

2

Referencial Teórico

Neste capítulo é descrita a fundamentação teórica da presente pesquisa, apresentando definições e conceitos centrais de cada um dos assuntos que compõem o arcabouço teórico necessário para a coleta, análise e interpretação dos dados.

2.1

Sistema Integrado de Gestão (ERP)

Com o crescimento das empresas, e conseqüentemente com o crescimento do volume de informações gerada pelas mesmas, seus gestores passaram a se preocupar em como gerir a informação produzida, de modo a sustentar e melhorar o desempenho organizacional.

Por décadas as empresas quiseram integrar seus sistemas, e então surgiu o conceito de sistema integrado de gestão para atender a esta demanda. Segundo Bancroft, Seip e Sprengel (1998), a idéia de sistemas de informação integrados existe desde o início da utilização dos computadores na década de 60, porém uma série de dificuldades, inclusive tecnológicas, não permitiu que a idéia fosse posta em prática pelas empresas naquela ocasião.

Os Sistemas Integrados de Gestão (ERP – Enterprise Resource Planning) originaram-se na Europa, na indústria de manufatura, em 1979, quando a empresa alemã SAP (System, Anwendungen, und Produkte in Datenverarbeitung) desenvolveu um software que integrava os dados operacionais e financeiros e um único banco de dados. Os ERPs consistem de uma série de módulos como finanças e contabilidade, compra de bens e serviços, venda, produção, recursos humanos, e etc (de acordo com a necessidade da empresa). Ao invés de se concentrar em áreas funcionais isoladas, o sistema tem seu foco nos processos de negócio que podem se estender por várias áreas funcionais da empresa.

A partir da década de 90, diversas empresas iniciaram a adoção de sistemas ERP em função da necessidade da atualização tecnológica dos sistemas de informação, das pressões competitivas, da necessidade de reduzir custos e de

integrar suas informações. Os ERPs cumprem o papel tecnológico, agilizando o fluxo de informações na organização com a adoção de padrões de troca e compartilhamento de dados, implantando processos amplos e interfuncionais nas organizações (Carneiro, 2005).

A tecnologia tem um papel fundamental nos processos empresariais, e influencia tanto a forma de realizar o trabalho quanto a maneira de gerenciá-lo, mas a real integração de uma organização depende da gestão eficiente de processos e da adoção de um estilo de gestão que incentive o compartilhamento das informações. Por causa disso os ERPs são considerados uma ferramenta de redesenho de processos por excelência. Quando utilizados de forma consistente na empresa, permitem que as pessoas assumam mais responsabilidades, adotem mecanismos mais eficazes de participação na realização do trabalho e empreguem melhores meios de comunicação e produção (Carneiro, 2005).

Um dos grandes problemas encontrados em projetos de implementação do ERP é o foco dado ao aspecto do sistema de informação e da substituição dos sistemas existentes, deixando de lado a revisão e ajuste de seus processos internos (Davenport, 2002). Segundo Gupta (1998) um projeto de ERP não trata de implementação de tecnologia. Trata-se de gestão de mudanças, que cuida de pessoas, processos, procedimentos e novas maneiras como as tarefas se desenvolvem.

Os ERPs, por serem “pacotes de solução”, requerem procedimentos de ajuste para que possam ser utilizados em uma empresa específica: parametrização, customização e localização. A parametrização é o processo de adaptação do sistema com os valores já disponibilizados por ele; a customização é a modificação que é feita pela empresa para ajustar o sistema às suas necessidades; e a localização é a adaptação do sistema ao país onde está sendo implantado (para atender a questões legais).

A implantação de um ERP é realizada em etapas bem definidas: decisão e seleção, implementação, estabilização e utilização (Souza e Zwicker, 2002).

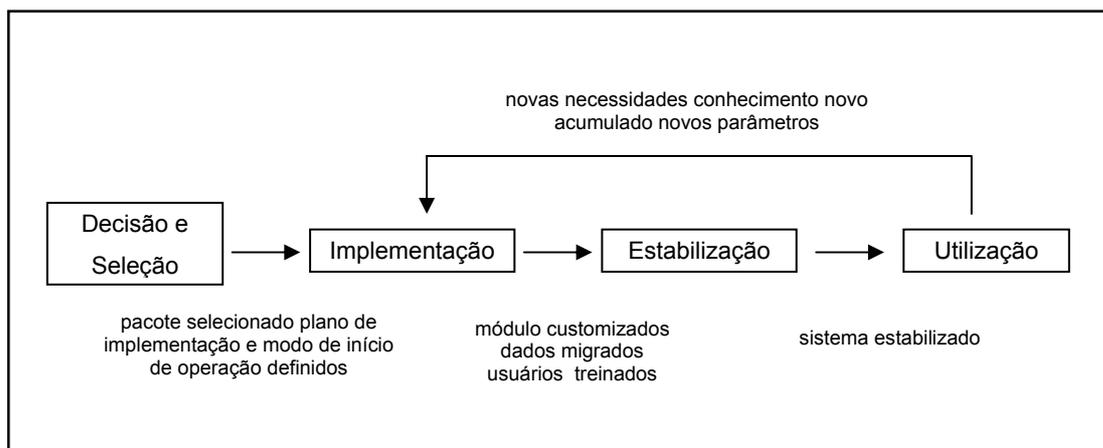


Figura 2 – Ciclo de vida dos sistemas ERP (Souza e Zwicker , 2002)

Na primeira etapa a empresa decide pela implementação de um ERP e escolhe o fornecedor. Na etapa seguinte, a empresa define os módulos do sistema que serão colocados em funcionamento, parametriza e/ou customiza o ERP (eventualmente adapta os processos de negócio ao sistema), prepara a carga inicial dos dados, treina os usuários e disponibiliza suporte para a entrada em operação. Após a entrada em operação, na fase de estabilização, todo o corpo gerencial e técnico da empresa cuidam para que o sistema funcione normalmente, resolvendo rapidamente qualquer problema que possa surgir. Finalmente, na última etapa o sistema já faz parte das operações da empresa e dia após dia seus ganhos são percebidos pelos seus usuários, durante o processo de utilização.

Segundo Willis e Willis-Brown (2002), existem duas ondas no processo de adoção de um ERP. A primeira se refere às mudanças que a organização precisa realizar para se ajustar em um ambiente de negócios integrado do ERP. A segunda onda se refere às ações subseqüentes à implantação do sistema que possibilitam a organização alcançar todo o potencial do sistema e os benefícios prometidos. Esta fase, consiste na extensão do ERP através da implantação de novos módulos não contemplados na solução original, como por exemplo pacotes de BI (Business Intelligence).

2.2

Os Sistemas de Apoio à Decisão

No começo do século XX, os critérios de tomada de decisão eram centrados no executivo principal, em geral o proprietário, que detinha as prerrogativas de escolher o que julgasse ser o melhor para sua empresa e para os seus trabalhadores (Pereira e Fonseca, 1997). Com o ambiente estável e as informações restritas, supunha-se que os decisores tivessem conhecimento suficiente das alternativas e suas conseqüências, de modo a escolher a melhor alternativa para cada problema.

A partir da década de 60 os trabalhadores, que antes eram tratados apenas como “mão-de-obra”, passaram a ser considerados como seres humanos capazes de pensar, de decidir e de serem motivados (Pereira & Fonseca, 1997). Com isso a informação precisou ser compartilhada pelos diversos níveis hierárquicos da organização, e ferramentas surgiram para apoiar o processo decisório nos diversos níveis.

As decisões acontecem a todo o momento em uma organização e a função dos administradores é essencialmente decisória (Pereira & Fonseca, 1997). Podem ser decisões simples ou complexas, mas afetam diretamente a sobrevivência da empresa e a vida das pessoas que giram em torno dela, sejam empregados, acionistas, clientes ou fornecedores. Na tomada de decisão, as dificuldades surgem devido à complexidade das alternativas. Decisões podem ser tomadas sob pressão ou então podem ser influenciadas pelo estado emocional do decisor, e quando faltam informações, as decisões acabam sendo tomadas longe da forma ideal. Desta forma, quanto mais informações de qualidade tiver o tomador de decisão, e quanto mais lógico e sistemático for o tratamento dado à forma de tomar decisões, maiores serão as chances de se tomar a melhor escolha (Parrini, 2002).

O processo decisório é sempre estressante para os decisores (Pereira e Fonseca, 1997). Existem alguns fatores que influenciam diretamente ou indiretamente na solução do problema:

Fatores que exercem influência individual ou coletiva
1) inteligência e nível cultural
2) nível social
3) sexo
4) costumes e crenças
5) ética moral e ética profissional
6) saúde física e mental
7) influência familiar
8) fator emocional
Fatores que exercem influência e que são inerentes das empresas
1) necessidade de produtos e serviços de qualidade
2) necessidade de atendimento rápido e personalizado
3) necessidade de preços competitivos
4) condições impostas por clientes e/ou fornecedores
5) exigências do governo e da sociedade
6) concorrência
7) notícias da mídia
8) cultura organizacional
9) tecnologias utilizadas
10) melhor emprego dos recursos existentes
11) normas da empresa
12) legislação em vigor
13) meio ambiente

Tabela 3 – Fatores que exercem influência na tomada de decisão (Bispo, 1998)

Administrar todos os fatores acima no momento da decisão é o desafio de todo gerente no mundo moderno. Desta forma, tornou-se um fator importante na gestão das empresas a utilização de ferramentas que forneçam as informações adequadas para suportar o processo decisório.

Segundo Gates (1997) a informação é um produto valioso no mundo contemporâneo, embora não seja tangível e nem mensurável, uma vez que proporciona poder. A busca de informações é atualmente o alvo da maior parte das empresas, e é através do seu uso que é possível ter um suporte adequado ao processo decisório. É função das ferramentas que suportam este processo, disponibilizar as informações necessárias de forma rápida e confiável, e exibi-las de forma compreensível, facilitando assim a tomada de decisão.

Os sistemas de apoio à decisão (SAD) começaram a ser desenvolvidos no início da década de 70, e passaram a ser caracterizados como sistemas

computacionais interativos que apoiavam no processo decisório de problemas considerados não-estruturados (Sprague e Watson, 1991 apud Bispo, 1998). Os primeiros SADs eram muito caros, de uso muito específico e difíceis de operar. Normalmente eram construídos especificamente para a empresa, por ela mesma ou por terceiros, para auxiliar no gerenciamento de problemas específicos, e posteriormente eram adaptados para englobar outras necessidades. Com isso tornavam-se grandes e complexos, e de difícil manutenção. Os principais problemas destes sistemas eram a inadequação da arquitetura das bases de dados para a realização de consultas não estruturadas (*ad hoc*), ausência de histórico de dados e a dificuldade na geração de relatórios.

Com o surgimento dos Sistemas Gerenciadores de Banco de Dados (SGDB), na década de 80, a tarefa de armazenamento e recuperação dos dados gerados pelas empresas ficou facilitada, porém a geração de relatórios ainda era feita de forma artesanal, colecionado e formatando as informações de acordo com a necessidade em questão. Os ambientes de informação tinham sua estrutura (modelagem de dados) baseada para suportar os processos operacionais da empresa, o que dificultava a obtenção de informações pelo corpo gerencial. A elaboração de relatórios contendo informações de áreas diferentes (exemplo: vendas e situação das contas a receber; compras e contas a pagar, etc) era lenta, e ainda faltava uma ferramenta que realmente auxiliasse os gerentes no processo de obtenção de informações, fornecendo-as de forma fácil, rápida e confiável.

Nesta ocasião, começaram a ser construídos os primeiros Sistemas de Informação para Executivos – EIS, que eram sistemas desenvolvidos especialmente para atender à demanda por informação dos executivos da empresa, mas ainda assim as necessidades de informação eram mais urgentes do que a velocidade com que estes sistemas conseguiam ser desenvolvidos.

Nas décadas seguintes, pesquisadores e desenvolvedores ampliaram o conceito de sistemas de apoio à decisão para qualquer sistema que contribuísse com o processo decisório, e que possuísse as seguintes características (Bispo, 1998):

- Ser voltado para problemas menos estruturados;
- Combinar o uso de modelos ou técnicas analíticas às funções tradicionais de acesso e recuperação de informações;

- Utilizar recursos para facilitar o acesso por usuários não especializados em computação e
- Possuir flexibilidade e adaptabilidade para acomodar mudanças no ambiente e na abordagem do processo decisório.

Durante as décadas de 80 e 90, foram desenvolvidos diversos sistemas para dar suporte às operações das empresas e dentre eles, conforme já foi mencionado anteriormente, surgiu o ERP, que é um sistema de gestão integrada. Nesta época surgiu também uma nova geração de ferramentas que se complementam, dando suporte à decisão: o data warehouse, o OLAP⁴ (On-Line Analytic Processing) e o data mining⁵ (Bispo, 1998). A utilização destas ferramentas possibilita que os tomadores de decisão gastem um tempo menor com a manipulação de dados, e permite que sejam feitas análises necessárias para a viabilizar a geração de soluções para seus questionamentos.

No contexto da tomada de decisão, as ferramentas OLAP permitem aos analistas, gerentes e executivos uma otimização no acesso às informações, apoiando o processo decisório. Esta tecnologia permite a realização de diversas operações que otimizam o processo de análise das informações como: definição de filtros e condições baseadas em regras de negócio, inversão de coordenadas e eixos de pesquisa, expansão nas linhas e colunas, ajuste do lay-out, etc.

Uma ferramenta OLAP, no entanto, é apenas um instrumento com a qual o tomador de decisão pode contar, e um processo decisório envolve, além de outras ferramentas, uma série de atividades e abordagens que não são cobertas no escopo desta ferramenta.

Para finalizar, independente da ferramenta, a informação é um recurso importante para o processo decisório, e é função das ferramentas de suporte à decisão, disponibilizar informações íntegras e confiáveis, e de forma rápida e amigável para seus usuários.

⁴ OLAP, processamento analítico on-line, é um termo utilizado para descrever uma tecnologia que permite que os dados possam ser analisados por dimensões de análise diferentes, abordagem multidimensional, suportando o processo decisório (Kimball, 1998).

⁵ Data mining é uma tecnologia que busca encontrar padrões, regularidades e tendências nos mais diversos bancos de dados, identificando correlações entre diversas variáveis de um universo, permitindo descobrir informações novas, extraindo assim mais valor dos dados contidos em um data warehouse (Serra,2002).

2.3

Ambiente de Data Warehouse

As empresas atualmente padecem de um mal clássico. Possuem uma enorme quantidade de dados, mas enfrentam uma grande dificuldade na extração de informações. Mesmo as empresas que implantaram um ERP possuem esta dificuldade uma vez que estes sistemas não trazem a informação na sua forma mais palatável. Ao contrário, as informações vitais para a tomada de decisão estão escondidas em milhares de tabelas e arquivos inacessíveis, dificultando o acesso para os que necessitam da informação (Barbieri, 2001).

A Evolução do ambiente de informações gerenciais

No início da década de 70 surgiu uma nova tecnologia de armazenamento e acesso a dados em discos e com ela, surgiu um novo tipo de software conhecido como SGDB ou sistema gerenciamento de banco de dados. Em meados da mesma década, as aplicações começaram a ser desenvolvidas acessando *on-line* os dados armazenados nestes bancos de dados, o que causou uma ampliação no escopo de atuação dos sistemas de informação (Inmon, 1997). No início da década seguinte, com o surgimento dos PCs (computadores pessoais) e com a descentralização das aplicações o usuário final passou a ter maior controle dos sistemas e dados disponíveis. Surgiram então os sistemas de informações gerenciais, ou MIS – Management Information Systems, também conhecidos como sistemas de apoio à decisão - SAD (Inmon, 1997) , que tinham como objetivo direcionar as decisões gerenciais.

Com o passar dos anos, os bancos de dados foram evoluindo e um dos fatores que contribuíram para esta evolução foi a queda crescente nos custos de armazenagem (Raden, 1998). Com isso, surgiu um novo conceito de banco de dados, com uma forma de armazenamento fisicamente diferente da anterior (dos bancos de dados destinados ao processamento operacional), cujo objetivo era atender as necessidades informacionais ou analíticas de seus usuários. Segundo Inmom (1997) as necessidades informacionais ou analíticas compreendem uma comunidade de usuários diferente, uma vez que os usuários do ambiente de

informações analíticas utilizam as informações armazenadas para apoiar o processo decisório, enquanto os usuários das informações operacionais utilizam o banco de dados para as operações do dia-a-dia.

O sistemas destinados ao processamento analítico atendem às necessidades dos gerentes durante o processo de tomada de decisão. Este tipo de processamento envolve uma grande quantidade de dados e permite a análise das informações armazenadas, a geração de indicadores, a análise de tendências, etc. Nestes sistemas os dados raramente são atualizados pelos usuários, que apenas os consultam. Os requisitos de tempo de resposta também são bastante atenuados, em comparação com os tempos exigidos em um sistema transacional.

Para popular os bancos de dados destinados ao processamento analítico surgiram os programas de extração, que deram aos usuários finais um controle maior dos dados armazenados em suas bases de dados. Com a proliferação das bases de dados e dos programas de extração, chamados por Inmon (1997) de desenvolvimento espontâneo, a arquitetura de informações começou a se tornar confusa, gerando descrédito nos números apresentados e diminuindo a produtividade das equipes na geração destes números. Cada departamento da empresa gerava um número, que era diferente de um outro apresentado por outra área em função do tempo de processamento, ou do programa de extração utilizado. Além destes problemas, ainda existia a dificuldade de passar do dado à informação. Segundo Inmon, 1997, uma informação é um dado “trabalhado” e que carrega consigo algum conteúdo, e para o ele os dados encontrados na arquitetura gerada pelo desenvolvimento espontâneo não são adequados à tarefa de apoio às necessidades de informação em virtude da falta de integração e da diferença entre os horizontes de tempo existentes em cada uma das bases de dados, necessárias ao processamento analítico. Desta forma existe a necessidade da criação de um novo ambiente, e assim surge o ambiente de data warehouse. É importante ressaltar ainda, que Inmon define duas categorias de dados – os dados primitivos e os dados derivados. Os dados primitivos são dados detalhados, utilizados nas operações diárias das empresas. Podem ser atualizados (nas bases de dados operacionais) e atendem às necessidades funcionais, sendo operados por rotinas executadas de forma repetitiva. Os dados derivados são gerados (calculados ou resumidos) com o objetivo de atender às necessidades gerenciais

da empresa. Não são atualizados pelos usuários, em atividades do dia-a-dia, e são geralmente compostos de dados históricos.

2.3.1

Data Warehouse

Um data warehouse é um conjunto de dados baseados em assuntos, integrado, não-volátil, e variável em relação ao tempo, para apoio às decisões gerenciais (Inmon, 1997). Isso significa que sua base de dados é orientada por assuntos de interesse relacionados a cada área de negócio da empresa, que são armazenados de forma integrada, com um horizonte de armazenamento maior que o dos sistemas operacionais da empresa, permitindo análises de séries temporais, avaliação de tendências e comportamentos ao longo do tempo.

Segundo Inmon (1997) uma ferramenta de data warehouse deve ser orientada aos principais assuntos ou negócios da empresa; integrada, através da padronização de termos e estruturas usadas nos sistemas de informação da empresa; não volátil, ou seja, os dados armazenados no data warehouse não podem ser alterados pelos usuários (sejam eles primitivos ou derivados); e variável no tempo, ou seja, as informações armazenadas fazem sempre referência a um momento do tempo (o horizonte de armazenamento também é maior que nos sistemas tradicionais podendo conter até 10 anos de dados).

Em um data warehouse existem diferentes níveis de detalhe dos dados armazenados (muito detalhado, resumido, e altamente resumido ou consolidado). Quando os dados são transferidos do ambiente operacional para o ambiente do data warehouse geralmente eles passam por um processo de transformação, que podem alterar ou não a granularidade⁶ das informações armazenadas. Nos sistemas operacionais, os dados são armazenados no nível mais baixo de granularidade, mas em um data warehouse esta definição deve ser feita durante a fase do projeto, uma vez que irá afetar o volume de dados armazenados e a *performance* futura das consultas. Esta questão deve ser analisada com cuidado uma vez que apesar da baixa granularidade estar vinculada a um maior número de

⁶ A granularidade diz respeito ao nível de detalhe, ou seja, quanto maior o detalhe, mais baixa a granularidade da informação (Serra, 2002).

dados armazenados, com este nível de detalhamento é possível responder a qualquer pergunta.

Os dados, armazenados em um ambiente de data warehouse, oriundos dos sistema de informação transacionais, deverão ser trabalhados, estar integrados e disponíveis, permitindo diversas consultas através do uso de ferramentas de geração de relatórios. Pelos termos usados, entende-se:

- trabalhados – contempla todo o processo de identificação, catalogação, coleta, transformação e disponibilização dos dados em informações necessárias aos negócios da empresa;
- identificação – corresponde ao processo de modelagem de dados voltado para as áreas específicas a serem atendidas pelo data warehouse, na qual os dados serão analisados segundo uma ótica própria, onde os conceitos de dimensionalidade, agregação e desnormalização aparecem como diferencial;
- catalogação – objetiva a identificação clara e concisa das informações gerenciais (busca por palavra-chave ou meta-dado);
- Extração / transformação / carga – definirão os procedimentos de busca, transformação e carga dos dados oriundos das diversas fontes existentes na empresa, além as estratégias e rotinas de consolidação e sumariação e agregação dos dados.

Um dos benefícios proporcionados pelo data warehouse é a diminuição do tempo que os gerentes levam para obter as informações necessárias aos seus processos decisórios, permitindo que a decisão seja baseada em fatos (Bispo,1998). Ao reunir informações dispersas por diversos bancos de dados e plataformas distintas, o data warehouse permite que sejam feitas análises bastante eficazes, transformando dados esparsos em informações estratégicas, antes inacessíveis ou sub-aproveitadas (Bispo,1998).

A modelagem dos dados é uma etapa importante na construção de um sistema de informação. Primeiro desenvolve-se um modelo conceitual de alto nível do processo empresarial ou da atividade que se deseja modelar. Depois utiliza-se este modelo para derivar um modelo lógico, no qual os dados são abordados com mais detalhes. Finalmente, a partir do modelo lógico elabora-se o

modelo físico que deve prover todos os detalhes para a implementação do banco de dados.

No desenvolvimento de um data warehouse, a construção da base de dados se processa em uma arquitetura diferente das encontradas nos sistemas tradicionais. É utilizada uma modelagem orientada para os fatos ocorridos e que podem ser encontrados através de dimensões de análise.

Abaixo foram incluídos dois exemplos, um da modelagem tradicional e outro da modelagem multidimensional, para mostrar as diferenças existentes entre as duas abordagens.

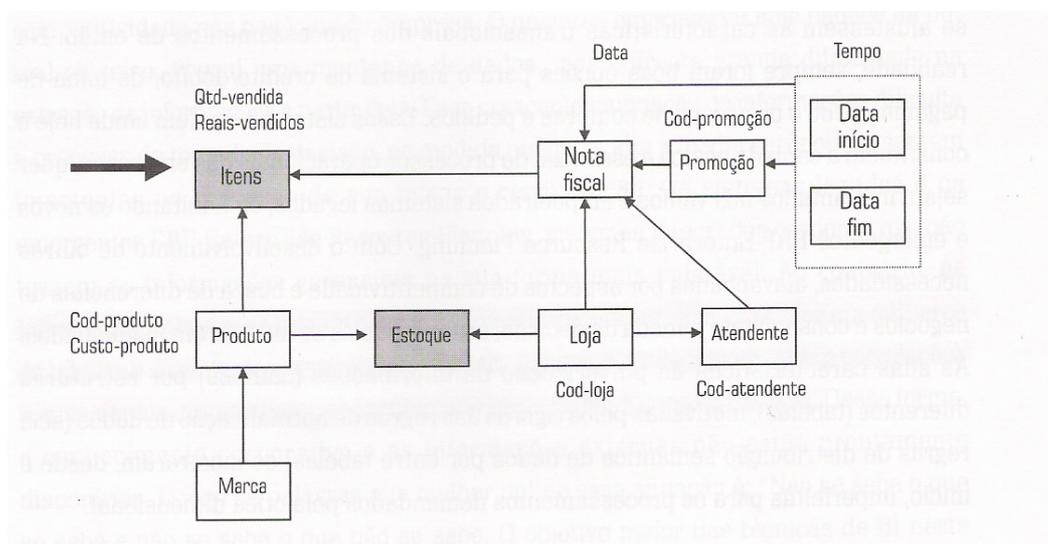


Figura 3 – Exemplo de modelo de Entidade e Relacionamentos (Barbieri, 2001)

O modelo acima representa um caso de sistema de vendas a varejo, cuja principais entidades são: LOJA, que emite várias NOTAS FISCAIS, cada uma com vários ITENS, onde cada um está associado a um PRODUTO, que possui sua quantidade registrada no ESTOQUE.

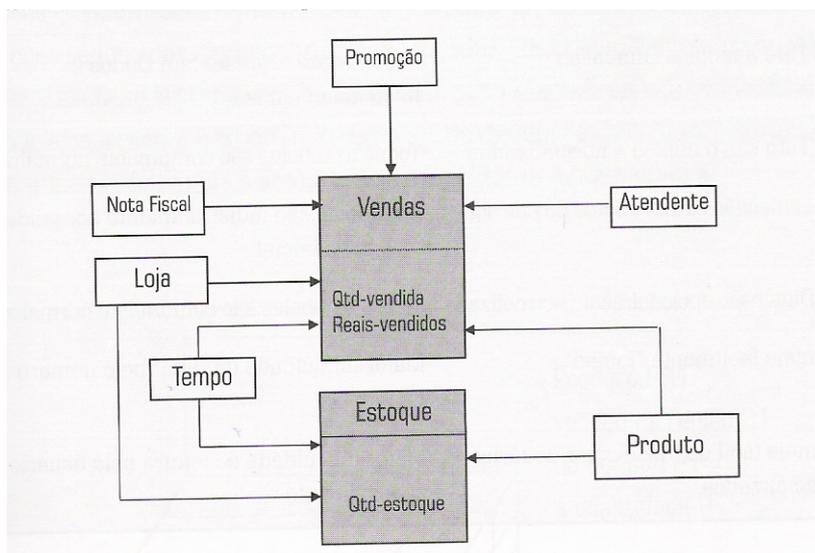


Figura 4 – Exemplo de modelo de Dimensional (Barbieri, 2001)

No modelo dimensional, mostrado na figura 4, existe uma tabela **fato** e seis tabelas de **dimensão** (tempo, produto, loja, promoção atendente e nota-fiscal). A tabela fato representa as vendas de um produto em um determinado momento, feito numa promoção, numa loja, realizada por uma atendente e registrada em uma nota fiscal. Existe também uma outra tabela fato, que surgiu para registrar os aspectos de estoque, e que também se relaciona com algumas dimensões já identificadas com a primeira.

A diferença para o usuário está na forma de acesso. No modelo dimensional os “fatos” (quantidade vendida, quantidade em estoque) já estão relacionados com as dimensões de análise (informações que desejam ser analisadas pelos usuários), facilitando o processo de obtenção das informações. Todos os relacionamentos já foram preparados e o usuário só precisa informar os campos que deseja consultar. A complexidade do modelo irá depender da necessidade de relacionamentos (muitos para muitos) entre as tabelas fato e tabelas dimensão, bem como a resolução de estruturas recursivas nas dimensões.

A arquitetura do data warehouse evoluiu muito até o final dos anos 90. Novos componentes e novas funcionalidades foram aos poucos sendo incorporados, permitindo um melhor aproveitamento da base de dados única (Murayama, 2004). Esta arquitetura, chamada de CIF – *Corporate Information*

Factory (Inmon e Imhoff,2001), possibilita uma visão holística de todo o ambiente de informações envolvendo processos, ferramentas de extração, infraestrutura de armazenamento e aplicações baseadas em data warehouse.

A figura 5, apresentada por Inmon e Imhoff em 2001, mostra as principais camadas que devem estar presentes nesta arquitetura.

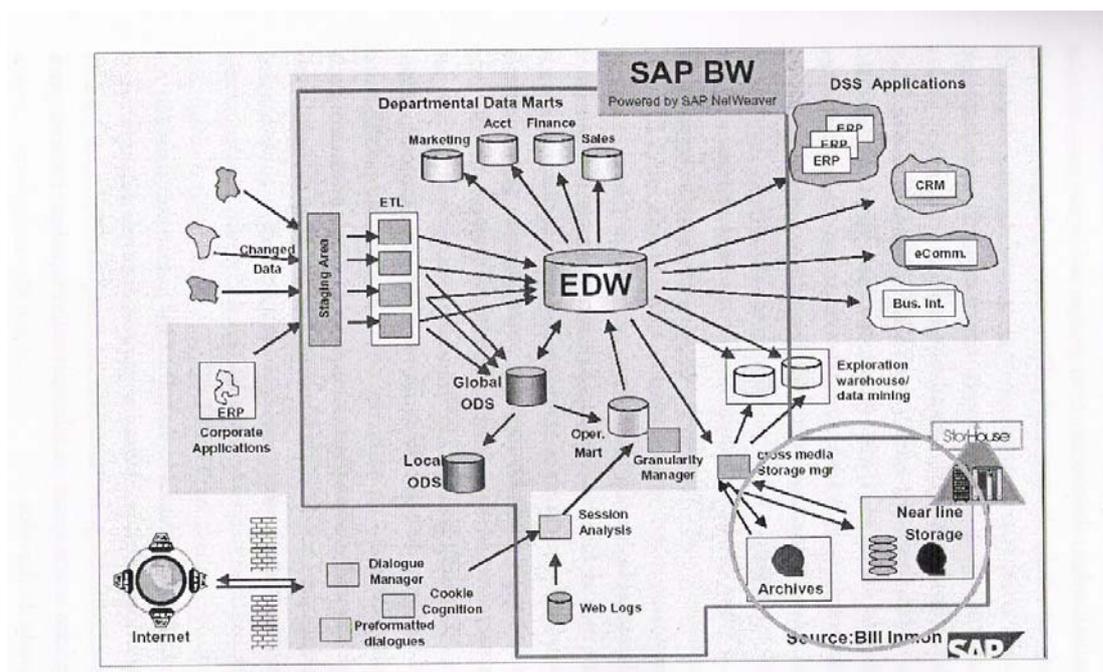


Figura 5 – Arquitetura típica de um data warehouse corporativo (Inmon, 2001)

Para os autores, o coração do CIF é o data warehouse, que é um depósito único de dados da empresa, orientado por assunto, integrado e alimentado pelos sistemas transacionais da empresa, variável no tempo e não volátil, onde podem ser obtidas informações para o conhecimento, inteligência e extensão do aprendizado do negócio. Os dados são organizados de forma única e normalizada, na área de dados operacionais – ODS (*Operational Data Store*), podendo derivar outras visões de dados em diferentes formatos – os *data marts*⁷, que permitem que a empresa disponha de acesso muito flexível ao dados, fatiando, agrupando e

⁷ *Data mart* ou sistemas de gerenciamento de dados multidimensionais, são estruturas que permitem flexibilidade no acesso aos dados, que são armazenados de forma multidimensional e orientadas para um assunto específico do negócio (Inmon, 1997).

explorando dinamicamente o relacionamento entre dados resumidos e detalhados.

A arquitetura CIF é composta por algumas camadas como:

- Camada de extração – sua finalidade é permitir a captura de dados residentes em diversos ambientes, oriundos de diversas fontes heterogêneas. Para tanto, esta camada é composta por mecanismos de manipulação dos dados que permitem a transformação, limpeza, consistência, agregação e etc;
- Camada de Armazenamento – é a camada responsável pelo armazenamento dos dados após sua extração dos demais sistemas. Os dados podem ser armazenados na forma relacional, com dados mais detalhados (ODS – *Operational Data Store*) ou então no formato multidimensional (*data marts* – visões multidimensionais de dados relacionados a um assunto específico);
- Repositório central de dados – sua função é documentar e padronizar o entendimento dos dados técnicos (metadados), permitindo o gerenciamento do ambiente pelos administradores em relação ao conteúdo dos objetos, forma e significado.

Para as empresas que já passaram pela implementação de um ERP, a construção e a manutenção de um data warehouse se torna uma tarefa muito difícil por uma série de fatores (Murayama, 2004):

- O ERP possui centenas de tabelas e suas estruturas e relacionamentos nem sempre são conhecidos;
- O ambiente é mutável, crescendo e se modificando a cada nova versão;
- A extração e a modelagem dos dados pode sofrer impactos em função do tratamento dos dados feitos no ERP;
- A administração dos metadados⁸ é difícil devido aos milhares de objetos existentes no ERP;
- O conhecimento das tabelas e estruturas do ERP fica restrito ao seu fornecedor, dificultando e encarecendo o desenvolvimento de aplicações para disponibilização de informações.

⁸ Metadado é a descrição que é feita a respeito dos dados (Serra, 2002).

Para amenizar os problemas mencionados acima, e também expandir sua área de atuação, as principais empresas fornecedoras de ERP entraram no mercado de data warehouse e passaram a oferecer soluções para esta área. A adoção deste tipo de solução, utilização do data warehouse comercializado pela mesma empresa fornecedora do ERP, traz alguns benefícios como:

- As ferramentas de extração já estão preparadas para as principais estruturas e tabelas do ERP, facilitando esta etapa do processo;
- Já existem disponíveis na solução de data warehouse alguns modelos de negócio pré-configurados para cada uma das áreas (produção, vendas, compras de materiais e serviços, contabilidade, recursos humanos, etc) que facilitam a etapa de modelagem conceitual;
- A integridade das informações é assegurada pela integração existente entre o ambiente ERP e o *data warehouse*;
- O tempo para implementação dos projetos de data warehouse fica reduzido, e aumenta a possibilidade de sucesso da implantação.

No caso da Petrobras, objeto deste estudo, foi adotada a ferramenta de data warehouse da SAP, o Business Information Warehouse – BW, considerada por Inmon e Imhoff (2001) como aderente a arquitetura CIF, por possuir todas as vantagens mencionadas acima com relação ao ambiente ERP – SAP R/3, implementado na empresa.

2.4

SAP BW – Business Information Warehouse

Um componente da nova dimensão de produtos da SAP, o Business Information Warehouse (BW) tem como objetivo a disponibilização de um ambiente dedicado exclusivamente ao tratamento de informações. O BW é considerada uma solução *end-to-end* (da extração dos dados à entrega da informação) de data warehouse com estruturas otimizadas para relatório e análise das informações. Para auxiliar o processo de disponibilização das informações, o BW possui em sua versão padrão uma série de *templates* para cada uma das áreas de negócio da empresa, que abrangem desde a extração, transformação e carga

dos dados até entrega das informações para os usuários, através do uso de ferramentas OLAP nativas, facilitando e agilizando o processo de construção do ambiente analítico.

O SAP BW concentra as informações da empresa em um único local, e possibilita aos seus usuários analisar dados oriundos do SAP R/3 ou de qualquer outra aplicação utilizada pela empresa, incluindo fontes de dados de externas da Internet. O BW utiliza como front-end o *Microsoft Excel* ou a *Web*, onde o usuário pode criar, formatar, e analisar relatórios.

Construído para possuir uma boa performance, o BW é instalado em um servidor dedicado, de modo que os processos operacionais da empresa não comprometam o desempenho das atividades de relatório, e vice-versa.

BW - Visão Geral da Arquitetura

Como mostrado na figura a seguir, existem três camadas principais na arquitetura do Business Warehouse: a camada de extração de dados, a de tratamento e armazenamento das informações, e a camada de disponibilização e acesso às informações para os usuários.

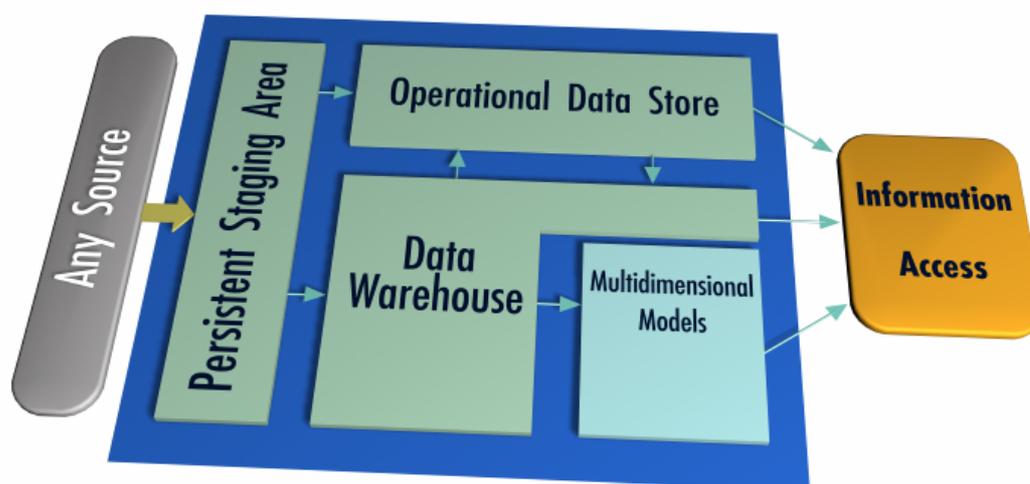


Figura 6 – Camadas e Arquitetura básica do SAP BW (Inmon, 2003)

Na figura a seguir é apresentado um detalhamento maior de cada uma das camadas mencionadas acima.

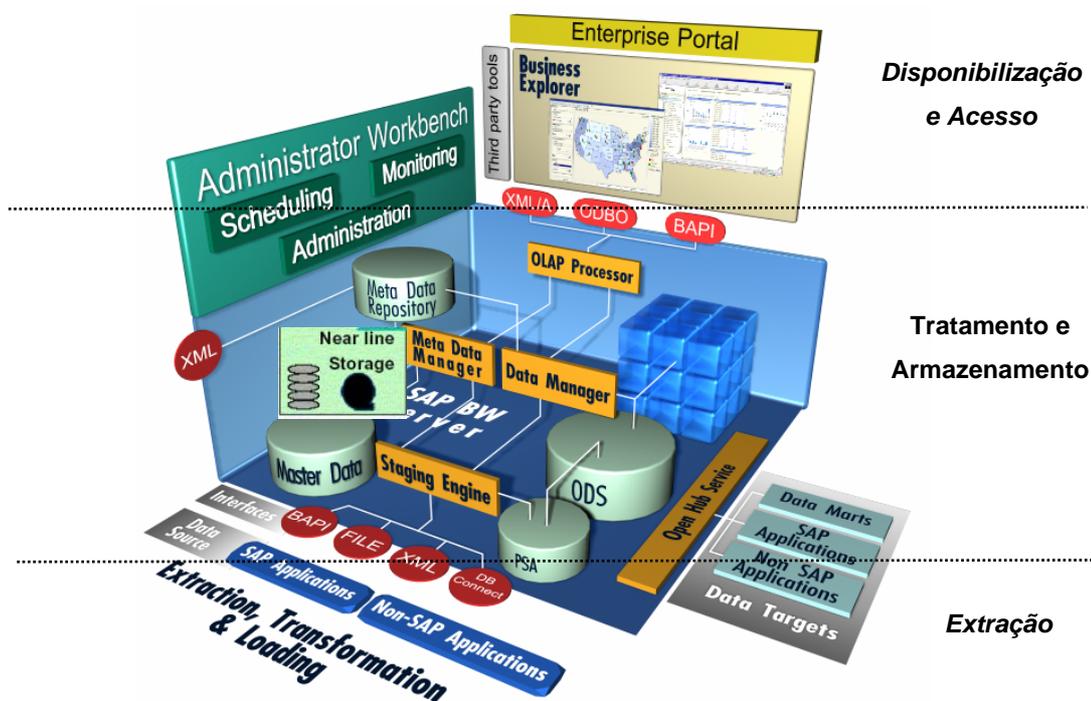


Figura 7 – Arquitetura atual da versão 3.0 do SAP BW (Inmon, 2003)

Camada de Extração

Na camada de extração é onde ocorre a extração dos dados identificados como relevantes para compor o ambiente de data warehouse. Os sistemas origem dos dados são conhecidos como “sistemas fonte” (“source system”) e no BW existem várias formas (nativas da solução ou não) de extração dos dados a partir de cada um deles. O BW suporta vários tipos de sistemas da fonte:

- Sistema SAP R/3
- Sistemas Não-SAP ou Sistemas Externos
- Sistemas de Arquivo (chamados de flat files)
- Outro(s) SAP Business Information Warehouse (BW)

Durante o processo de extração os dados podem ser transformados (calculados ou sumarizados, por exemplo) e armazenados no ambiente BW de diversas formas, mais operacionais ou mais consolidados, dependendo do requisito do negócio a que pretende atender.

As interfaces de programação para extração dos dados contemplam a carga de dados através do uso de arquivos texto, de ferramentas nativas para extração de dados do SAP R/3 (extratores standard e genéricos de dados do R/3), de ferramentas ETL (Extraction and Transformation Load), entre outras que fazem com que o ambiente BW possua um padrão aberto de comunicação com outros sistemas.

Camada de Tratamento e Armazenamento das informações

A camada de aplicação que permite o tratamento, armazenamento e a recuperação dos dados é conhecida como Business Information Warehouse Server. Esta camada abriga os seguintes componentes principais:

- Administrator Workbench
- Meta Data Repository
- Staging Engine
- InfoCube
- Operational Data Store (ODS)
- OLAP Processor

Administrator Workbench

Atua como o sistema de controle do BW e permite que o desenvolvedor execute todas as tarefas de modelagem e de manutenção dos dados em um único ambiente. O *Administrator Workbench* possui os seguintes componentes:

- *Meta Data Maintenance* – Permite que o administrador especifique e mantenha o data warehouse, definindo os modelos multidimensionais, os ODS - *Operational Data Store* (áreas contendo dados mais operacionais), além de outras informações técnicas.
- *Scheduler* - Sua função é agendar a transferência dos dados de um sistema fonte para o BW, em intervalos regulares de tempo.
- *Data Load Monitor* - Supervisiona os processos da carga e preparação dos dados.
- *Data Access Monitor* – Obtém estatísticas de utilização do BW.

Meta Data Repository (Repositório de Metadados)

Os metadados são a “informação da informação”. Este é o conceito central de um data warehouse, e um repositório de metadados ativo documenta e controla o ambiente inteiro do data warehouse. Os metadados são usados para descrever a origem, a história, e os muitos outros aspectos dos dados. O repositório de metadados contém informações relacionadas ao negócio (como exemplo, definições, descrições e regras usadas para os relatórios) e informações mais técnicas (como exemplo, a estrutura e o mapeamento das regras de extração e de tratamento dos dados).

Staging Engine

O Staging Engine é o mecanismo existente no ambiente BW para implementar o mapeamento, a transformação e a carga automatizada dos dados, processo orientado pelos metadados. O Application Link Enabling (ALE), é o *middle-ware* persistente de troca de mensagens entre os componentes, é usado para transportar os dados das plataformas OLTP (sistemas transacionais) ao servidor de BW. Iniciado pelo *Scheduler*, o *Staging Engine* solicita a extração de um sistema transacional. Os dados são então submetidos a uma série de transformações, nas etapas de mapeamento e condensação dos dados, antes que possam ser armazenados no data warehouse, nas áreas mais operacionais (ODS) ou nos modelos multidimensionais (ou “infocubos”).

InfoCubos

Os recipientes centrais os dados que formam a base para os relatórios e análises no BW são os modelos multidimensionais chamados no BW de “infocubos”. Os “infocubos” contêm dois tipos específicos de dados: as características (conhecidos em outras ferramentas como campos) e as *key figures* (campos de quantidade ou valor). Um “infocubo” é um conjunto de tabelas relacionadas e que são organizadas de acordo com a modelagem multidimensional mostrada na seção 2.3.1, (o BW utiliza uma variação desta modelagem - o

Extended Star Schema), contendo no centro do modelo uma tabela fato (fact table), específica para armazenar os “fatos” (exemplo: quantidade vendida de um determinado produto, valor da venda, etc), e diversas tabelas de dimensões que contém características relacionadas com estes fatos (exemplo: cadastro de produto, estoque, cliente, etc). A “tabela do fato” contém os valores e quantidades das transações (sumarizados de acordo com as regras de extração), enquanto as tabelas das dimensões contém as informações de cadastro necessárias e relacionadas à tabela fato do modelo.

Operational Data Store - ODS

O ODS é um repositório de dados dentro do sistema de BW, contendo dados não agregados, organizados em um conjunto de tabelas relacionais. O ODS armazena dados transacionais de diferentes sistemas fonte e pode ser utilizado como um repositório intermediário no processo de carga e tratamento dos dados, antes da carga para os modelos multidimensionais.

Processador OLAP – On-Line Analytical Processing

O processador de OLAP é o motor analítico do BW. Utiliza a base de dados multidimensional para análise e apresentação de vários tipos de dados, permitindo ao usuário vários tipos diferentes de consulta às informações armazenadas (por exemplo, o drilldown aos vários níveis de detalhe).

Camada de Acesso às Informações

Na camada superior da arquitetura do BW, temos o Business Explorer - BEx é o ambiente de relatório (apresentação e análise) para usuários finais. Com sua interface gráfica de fácil utilização, o usuário pode acessar as informações disponíveis no ambiente BW usando o Microsoft Excel (BEx Analyzer) ou então consultas disponibilizadas na Web (intranet da empresa ou então na própria Internet).

O Business Explorer Analyzer é um ambiente interativo baseado no Microsoft Excel onde as consultas (queries) são definidas através da seleção de características e *key figures* (campos texto e campos de valor). Durante a execução, os dados são selecionados no data warehouse de acordo com os filtros especificados pelo usuário e com seu perfil de autorização às informações solicitadas. A utilização do Excel possibilita o uso de recursos adicionais desta ferramenta aos recursos disponíveis do ambiente BW.

O acesso às informações do BW via *web* também é bastante amigável e na sua construção permite a utilização de recursos de ferramentas de aplicação *web* (Html, Java, Asp, etc) que associadas às funcionalidades disponíveis no BW permitem a construção de painéis para acompanhamento de indicadores, além de relatórios de acompanhamento, muito utilizado pelos executivos das empresas.

Desta forma, de acordo com as informações apresentadas nesta seção a ferramenta BW é uma plataforma para a construção de aplicações analíticas de suporte a gestão e a tomada de decisão. Alinhada com a arquitetura CIF proposta por Inmon e Imhoff (2001), o BW veio evoluindo desde sua primeira versão em 1998 e atualmente é considerada a solução da SAP para aplicações de *Business Intelligence*, uma vez que permite o armazenamento de uma grande quantidade de dados, além de viabilizar a exploração e análise destes dados, que se traduzem em informações úteis e necessárias às decisões a serem tomadas na organização.

2.5

Conceituação do Objeto Informação

Apesar de ser um tema bastante debatido, ainda existe muita controvérsia com relação às definições de dado e informação.

Checkland e Holwell (1998) iniciam sua discussão sobre este tema definindo duas palavras:

- o dado - ou *data*, originado da palavra latina *dare* que significa “dar” - está associado ao ponto de partida do processo mental, uma vez que ele representa um fato existente no mundo real;
- *capta* - da palavra latina *capere* que significa “captar” – é um subconjunto de dados que alguém seleciona ou presta algum nível de atenção.

Para os autores, transformar *data* em *capta* é um processo mental natural que é feito quase que o tempo todo pelos seres humanos. Como *capta* é um dado revestido de características cognitivas, pode dar origem à informação desde que seja revestido de contexto e significância, através de um processo individual ou coletivo. Uma informação é entendida então como um fato significativo, contextualizado e dotado de algum tipo de motivação. Segundo Checkland e Holwell (1998) os atuais sistemas eletrônicos de informação provêm *capta* para os seus usuários, já que a verdadeira informação é criada na mente das pessoas após sua devida contextualização e atribuição de significado.

Para Urdaneta (1992), existem quatro classes diferentes de informação que são: dados, informação, conhecimento e inteligência. Estas quatro classes são diferentes e possuem um valor diferente no processo decisório. Os dados são considerados a matéria-prima a ser utilizada na produção da informação. Através de algum tipo de processamento são transformados (formatação, fusão, tradução, etc) para serem exibidos de forma inteligível para aqueles que vão utilizá-la – geração da informação. Em seguida, através da interpretação e integração de vários dados e informações, ocorre a geração do conhecimento. O conhecimento não é estático e modifica-se mediante a interação com o ambiente, sendo este processo denominado de aprendizado. A inteligência resulta então da síntese do conhecimento, sendo usada no julgamento e na intuição daquele que toma decisões. A transformação de conhecimento em inteligência é realizada baseada na experiência e na intuição humana, e não pode ser reduzida a procedimentos ou regras. Desta forma, a inteligência é considerada a base para o processo decisório.

Para McGee e Prusak (1994) informação são dados coletados, organizados, ordenados, aos quais são atribuídos significado e contexto; representa dados em uso, o que implica na existência de um usuário. A transformação de dados em informação requer que a pessoa responsável pelo processo decisório receba-os de tal forma que possa relacioná-los e de que possa atuar sobre eles. Assim sendo, segundo os autores a informação deve ser discutida no contexto de usuários e responsáveis por decisões específicas.

Barreto (1996) conceitua informação como “estruturas significantes com a competência de gerar conhecimento no indivíduo, em seu grupo, o na sociedade”, caracterizando-a como instrumento modificador da consciência do homem e de

seu grupo social. Moresi (2000) em estudo sobre o valor de um sistema de informação restringe o conceito de informação definido por Barreto (1996) para o âmbito do indivíduo e das empresas.

Callaos e Callaos (2002) em uma abordagem sistêmica do conceito de informação definem “dado” como o lado objetivo e material, enquanto que a “informação” é o lado subjetivo e mental.

Sá Carvalho (2003), sintetiza as abordagens de Davenport, Checkland e Barreto com a seguinte conceituação de informação organizacional: “Representação simbólica de fatos e/ou idéias potencialmente capaz de alterar o estado de conhecimento de um usuário visando uma ação no âmbito de uma organização”.

Esta definição aborda que o fenômeno da informação ocorre tão logo os dados apresentados a alguém sejam considerados relevantes e dotados de algum significado, além de acrescentar a forma e o objetivo como elementos indissociáveis, além do próprio conteúdo.

2.6

Qualidade ou Valor da Informação

A informação organizacional possui características que podem ou não torná-la um recurso efetivo para os seus consumidores. Na literatura existente, observou-se uma série de abordagens sobre o tema, mas não foi encontrado um consenso sobre o conjunto de características que devem ser consideradas para avaliação da qualidade da informação disponibilizada para os usuários.

O objetivo deste tópico é apresentar então o resultado da pesquisa bibliográfica que destaca diversas características e formas de abordagem do tema, e que podem apoiar nesta avaliação.

Segundo Nehmy e Paim (1998), vários termos podem ser utilizados para a abordagem da avaliação da informação de forma similar à qualidade, mas o mais freqüente deles é o uso do “valor” como equivalente à qualidade. Na tentativa de classificar o valor da informação, Cronin (1990) identificou os seguintes tipos:

- valor de uso – depende da utilização final que terá a informação;
- valor de troca – depende da oferta e da demanda, podendo ser considerado como o valor de mercado;
- valor de propriedade – informações geradas pela empresa e que são guardadas sem a obrigatoriedade da exploração ou uso;
- e valor de restrição – surge para o caso de informações secretas ou de interesse comercial.

Embora não haja uma delimitação do enfoque a ser utilizado para o “valor”, de acordo com Repo (1989), quando os usuários tomam este termo, na verdade eles estão atribuindo uma significação mais próxima ao “valor de uso”, deixando o “valor de troca” para os economistas.

Para um sistema de informação, a percepção de valor da informação pode ser traduzida de forma metafórica como uma equação na qual o valor é variável e dependente de outros quatro fatores : qualidade, o portfólio de produtos e serviços oferecidos, o custo e o tempo de resposta (Moresi, 2000). Para que estes fatores possam ser qualificados, com enfoque particular na essência do sistema, que é a própria informação, é necessário encontrar atributos que permitam quantificar estes fatores.

Complementando, Moresi (2000) ainda menciona que o valor da informação para a organização depende de sua finalidade, classificação, de seu tipo e de seu valor econômico, e que o sucesso da organização pode estar diretamente ligado à gestão e aproveitamento da “informação”, que atualmente é considerado um recurso importante para a competitividade, equiparando-se aos recursos de produção, materiais e financeiros.

Para Wagner (1990), existe um problema de terminologia, e o valor da informação - valor de uso, valor agregado e valor de troca da informação, é que são os conceitos preferidos, ao invés da “qualidade”. Em suas palavras:

Na era da informação é uma profunda ironia a falta de um corpo sólido de trabalho teórico sobre qualidade e valor da informação. Essa área de conhecimento carece de síntese ou mesmo compêndio que reúna os estudos teóricos.

Chiavenatto (1999) identifica nas organizações três níveis organizacionais distintos: o operacional, o intermediário e o institucional. Desta forma a informação deve atender às necessidades diferenciadas dos diversos níveis da organização, independente de sua natureza ou tamanho:

- Nível Operacional – A informação é dirigida para a execução de tarefas do dia-a-dia operacional;
- Nível Intermediário ou gerencial – A informação é utilizada para monitorar e avaliar os processos, o planejamento e a tomada de decisão do nível gerencial;
- Nível Institucional (nível mais alto da empresa) – Neste nível, onde as decisões são tomadas e são estabelecidos os objetivos da organização, as informações são utilizadas com a finalidade de monitorar e avaliar o desempenho da organização, o planejamento e as decisões de alto nível.

É preciso reconhecer a dificuldade de se disponibilizar a informação correta, para as pessoas corretas, no momento exato da tomada de decisão. Poucas decisões são tomadas com a informação perfeita, portanto, podemos dizer que o valor da informação é uma função do efeito que ela tem sobre o processo decisório (Wetherbe,1987). Segundo Olaisen (1990), apenas quando o usuário utiliza uma fonte de informação é que se torna capaz de julgar se suas expectativas foram atendidas ou não. Mais uma vez cabe ressaltar que o valor da informação também depende do contexto da organização e que uma mesma informação pode ter valor para uma organização e ser considerada sem interesse para a outra. É importante, nesta análise de valor, olhar para o cliente, para o nível organizacional a que pertence, para a finalidade de utilização da informação e para os resultados esperados.

Davis e Olson (1985 apud Amaral, 1994) também ressaltam a importância da destinação e da oportunidade da informação, necessárias para fazer chegar informações de qualidade para as pessoas certas da organização, no tempo certo. Para os autores este ponto é importante, caso contrário a informação é considerada desnecessária ou sem utilidade real para seu usuário.

Segundo Levy (1993 apud Gewandszjder, 2005) a informação faz parte de processo de redução de incerteza visando a melhoria da tomada de decisão, e que existem dois tipos de informação: a informação digital, constituída por fatos e números, e a informação analógica, constituída por uma opinião não fundamentada, por pareceres pessoais, preconceitos, suposições e subjetividade.

De acordo com Nehmy e Paim (1998), os termos qualidade e valor são considerados como equivalentes e não há elaboração teórica suficiente para diferenciá-los da noção do senso comum. Segundo Schwuchow (1990) e Saracevic (1992), um critério para a avaliação da informação é a eficácia, e que esta depende da adequação da informação às necessidades do usuário. Saracevic aborda também outros critérios como a relevância, a utilidade, seletividade, precisão e poder de síntese, para avaliação do valor da informação.

Na visão de Menou (1993), o valor essencial da informação refere-se a um acréscimo de conhecimento em relação ao estágio anterior, considerado o lado positivo da informação. Por outro lado, segundo Nehmy e Paim (1998), pouco tem sido falado com relação ao lado negativo da informação – o erro, a desinformação, ... – e os discursos sobre qualidade e suas noções correlacionadas acabam impedindo que se trate deste lado negativo da informação. Para Nehmy e Paim (1998) o usuário é o contraponto necessário no julgamento de qualquer valor da informação.

A informação possui duas finalidades básicas: a obtenção do conhecimento do ambiente externo e interno da organização, e permitir sua atuação nestes dois ambientes (Chaumier, 1986). Segundo Moresi (2000) de acordo com sua relevância, a informação pode ser classificada em:

- Informações “críticas” - são assim definidas por serem consideradas fundamentais para a sobrevivência da organização;
- Informações “mínimas” - são aquelas informações necessárias à gestão;
- Informações “potenciais” - são as que trazem vantagem competitiva com relação as demais organizações do ambiente;

- Informações “sem interesse” ou “lixo” - são aquelas onde não se deve envidar esforços organizacionais, evitando desperdícios de tempo e de dinheiro na obtenção de informações que não trazem benefício para a organização.

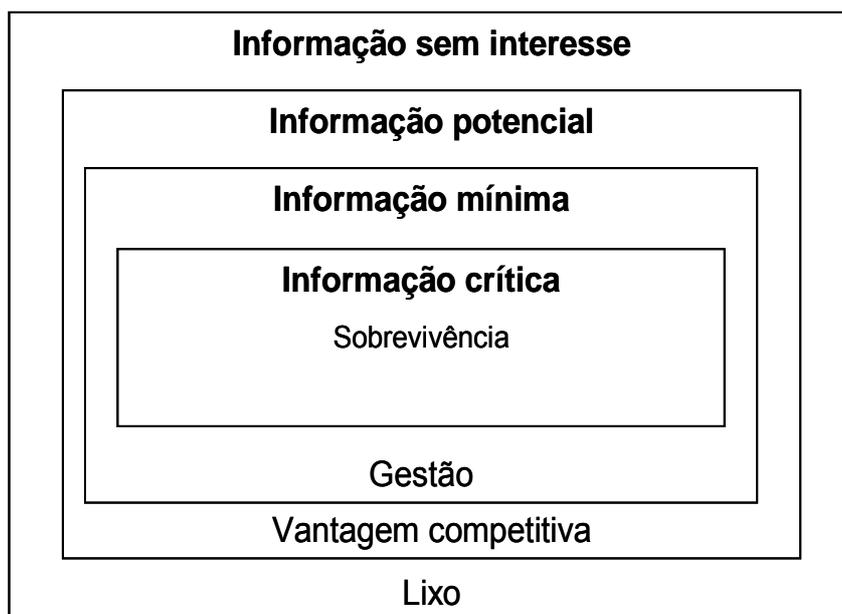


Figura 8 – Classificação da Informação segundo sua finalidade (adaptação de Amaral, 1994 apud Moresi, 2000)

De acordo com esta classificação é possível identificar o grau de relevância e prioridade que a informação exerce em cada nível da organização⁹. Este conceito foi desenvolvido por Saracevic (1999) que define a relevância da informação como função de uma relação que assume significados específicos em aplicações e contextos distintos. Segundo o autor existem cinco tipos de relevância:

⁹ O conceito de relevância para a Ciência da Informação está relacionado ao atributo ou critério que reflete a efetividade de troca de informações entre pessoas e sistemas de recuperação de informação (Gewandsznajder, 2005).

- Relevância do sistema ou algoritmo – relação entre uma pesquisa e objetos de informação de um sistema, através de um procedimento ou um algoritmo;
- Relevância do tópico ou do assunto – relação entre o assunto ou tópico expresso em uma pesquisa;
- Relevância cognitiva ou pertinência – relação entre o estado de conhecimento e a necessidade cognitiva de informação do usuário e dos textos recuperados;
- Relevância situacional ou utilidade – Relação entre a situação, tarefa ou problema e os textos recuperados;
- Relevância motivacional ou afetiva – relação entre intenções, metas e motivações de um usuário e os textos selecionados.

No âmbito desta pesquisa destacam-se a relevância cognitiva, relacionada a interpretação e qualidade da informação, e a relevância situacional, voltada para a utilidade da informação na tomada de decisão, na redução de incerteza e na adequação de informação para a resolução de problemas.

Conceituando a Qualidade da Informação

Wormell (1990) observa que

as definições sobre a qualidade da informação têm sido feitas sob o ponto de vista de definições específicas e subjetivas, seguidas por definições ad hoc. Isso tem resultado em inúmeras interpretações pouco claras do conceito, imperfeitas e de alguma forma caóticas.

Na tentativa elucidar o conceito da qualidade da informação, Nehmy e Paim (1998) fazem um resumo das cinco tendências de abordagem do conceito de qualidade da informação proposto por Marchand (1990):

- **Abordagem “transcendente”** – é aquela que tende a perceber o valor da informação como absoluta e universalmente reconhecido.

Exemplo - Cooney (1991) : “Pode-se considerar a informação do ponto de vista do valor intrínseco que ela possui. Um poema ou uma fórmula matemática, por exemplo, podem possuir qualidades tais como elegância,

introspecção, expressão precisa..., que lhes conferem um valor indiscutível, pelo menos no sentido metafísico. Entretanto, o valor econômico de tal informação, ou seja, o que o mercado se dispõe a pagar por ela pode aproximar-se de zero”.

- **Abordagem “baseada no usuário”** – Segundo Eaton e Bawden (1991) e Moresi (2000), a informação, por ser abstrata e intangível, não possui um valor intrínseco pois o valor está na dependência do contexto e da sua utilização por usuários particulares, em ocasiões particulares.

Para Connel (1981), a informação não varia em valor intrínseco, e seu valor é inteiramente subjetivo, estando na mente do usuário.

Para Casanova (1990), a informação não é uma opinião, e possui características intrínsecas, como responsabilidade, confiabilidade, objetividade, abrangência, precisão, capacidade de ser transmitida, suporte material.

- **Abordagem “baseada no produto”** – é compreendida como a abordagem que tende a ver a qualidade da informação em termos precisos e identificáveis, sendo passíveis de serem mensurados e quantificados; para Nehmy e Paim (1998) esta visão é muito subjetiva, permitindo pouca possibilidade de operacionalização.

- **Abordagem “baseada na produção”** – tende a ver a qualidade como adequação a padrões estabelecidos de necessidade de informação do consumidor.

Huang e Wang (1999) definem qualidade da informação como aquela que está adequada para o uso dos consumidores da informação.

Para Kahn e Stong (1998) qualidade da informação é a característica da informação de alcançar ou superar as expectativas dos usuários finais.

- **Abordagem “baseada em atributos de valor”** – considera a qualidade como um dos aspectos da definição de “valor agregado”. O valor é pensado como a categoria mais abrangente e a qualidade como um de seus atributos. Esta visão reforça a percepção de ambigüidade do uso nos termos na literatura.

Taylor (1985) considera a qualidade como uma das dimensões do “valor agregado” da informação; isso significa que a informação se torna mais valiosa quando é organizada, sintetizada e julgada. Ele considera qualidade as

seguintes características : precisão, abrangência, atualidade, confiabilidade, validade dos dados e informações de um sistema. As outras dimensões, paralelas a qualidade são: acesso, facilidade de uso, adaptabilidade, economia de tempo e de custo.

Em função das limitações de cada uma das abordagens apresentadas acima, Marchand (1990) propôs uma tipologia para definir a qualidade da informação, onde agrega o conceito em oito dimensões inter-relacionadas. Acerca de cada uma delas, ele fez alguns comentários, conforme apresentado a seguir:

- Valor Atual – se refere à percepção do valor da informação e depende do estilo individual de tomada de decisão;
- Exatidão e Completeza – são características da informação, e que podem possuir pesos diferentes, dependendo de sua utilidade no contexto da tomada de decisão;
- Confiabilidade – se refere a uma atitude de confiança com relação às fontes de informação, ou seja, que o uso da informação não levará a uma decisão errada;
- Significado ao longo do tempo - capacidade da informação de ser útil ao longo do tempo; a informação possui um ciclo de vida, o que significa que a partir de um determinado momento a informação pode não ser mais útil para o contexto da tomada de decisão;
- Relevância – está relacionada ao significado ou ao grau de conformidade da informação aos critérios estabelecidos pelo usuário;
- Validade – aborda a variação da percepção da validade da informação, que depende de quem a fornece e de como é apresentada;
- Estética – é um critério muito subjetivo e que está relacionado a forma de apresentação e entrega da informação;
- Valor percebido – representa o valor que o usuário percebe que a informação tem. A reputação da origem da informação muitas vezes legitima este valor, mesmo quando existe alguma deficiência em alguma outra dimensão.

Diversos estudos sobre qualidade da informação chegaram a um consenso de que a sua avaliação deve ser feita por um constructo multidimensional. No entanto, cada autor define e agrupa de forma diferente os atributos e dimensões que irão especificar a métrica da qualidade da informação. Wang e Strong (1996) propuseram uma estrutura de medição que consiste de quatro categorias de atributos, como apresentado na tabela a seguir.

Categorias de Qualidade	Dimensões de Qualidade
Qualidade Intrínseca	Precisão, Objetividade, Confiabilidade, Reputação
Qualidade de Acessibilidade	Acessibilidade, Segurança
Qualidade de Contexto	Relevância, Valor Agregado, Atualidade, Completeza, Quantidade de Dados
Qualidade de Apresentação	Interpretabilidade, Compreensibilidade, Consistência e Consistência

Tabela 4 – Modelo de Medição da Qualidade da Informação (Wang e Strong, 1996)

De acordo com Wang (1993) diferentes usuários podem desejar padrões de qualidade distintos e por isso diferentes aplicações podem requerer dimensões diferentes para a análise. Desta forma, a noção de qualidade depende do uso dos dados, ou seja, o que é considerado adequado para um caso pode não ser para algum outro.

Königer e Reithmayer (1998) propuseram um outro modelo de avaliação da qualidade, ampliando as categorias de avaliação do modelo de Wang. Neste modelo as categorias e dimensões do modelo de Wang são chamadas de dimensões e critérios, como apresentado na tabela 5 (as descrições dos critérios foram adaptadas pela autora desta pesquisa para facilitar o entendimento).

Dimensão	Critério
Qualidade Intrínseca	Precisão – se refere ao rigor / exatidão da informação disponibilizada para o usuário
	Objetividade – as informações disponíveis são objetivas e livres de ambigüidade ou redundância
	Confiabilidade – a informação é confiável e pode ser utilizada sem risco para a tomada de decisão
Qualidade de Acesso	Acessibilidade – a informação pode ser acessada e utilizada com facilidade
	Confidencialidade – se refere à segurança das informações disponibilizadas
Qualidade de Contexto	Relevância – se refere ao grau de conformidade e adequação da informação aos critérios estabelecidos pelo usuário
	Valor Agregado – se refere à percepção do valor da informação para o usuário
	Atualidade das informações - as informações disponibilizadas estão sempre atualizadas
	Completeza – se refere ao grau de abrangência da informação, que é necessária para a tarefa a ser desempenhada
Qualidade de Apresentação	Compreensibilidade – a informação disponibilizada é clara, útil e de fácil compreensão
	Qualidade de Apresentação – a informação é apresentada de forma compreensível e que facilita sua interpretação
	Concisão e consistência – a informação está íntegra, consistente e livre de erros
Qualidade da Meta Informação	Existência – a documentação sobre informação existe e é apropriada para a necessidade do usuário
	Disponibilidade - a documentação da informação está disponível no momento exato em que se precisa dela
Qualidade de Estruturação	Existência – a documentação sobre informação existe e é apropriada para a necessidade do usuário
	Disponibilidade - a documentação da informação está disponível no momento exato em que se precisa dela
	Compreensibilidade – a informação disponibilizada é clara, útil e de fácil compreensão

Tabela 5 – Modelo de Medição da Qualidade da Informação (adaptado de Königer e Reithmayer, 1998)

Olaisen (1990) também propõe um modelo, onde aspectos usuais da qualidade são grupados em quatro categorias:

- Qualidade Cognitiva – depende de como a fonte de informação é valorizada pelo usuário. Nesta categoria estão incluídos: credibilidade, relevância, confiança, validade, e significado no tempo;
- Qualidade do Desenho da Informação – são os aspectos referentes à forma, flexibilidade e seletividade;
- Fatores referentes ao Produto da Informação – relativos ao valor real e a abrangência;
- Fatores relativos à Qualidade da Transmissão - acessibilidade.

Segundo o autor, a qualidade deve ser avaliada por várias dimensões, mas esta avaliação só é possível no momento do uso, ou seja, no momento em que usuário necessita da informação é que pode verificar se sua expectativa foi atendida ou não.

Existem várias formas de abordagem nas quais os atributos de qualidade são derivados e agrupados. A abordagem intuitiva ou empírica é a mais usada delas desde o início das pesquisas sobre qualidade da informação (Liu e Chi, 2002), no entanto, isso gera uma lacuna teórica para justificar o agrupamento dos constructores de uma maneira ou de outra. Como exemplo, Wang e Strong (1996) sugeriram categorias a partir de aspectos como armazenamento de dados (acessibilidade e apresentação) e aplicação (qualidade do contexto). Um ano mais tarde, em 1997, Strong sugeriu quatro categorias diferentes de agrupamento: intrínseca, acessibilidade, contexto e apresentação. Como não existe nenhuma razão teórica para justificar a classificação, não fica claro o motivo do agrupamento em quatro categorias, e não três ou cinco (Liu e Chi, 2002).

Em uma revisão sobre *frameworks* criados nos últimos dez anos para avaliar a qualidade da informação, Eppler e Wittig (2000), depois de escolherem os sete mais representativos e oriundos de autores de locais geograficamente diferentes no mundo, identificaram alguns elementos comuns a cada um deles:

Frameworks avaliados: Lesca & Lesca (1995); Redman (1996); Wang e Strong (1996); Russ_Mohl (1998); Königer e Reithmeyer (1998); English (1999); e Alexander e Tate (1999).

Avaliação:

- Dimensão Tempo - Seis dos sete *frameworks* avaliados consideraram a dimensão tempo. Alguns avaliaram a rapidez no processo de entrega da informação, e outros consideraram o aspecto da atualidade da informação, onde a rapidez na entrega é necessária mas não suficiente;
- Acessibilidade - Cinco consideraram importante a acessibilidade como critério para definição da qualidade;
- Objetividade - Quatro utilizaram o critério objetividade, que é considerado um dos critérios mais difíceis de em termos de definição;
- Relevância – Cinco deles consideraram a relevância como critério e a definiram como a pertinência da informação para o usuário final;
- Precisão – foi identificada como critério em apenas quatro dos *frameworks*, no entanto, os demais utilizaram termos correlatos para esta classificação;
- Consistência – apareceu explicitamente em apenas três *frameworks* avaliados, e se refere ao sistemático, não ambíguo formato e conteúdo da informação;
- Completeza – foi considerado critério crucial em quatro dos sete avaliados e é visto conjunto de informações que atende completamente aos requisitos do usuário.

De acordo com sua análise, Eppler e Wittig identificaram que dentre todos os frameworks avaliados, o definido por Wang e Strong (1996) é o que oferece o maior equilíbrio entre consistência teórica e aplicabilidade prática, podendo ser utilizado em qualquer tipo de aplicação. O modelo de Königer e Reithmeyer (1998) também foi considerado adequado para qualquer contexto de aplicação. Na presente pesquisa, será considerado o modelo de Königer e Reithmeyer (1998) para avaliar a qualidade da informação do ambiente analítico SAP BW da Petrobras.

Cadeia de Valor de um Sistema de Informação

Depois de definidas algumas formas de classificação da qualidade da informação, é preciso definir o “valor” desta qualidade em um sistema de informação da organização. Os sistemas de informação têm sido criados para otimizar o fluxo de informações relevantes no âmbito de uma organização. Para agregar valor para a organização, um sistema de informação deverá ter qualidade, gerando o benefício desejado (serviço), a um custo adequado e no tempo certo. Seu portfólio compreenderá todos os produtos tangíveis (documentos, relatórios, etc.) e intangíveis (indexação, recuperação, etc.) relacionados à matéria-prima informação.

Para Serra (2002) a qualidade da informação pressupõe a qualidade do dado, do sistema de informação e do ambiente computacional e para o autor um dos grandes desafios é determinar pontos de controle para avaliar a qualidade do sistema depois de sua construção.

Uma má qualidade da informação pode trazer impactos severos em todo o desempenho organizacional (Matsumura e Shouraboura, 2002). Desta forma, este ponto tem recebido bastante atenção das empresas nos últimos anos. Inúmeros artigos tem sido escritos relatando estes aspectos da qualidade da informação, e uma nota do The Wall Street Journal (1992) sobre uma pesquisa feita pelo MIT relata que *“grandes bases de dados podem ser mineradas para prospectar vendas pelos clientes, analisar custos, simular tendências futuras... no entanto, existe um problema grave: estas grandes base de dados podem estar cheias de lixo”*.

Atualmente, a tecnologia da informação e os sistemas de informação tem afetado os processos de trabalho das organizações, e a qualidade da informação disponível se tornou critica para o posicionamento das empresas em mercados muito competitivos. Para tomar uma decisão os executivos precisam confiar nas informações disponibilizadas e é imperativo que os problemas com a má qualidade das informações nos sistemas de informação da empresa sejam endereçados e resolvidos o quanto antes.

Uma das tarefas críticas de qualquer sistema de informação é a disponibilização da informação correta, às pessoas certas, e cada tomador de

decisão necessita apenas de uma parte das informações disponibilizadas para apoiá-lo neste processo.

Segundo Serra (2002) é preciso distinguir dois tipos de informação quanto à sua finalidade: a informação operacional e a informação gerencial. A primeira destina-se a permitir que as operações da empresa continuem a ocorrer, enquanto que a segunda destina-se a “alimentar” os processos de tomada de decisão. Para o caso das informações classificadas como gerenciais, o autor menciona que para cada um dos níveis gerenciais existe uma necessidade diferente e desta forma uma organização deve conhecer suas necessidades de informação gerenciais em todos os níveis. Somente então um sistema de informação poderá atender adequadamente estas necessidades.

Para Avison & Taylor (1997) o planejamento e o desenvolvimento de um sistema de informação de uma organização deve ser analisado sob os seguintes aspectos: estratégico (relacionados à competitividade, posicionamento e diferencial de mercado); organizacional (necessidade de informação para a coordenação operacional e atendimento às demandas externas); econômicos (redução de custos, melhoria de produtividade e ganhos financeiros); e capacitação da organização na utilização da tecnologia disponibilizada pelo sistema de informação. Desta forma, o sistema de informação poderá atender aos objetivos previstos, proporcionando melhoria e otimização do processo decisório.

Segundo Moresi (2000), cada sistema de informação possui uma cadeia de valor associada, ou seja, um conjunto de atividades que são executadas com o objetivo de produzir e transferir informação, sustentando o processo decisório.

Os sistemas de informação têm sido desenvolvidos para otimizar o fluxo de informações dentro de uma organização, apoiando o processo decisório da organização. Desta forma, o fluxo da informação em uma organização é um processo de agregação de valor, que cresce com um grau de proporcionalidade direta em relação à pirâmide organizacional. O nível mais alto necessitará de informações de alto valor agregado, visando obter uma visão global da situação, enquanto os níveis mais baixos precisarão de informações de menor valor agregado, para atender às tarefas rotineiras, do dia-a-dia (Moresi, 2000).

Cabe mencionar ainda alguns fatores importantes na percepção de valor da informação e do sistema de informação : confiabilidade e precisão das

informações; “facilidade” na transferência da informação entre emissor e receptor através do uso de um sistema de informação; apoio ao processo decisório; informação vista como instrumento de vantagem competitiva; tempo de resposta do sistema; relação custo-benefício; e relação custo-efetividade.

Para a empresa objeto deste estudo, é de fundamental importância que as informações estejam integras e confiáveis, suportando o processo de gestão neste momento de expansão para o mercado internacional. Desta forma, espera com a implantação do SAP R/3 e do SAP BW, garantir a qualidade das informações desde o processo de geração dos dados até a sua disponibilização no ambiente de data warehouse, apoiando o processo decisório da empresa em todos os níveis organizacionais.