significa muito mais do que vender para seus clientes ou para outras empresas de forma eletrônica. Na verdade para as organizações que desejam operar neste cenário, comércio eletrônico significa adequar as estratégias de mercado para alinhar os interesses da empresa com os interesses dos clientes, tratar a questão da adoção e do relacionamento com seus clientes e fornecedores, adequar produtos e serviços para este cenário, alterar a estratégia e comprometimento organizacional e principalmente tratar a questão da tecnologia para assegurar tanto clientes e fornecedores quanto a própria organização. Neste sentido, as organizações, que operam seus negócios nesse cenário, estão abrindo um canal extremamente poderoso para a criação e o desenvolvimento de relações com seus clientes. O sucesso dessas organizações vai depender em grande parte da qualidade da resposta fornecida a seus clientes individuais. Desta forma, para se alcançar essa qualidade da resposta, é necessário um investimento no que diz respeito ao gerenciamento do relacionamento com os clientes, ou seja, a utilização de uma estratégia de *marketing* no comércio eletrônico.

As empresas que operam de forma B2C têm uma excelente oportunidade de coletar dados valiosos sobre o cliente, uma vez que a interação dos usuários com os *sites* de comércio eletrônico, através de seqüência de cliques, fornece uma imensa fonte de dados comportamentais como hábitos de compra, preferências e necessidades referente aos mesmos. As organizações que não possuem recursos necessários para gerenciar essa enorme quantidade de informações certamente não possuem vantagem competitiva. Com o advento da TI torna-se possível gerar recursos que atendam às necessidades destas organizações, ou seja, entender a seqüência de cliques, capturá-las e armazená-las em um banco de dados com fins analíticos. A TI adequada para o armazenamento e gerenciamento tipo de informação é o Data Webhouse ou Data Warehouse capacitado para a Web.

A análise da aplicabilidade do Data Webhouse proposta por KIMBALL & MERZ (2000) aliada a estratégia de CRM, mostrou que pode auxiliar a empresa e entender melhor os hábitos e necessidades de compra de seus consumidores e identificar os CMV e CMP. Os resultados são negócios bem posicionados para manter a empresa competitiva e com um diferencial estratégico.

Para se atingir a análise da aplicabilidade do Data Webhouse enfatizando o CRM, inicialmente este trabalho apresenta os principais conceitos, formas e aspectos do comércio eletrônico. A partir de então as empresas deparam não apenas com as contribuições que o comércio eletrônico apresenta, mas também com diversas problemáticas ligadas à adoção, ao relacionamento humano, à adequação, às estratégias, ao comprometimento organizacional, à questão da privacidade e segurança, aos sistemas eletrônicos de pagamento, aos aspectos legais e também aos aspectos de implementação. Para cada caso, diversas são as medidas que vêm sendo providenciadas. Muitos desses problemas requerem estratégias distintas.

No caso da segurança de dados que transitam no comércio eletrônico, as empresas cada vez mais, vêm adotando medidas distintas de acordo com sua necessidade, para que o consumidor se sinta de alguma forma “seguro” ao realizar suas transações econômicas. Apesar dos consumidores estarem cientes que medidas de segurança são tomadas, o comércio eletrônico não atinge seu potencial devido a “desconfiança” dos usuários.

Após o cliente estar possivelmente “seguro” para realizar suas compras no comércio eletrônico, outro problema surge para as empresas: o relacionamento humano. A questão de se relacionar com os clientes no comércio eletrônico é de grande importância. No comércio tradicional, existe um contato físico, onde o vendedor é capaz de conhecer as preferências e necessidades dos consumidores através de algumas perguntas. Já no comércio eletrônico o contato é virtual.

A medida que a empresa é capaz de acessar suas informações de forma precisa utilizando a TI, ela passa a ter melhores condições de desenvolver um relacionamento com seus clientes. Com base na literatura pesquisada, através da TI apoiada por uma estratégia de *marketing*, as empresas que atuam no comércio eletrônico são capazes de “conhecer” seus consumidores, ou seja, são capazes de identificar, diferenciar, interagir e personalizar seus produtos e serviços de acordo com o perfil de seus principais clientes.

A grande motivação deste trabalho é mostrar que existe uma ferramenta, o DW que é uma TI capaz de integrar e armazenar dados históricos de “seções de cliques” realizadas por consumidores no comércio eletrônico, apoiada por uma estratégia, o CRM, capaz de realizar a diferenciação de clientes. O resultado da pesquisa bibliográfica apontada neste trabalho, mostrou que as empresas estão conscientes da necessidade de intensificar o foco nos seus principais clientes para garantir a fidelidade dos mesmos, utilizando para tal a integração da TI ao *marketing*.

Personalização do *site*, aumento da interação entre cliente e empresa, *marketing* voltado para o consumidor são algumas das tendências previstas no passado, por “gurus” da Administração, que atualmente vêm se concretizando dentro das organizações. É na era do

cliente, onde o indivíduo triunfa, que os executivos precisam manter a sua intuição sobre a tendência do mercado e das atitudes dos indivíduos.

Conhecer os clientes é fundamental às empresas para realizar a personalização de seus produtos e serviços e até mesmo realizar programas de promoções com o objetivo de garantir a fidelidade dos mesmos, conseqüentemente adquirir vantagem competitiva para a empresa.

De acordo com esse objetivo, a proposta de analisar de forma crítica a aplicabilidade da ferramenta DW no comércio eletrônico, utilizando a estratégia do CRM, torna -se de grande importância para a compreensão de registros comportamentais. Para as empresas, a análise completa do comportamento de um cliente em um *site* de comércio eletrônico, possui um grande significado econômico. A habilidade para desenvolver e utilizar estratégias com o CRM, desempenha um papel preponderante e auxilia as empresas a atingir seus objetivos. Através do CRM a empresa pode conhecer o perfil de seu cliente, e a partir daí aplicar as estratégias do CRM para fidelizar um cliente no comércio eletrônico.

A integração entre a tecnologia de DW e a estratégia oferecida pelo CRM, proporciona às empresas meios de se relacionarem com o cliente no comércio eletrônico. Não se trata de um novo conceito, e sim de uma “aliança” entre tecnologia e *marketing* que vem como resposta aos novos desafios do comércio eletrônico.

Embora as empresas que atuam no comércio eletrônico tenham descoberto a importância dessa “aliança”, existe ainda outras variáveis inseridas neste contexto, que são reconhecidas como problemas internos à empresa. São elas:

- Dificuldades para mudar a cultura atual da empresa;
- Dificuldade para mudar os processos de negócio existente;
- Dificuldades para se adaptar e atualizar seus sistemas;
- Custo do projeto.

Ao implementar a ferramenta DW utilizando a estratégia de CRM no comércio eletrônico, nota-se que quem constrói as relações com os clientes são os analistas e não a tecnologia. A tecnologia de informação fará o papel de suporte aos analistas de *marketing* e de negócio.

Por isso a necessidade de se envolver a alta administração da empresa, a área de informática, analistas de *marketing* e as necessidades dos clientes, num projeto com esse é fundamental e necessário - isso pode projetar o sucesso ou o fracasso do empreendimento. Desta forma, a implementação de projetos desse porte devem ser orientadas por etapas pequenas. Os passos dados devem ser pequenos, e sempre há a preocupação com o retorno dos investimentos realizados.

Portanto o DW bem estruturado de acordo com os objetivos da empresa que atua no comércio eletrônico juntamente com a estratégia de CRM é capaz de:

- contemplar as análises de campanhas de *marketing*;
- identificar o perfil do cliente;
- realizar análise de vendas;
- garantir lealdade;
- aumentar o nível de satisfação dos clientes.

Um fator a ser mencionado é que não se deve “enxergar” o Data Webhouse como uma solução milagrosa que vai tornar a empresa focada no cliente uma grande potência. É apenas uma tentativa de solução e um caminho confiável que abre espaço para a experiência, visão e intuição dos executivos de todas as empresas que operam seus negócios de forma B2C. Números e simulações devem ser usados para confirmar e comprovar a intuição, mas não se deve tornar “escravos” dos sistemas de informação e das poderosas ferramentas que estão à disposição. A TI deve trabalhar para o indivíduo e não o contrário. Só desta forma, as organizações se tornam aptas para atuar no ambiente competitivo e turbulento, criando e sustentando a vantagem competitiva. A informática demonstra que é cada vez mais, um ponto de apoio aos empresários e quem souber usá-la com certeza sairá na frente dos concorrentes.

Finalmente, considera-se que a partir desse trabalho pode-se originar novos trabalhos, como:

- A utilização da ferramenta Data Mining no comércio eletrônico para extrair do DW novas informações, não triviais e úteis a respeito do comportamento dos clientes. Como o volume de dados gerado diariamente pelo comércio eletrônico é muito intenso, a utilização de uma ferramenta para minerar os dados contidos no DW é de grande importância para realizar uma análise de dados “mais limpos” para então serem preparados para uma análise mais precisa;
- A utilização da ferramenta OLAP no comércio eletrônico para verificar se as vendas realizadas pelo processo *cross-selling* são bem sucedidas, ou seja, se os clientes realmente estão comprando através desse procedimento;
- Analisar se através dos *cookies* as empresas realmente estão desenvolvendo um relacionamento com o cliente e não com a máquina;
- Analisar o impacto gerencial da adoção do Data Webhouse nas empresas: quais os fatores mais importantes, estudo de casos antes e depois da implementação, dificuldades, etc.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERTIN, A.L. (1998). Comércio eletrônico: seus aspectos de segurança e privacidade. **Revista de Administração de Empresas**, v.38, n.2, p.49-61.
- ALBERTIN, A.L. (1998a). Comércio eletrônico: benefícios e aspectos de sua aplicação. **Revista de Administração de Empresas**, v.38, n.1, p.52-63.
- ALBERTIN, A.L. (2000). O comércio eletrônico evolui e consolida-se no mercado brasileiro. **Revista de Administração de Empresas**, v.40, n.4, p.94 -102.
- ALBERTIN, A.L. (2000a). **Comércio eletrônico: modelo, aspecto e contribuições de sua aplicação**. 2.ed. São Paulo, Atlas.
- ALCOFARADO, I.; SÊMOLA, M. (2000). **PKI: a chave para a segurança no e-commerce**. <http://www.modulo.com.br/noticia/a-semivan.htm> (24 Ago).
- ALMEIDA, S. (1997). **O Melhor sobre clientes: 500 citações, textos e comentários**. Salvador, Casa da Qualidade.
- ALVES, A. (1997). Prospecção. **Developers' Magazine**, v.1, n.6, p.40 – 41, fev.
- AMARAL, J.A. (1997). Desmistificando definitivamente o data warehousing. **Developers' Magazine**, v.1, n.6, p. 14 -17, fev.
- APPLEGATE, L.M. et al. (1996). Eletronic commerce: building blocks of new business oportunity. **Journal of Organizational computing and Eletronic Commerce**, v.6, n.1, p.1-10.
- BAPTISTELLA, M.M.T. (2000). Comércio eletrônico: motivos para utilização e tendências futuras. (CD-ROM) In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 20., São Paulo, 2000. **Anais**. São Paulo, TEP- UFF
- BASTOS, A. (1998). **Segurança da informação na Internet e Intranet oportunidades versus riscos**. <http://www.modulo.com.br/noticia/a-opxris.htm> (17 Dez).

- BERSON, A. (1997). **Data Warehouse, data mining and OLAP.**
<http://www.datawarehouse.da.ru> (31 Jan.).
- BISPO, C.A.F. (1998). **Uma análise da nova geração de sistemas de apoio à decisão.**
São Carlos. 144p. Dissertação (Mestrado) - Escola de Engenharia de São Carlos,
Universidade de São Paulo.
- BISPO, C.A.F.; CAZARINI, E.W. (1998). Conceitos básicos e a elaboração de um projeto lógico de um data warehouse. (CD-ROM) In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 18.,/CONGRESSO INTERNACIONAL DE ENGENHARIA DE INDUSTRIAL, 4., Niterói, 1998. **Anais.** Niterói, TEP- UFF.
- BISPO, C.A.F.; CAZARINI, E.W. (1999). Análises sofisticadas com o on -line analytical processing. **Developers' Magazine**, v. , n.32, p.28 -31, abr.
- BR BUSINESS (2000). **Redes Neurais .** <http://www.br-business.com.br> (7 dez.).
- BRETZKE, M. (1992). **O Marketing de relacionamento: integração entre informática e marketing na busca da vantagem competitiva.** São Paulo. 200p. Dissertação (Mestrado) – Fundação Getúlio Vargas.
- BRETZKE, M (1999). **Marketing de relacionamento transformando a organização para competir em tempo real: integração entre tecnologia de informação e marketing.** São Paulo. 268p. Tese (Doutorado) – Fundação Getúlio Vargas..
- BRETZKE, M. (2000). **Marketing de relacionamento e competição em tempo real com CRM (Customer Relationship Management) .** São Paulo, Atlas.
- CABRERA, M. (2000). **Noções sobre data warehouse.** <http://www.datawebhouse.com.br> (20 jul.).
- CAMERON, D. (1997). **Eletronic commerce: the new business platform of Internet.** Charleston, Computer Technology Research.

- CERVO, A.L.; BERVIAN, R.A. (1983). **Metodologia científica**. São Paulo, McGraw-Hill do Brasil.
- CERVO, A.L.; BERVIAN, R.A. (1996). **Metodologia científica**. 4.ed. São Paulo, Makron Books do Brasil.
- CHEIN, L.; CABRERA, M. (2000). **Como implementar idéias eficientes**. <http://www.intermanagers.com.br> (22 maio).
- CHEIN, L.; CASTRO F.R. (2000). **A Era da personalização do conteúdo**. <http://www.intermanagers.com.br> (30 out.).
- CHUCK, M. (1999). **Canal executivo**. <http://www.uol.com.br/canalexecutivo> (20 jul.).
- CIELO, I. (2001). **Business intelligence** <http://www.datawarehouse.inf.br> (31 Jan.).
- CIELO, I. (2001a). **Data warehouse X data mart: como eles podem ajudá-lo**. <http://www.datawarehouse.inf.br> (27 Jan.).
- CORRÊA, N. (2000). **Você está seguro para realizar e-commerce ?** <http://vil.nai.com/vil/vpe10175.asp> (15ago.).
- DAL'ALBA, A. (1998). **Site sobre data warehouse**. <http://www.geocities.com/SiliconValley/Port/5072> (2 mar.).
- DATA WAREHOUSE. (1997). *Revista Byte Brasil*, v.6, n.1, jan.
- DEPOMPA, B. (1998). **Companies that start whit data marts find it difficult to move to a warehouse: information week**. <http://techweb.cmp.com/iw/605/05olwa1.htm> (11 Nov.).
- DINIZ, D. (2000). **Tabelas dimensionais ou faces de um cubo**. <http://www.datawarehouse.inf.br> (17 dez.).
- DINIZ, D. (2000a). **Dimensão: como ela funciona**. <http://www.datawarehouse.inf.br> (15 jul.).

DW (2000). **Uma breve história do data mining**. <http://www.datawarehouse.inf.br/> (20 mar.).

EDELHERTZ, P. (2000). **A Convergência entre e-commerce e CRM**. <http://www.1to1center.com.br> (26 jun.).

ESCOLA REGIONAL DE INFORMÁTICA DE SÃO PAULO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. (2001). **Minicursos: coletânea de textos**. São Carlos, ICMC/USP.

FAYAD, U.; SHAPIRO, G.P. (1996). **Advances in knowledge discovery and data mining**. s.l., AAAI/ MIT Press.

FELTON, A.P. (1959). **Marking the concept work**. **Harvard Business Review**, v.37,p.65, July, p.65, July

FLETCHER, K. (1990). **Marketing management an information technology**. Cambridge, Prentice Hall International.

FONSECA, G.R. (1997). **Uma visão integrada de ambientes e repositórios de dados**. **Developers' Magazine**, v.1, n.6, p.18-22, fev.

GIL, A.C. (1991). **Métodos e técnicas de pesquisa social**. São Paulo, Atlas.

GONÇALVES, L.M. (2000). **E -business ou out-of-business: para onde vai a sua empresa?** **Developers' Magazine**, v.4, n.42, p.10-11, fev.

HODEL, A. (1998). **Data mining: a new weapon for competitive advantage**. <http://www.software.ibm.com/sq/issues/vol24/Data.htm> (16 Jan.).

IBM (1996). **Garimpendo dados**. <http://www.ibm.com.br> (13 out.).

IBRAHIM, F. (2000). **Uma visão macro do CRM analítico**. <http://cf.doctordata.com.br/crm/> (6 dez.).

- INMON, W.H. (1997). **Como construir o data warehouse**. Rio de Janeiro, Campus.
- INMON, W.H. (1992). **Building the data warehouse**. Local, John Wiley & Sons.
- KIMBALL, R. (1998). **Data warehouse toolkit**. São Paulo, Makron Books.
- KIMBALL, R. (1998a). **Meta meta data data**. DBMS Magazine.
<http://www.dbmsmag.com/9803d05.html> (27 Ago.).
- KIMBALL, R. (1999). **Clicking with your customer**.
http://www.intelligententerprise.com/db_area/archives/1999/990501/warehouse.shtml (01 May).
- KIMBALL, R.; MERZ, R. (2000). **Data webhouse: construindo um data warehouse para a Web**. Trad. por Edson Furmankiewicz e Joana Figueiredo. Rio de Janeiro, Campus.
- KOLISH, A.K.; JAWORSKI, B.J. (1990). Market orientation: the construct, research propositions and managerial implications. **Journal of Marketing**, v.54, p.1 -18, Apr.
- KÖCHE, J.C. (2000). **Fundamentos de metodologia científica: teoria da ciência e prática da pesquisa**. 18.ed. Petrópolis, Vozes.
- KOTLER, P.; ZALTAMANN, G. (1971). Social marketing: na approach to planned social changes. **Journal of Marketing**, v.35, p.12, July.
- KRISHNAMURTHY, R.; IMEILINSKI, T. (1991). **Research directions in knowledge discovery**. New York, SIGMOD Record.
- LAMBERT, B. (1997). **Data warehousing fundamentals: what you need to know to succeed**. Data Management Review. <http://www.data-warehouse.com/resource/articles/lambero.htm> (29 Dez.).
- LICO, A. E. O. P. (2001). **Comércio eletrônico como nova era de negócios: impacto nas empresas e fatores de sucesso**. Campinas. 99p. Dissertação (Mestrado) - Pontífca Universidade Católica.

- MAIA R.; MAIA F. (1981). **Programa auxiliar de pesquisa estudantil**. 3.ed. São Paulo, DCL. (Coleção PAPE, 3).
- MAY, P. (2000). **The Business of ecommerce: from corporate strategy to technology**. Cambridge University Press. New York, Sigs Books.
- MCELREATH, J. (1997). **Data warehouses: an architectural perspective**. Perspectives. http://www.csc.com./about/tech_dw_arch.html (29 Dez.).
- MCKARTHY, E.J.; WILLIAMS, D.P. (198 2). **Basic marketing**. 7.ed. Homewood, Rochard D. Irwin. p.6.
- MCKENNA, R. (1991). **Relationship marketing: succesful strategies for the age of the customer**. New York, Addison-Wesley.
- MORENO, C. (1998). SET: o caminho da maturidade para o comércio eletrônico. **Developers' Magazine**, v.3, n.25, p.22-25, set.
- NIMER, F. (1998). Analisando o retorno sobre o investimento de data warehouse. **Developers' Magazine**, n.18, p.16-17, fev.
- ORR, K. (1996). **Data warehousing technology**. The Ken Orr Institute. <http://www.kenorrinst.com/warehousing.html> (7 Dez.).
- PACE, M. (2000). **Foco no cliente não é só moda**. <http://www.intermanagers.com.br> (6 dez.).
- PARENTE, D. (2001). **Data warehouse: identificando os possíveis pontos que podem levá-lo ao fracasso**. <http://www.datawarehouse.inf.br> (30 jan.).
- PEPPERS, D. (2000). **Faça mais pelo seu cliente**. **Automação comercial, bancária e tecnologia de informação**, v.7, n.77, p.50, mar.

- PEPPERS, D.; ROGERS, M. (2000). **Um guia executivo para entender e implantar estratégias de customer relationship management**. São Paulo. Peppers & Rogers Group do Brasil. (CRM Series Marketing 1 to 1)
- PEPPERS, D.; ROGERS, M. (2000a). E a Internet reinventou os negócios. **HSM Management**, n.21, p.20-36, jul./ago.
- PEPPERS, D.; ROGERS, M. (2001). **Um guia executivo para entender e implantar estratégias de customer relationship management**. 2.ed. São Paulo, Makron Books. (CRM Series Marketing 1 to1).
- SANTANA, F.F. (2001). **Customer relationship management** – definição de CRM. http://www.tinfonews.com.br/tt_crm03pv.asp (24 Jan.).
- SILVA, E. C.; SACOMANO, J. B. (2001). Núcleo de Estudo como Apoio à Pesquisa na Graduação em Engenharia. XXIX COBENGE – Congresso Brasileiro de Ensino de Engenharia – da teoria à prática: experiências concretas no ensino de engenharia. PUC/RS, Porto Alegre-RS.
- SILVA, E.L.; MENEZES, E.M. (2001). **Metodologia da pesquisa e elaboração de dissertação**. 2.ed. Florianópolis, Laboratório de Ensino a Distância da UFSC.
- SILVA, F.C. (2000). **Utilização de datamining no comércio eletrônico como forma de criação de valor**. São Paulo. 128p. Dissertação (Mestrado) – Fundação Getúlio Vargas.
- SANTOS, J.; HENRIQUES, N.A.C.; REIS, V. (2000). **Data mining/data warehousing**. <http://students.fct.unl.pt/users/nach/dmdw/prologo> (1 Abr.).
- SOUZA, J.M.; SULAIMAN, J.A. (1997). Prospecção de conhecimento em banco de dados. **Developers' Magazine**, v.1, n.6, p.38-39, fev.
- STEVENS, A. (2000). **The oacle datamart suite - a case study**. <http://www.avanco.com/dmcase.htm> (12 Dez.).

- TAPSCOTT, D. (1996). **The digital economy**: promise and peril in the age of networked intelligence. New York, McGraw-Hill.
- TAURION, C. (1997). Data warehouse: estado da arte e estado da prática. **Developer's Magazine**, v.1, n.6, p.10-11, fev.
- TAURION, C. (1998). Sua empresa está preparada para o comércio eletrônico? **Developers' Magazine**, v.3, n.25, p.10-11, set.
- TAURION, C. (2000). **Os estágios de uma estratégia de CRM**.
<http://cf.doctorData.com.br/crm/> (23 nov.).
- TAURION, C. (2000a). **Preferências ao clicar do mouse**.
<http://www.computerworld.com.br> (18 dez.).
- TAURION, C. (2001). **Customer relationship management** – justificando o CRM na empresa. http://www.tinfonews.com.br/tt_crm04pv.asp (24 jan.).
- TEC. (1999). **Tecnologias data warehouse e OLAP**.
<http://jacques.ic.cti.br/ic/pqps/atps/dw.htm> (1 set.).
- THIOLLENT, M. (1983). Problemas de metodologia. In: FLEURY, A.C.C.; VARGAS, N. **Organização do trabalho**: uma abordagem interdisciplinar - sete casos brasileiros para estudo. São Paulo, Atlas.
- TIEZZI, G. (1997). O Planejamento estratégico da informação. **Developers' Magazine**, v.1, n.6, p.24-25, fev.
- TRONCHIN, V. (2000). **Como diferenciar sem discriminar?**
<http://www.1to1center.com.br/> (14 nov.).
- TRONCHIN, V. (2000a). **Marketing one-to-one versus marketing de massa**.
<http://www.1to1center.com.br/> (24 ago.).
- VIEIRA, F. (2001). **Alguns conceitos sobre DW**. <http://www.datawarehouse.inf.br> (30 jan.).

WELDON, J.L. (1998). **A career in data modeling.**

<http://www.byte.com/art/9706/sec7/atr3.html> (19ago)

WIDEL, L. (2000). **Bancando uma solução de customer relationship management .**

<http://www.1to1center.com.br> (5 dez.).