

---

## PROCESSOS DE GERAÇÃO DE INFORMAÇÕES DE QUALIDADE PARA O APOIO A DECISÃO EXECUTIVA

Décio Fonseca

**UFPE**

Antonio de Vasconcellos Carneiro Campello

**UFPE**

Jairo Simião Dornelas

**UFPE**

**SUMÁRIO:** 1. INTRODUÇÃO; 2. REFERENCIAL TEÓRICO; 2.1 TÉCNICA DE ARMAZENAMENTO; 2.2 QUALIDADE DOS DADOS; 2.3. EXPLORAÇÃO DE DADOS E INFORMAÇÕES; 3. METODOLOGIA DE PESQUISA; 4. PROCESSOS DE GERAÇÃO DE INFORMAÇÕES DE QUALIDADE PARA A TOMADA DE DECISÃO; 4.1 MODELO FUNCIONAL; 4.2 OS INSTRUMENTOS DE SUPORTE AO MODELO SÃO; 4.3 AS TÉCNICAS QUE PODEM SER UTILIZADAS NO SUPORTE AO MODELO SÃO; 4.4 OS PROCEDIMENTOS UTILIZADAS NO MODELO SÃO; 4.5 ELEMENTOS ENVOLVIDOS; 4.6 A BASE DE DECISÕES E MODELOS; 4.7. PROCEDIMENTOS DE CONTROLE DE QUALIDADE DOS DADOS; 5. CONSIDERAÇÕES FINAIS;

**RESUMO**

As decisões executivas constituem-se no principal desafio dos gerentes nas organizações modernas. O custo de uma decisão estratégica errada é incalculável, podendo até representar a falência de grandes organizações. As decisões precisam ser tomadas em tempos cada vez menores pela própria dinâmica do mundo atual. Para obter uma boa decisão o agente decisor necessita de informações de qualidade, oriundas de diferentes sistemas de informação, de diferentes fontes de informação, internas e/ou externas à organização. O que se propõe neste trabalho, construído a partir de um ensaio teórico, é definir processos para a geração de informações na tomada de decisão executiva, baseados em critérios de qualidade de dados, de forma a permitir um aumento da probabilidade de sucesso nas decisões.

**Palavras-chave:** Apoio à Decisão Executiva. Qualidade de Dados. Data Warehouse. Modelo. Processo.

**ABSTRACT**

The executive decisions consist in the main challenge of the controlling in the modern organizations. The cost of a missed strategical decision is incalculable, being able until representing the bankruptcy of great organizations. The decisions need to be taken in lesser times each time for the proper dynamics of the current world. To get a good decision the decisor agent needs information of quality, deriving of different systems of information, different sources of information, external interns and/or to the organization. What it is considered in this work, constructed from a theoretical assay, is to define processes for the generation of information in the taking of executive decision, based in criteria of quality of data, of form to allow an increase of the probability of success in the decisions.

**Key Words:** Executive Decision Support, Data Quality, Data Warehouse, Model, Processes.

## 1. INTRODUÇÃO

As decisões executivas num mundo globalizado constituem-se no principal desafio dos gerentes nas organizações modernas. O custo de uma decisão estratégica errada é incalculável, podendo até representar a falência de grandes organizações.

Observa-se na prática que os bancos de dados apresentam problemas de qualidade de dados armazenando registros com dados imprecisos, incompletos e ambíguos e o fato de existir uma grande necessidade de se obter dados de fontes externas para decisões estratégicas, torna-se ainda mais premente a necessidade de se considerar o uso de modelos de qualidades de dados para se assegurar as condições básicas para o processo de tomada de decisão executiva.

Para obter uma boa decisão o agente decisor necessita de informações de qualidade, oriundas de diferentes sistemas de informação, de diferentes fontes de informação, internas e/ou externas à organização (OLSON, 2003), armazenadas em arquivos convencionais, em bancos de dados de diferentes *SGBD* ou obtidas através de *sites* na *WEB*. O que se percebe na prática é que a maioria das organizações não se adotam práticas na geração e gestão dos dados considerando sistematicamente a qualidade dos dados.

Muitos sistemas legados permanecem em uso nas organizações porque são essenciais ao funcionamento das empresas, gerando muitas vezes, dados incompletos e de baixa qualidade para a tomada de decisão executiva. Os novos sistemas com a evolução das aplicações que obtém dados externos, têm enfrentado um grande desafio de tratar um grande volume e crescente de dados e informações. Neste contexto o processo de geração e gestão de dados com qualidade de dados assegurada, torna-se ainda mais premente a fim de se assegurar processos de decisão mais confiáveis.

A geração de informações de qualidade para a tomada de decisão utiliza também dados oriundos de diferentes sistemas legados, que são analisados, enriquecidos e integrados, gerando a base de dados de suporte à decisão. A partir desta base de dados são executados procedimentos de exploração de dados, como técnicas *OLAP*, que fazem parte das tradicionais ferramentas de suporte à decisão, como também através de técnicas de mineração de dados, que buscam descobrir padrões existentes na base de dados, possibilitando ao decisor um ganho adicional de conhecimento sobre a decisão.

O que se propõe neste trabalho é a definição da sistematização de um conjunto de processos de qualidade de dados, na tomada de decisão executiva, de forma a permitir um aumento da probabilidade de sucesso nas decisões.

## 2. REFERENCIAL TEÓRICO

Consideramos como base de referência teórica para a tomada de decisão executiva aspectos referentes ao armazenamento de informações, a exploração de dados e informações, e a qualidade de dados.

### 2.1 TÉCNICA DE ARMAZENAMENTO

Com relação ao armazenamento de informações, escolhemos a tecnologia de data warehouse por ser atualmente a mais apropriada e usada nas organizações. Um *data warehouse (DW)* pode ser definido como um banco de dados integrado e especializado, derivado de diversos bancos de dados corporativos e de diferentes fontes de informação, internas ou externas à organização, tendo como objetivo principal apoiar a tomada de decisão executiva.

O *DW* contém dados granulares integrados, oriundos de diferentes fontes de informação. A normalização é o estado ideal para estes dados, por tornar os dados extremamente flexíveis, pois podem ser moldados e remoldados, podendo ser resumidos ou agregados, para a criação de novas formas de dados (KIMBALL, 1998; KIMBALL, 1997; INMON, 2001):

Na construção de um *DW de apoio à decisão* as seguintes atividades são essenciais:

#### a) Identificação dos Requisitos de Decisão

Tem o objetivo é definir precisamente o problema que afeta o decisor na organização (PYLE, 1999). Os decisores definem, selecionam e complementam os requisitos de negócio envolvidos com a decisão, como: objetivos e escopo da decisão; fontes de informação; alternativas de solução; atributos de qualidade para os dados legados.

#### b) Avaliação das Fontes de Informação

Esta atividade prevê o levantamento, avaliação e seleção das diferentes fontes de informação relacionadas. Na avaliação é analisada a qualidade dos dados disponíveis, com base nos atributos de qualidade definidos (ENGLISH, 1999; PYLE, 1999).

### c) Integração de Dados

Visa integrar os dados selecionados, com base nos requisitos de decisão estabelecidos. Os dados não integrados, de aplicativos legados, são convertidos para uma estrutura integrada, que é compatível com dados que vêm de outros sistemas legados (INMON, 2001). São estabelecidas estratégias para a coleta dos dados das diferentes fontes e meios de armazenamento, através da definição de procedimentos de coleta e periodicidade de atualização. Uma das técnicas de integração se dá por conversão, como a padronização de unidades de medidas e de domínios:

### d) Pré-Processamento

Esta atividade consiste na identificação dos dados com problemas e na exclusão daqueles que não podem ser corrigidos. Estes problemas ocorrem devido à duplicação de registros; inconsistência e incompatibilidades entre campos; dados incompletos ou não existentes; provocados normalmente por erros gerados na origem dos dados e também pela redundância de informações (ENGLISH, 1999; SANTOS, 2002);

### e) Enriquecimento dos Dados

O enriquecimento dos dados é sempre a parte mais trabalhosa do processo, consistindo da correção dos dados errados. Muitas vezes são encontrados erros que só podem ser corrigidos com a interação com o negócio: códigos impróprios, valores fora de domínio, ausência de valores. Tudo isto causa impacto na qualidade dos resultados.

## 2.2 QUALIDADE DOS DADOS

A qualidade dos dados passou a ser uma exigência atual, principalmente pelo crescente uso das informações no processo de tomada de decisão. Constata-se a importância da qualidade dos dados nos sistemas legados, nos esquemas conceituais projetados e nas *bases de dados*.

Com a utilização cada vez maior das técnicas de suporte à decisão, a qualidade dos dados tem sido um tema que preocupa os especialistas, já que está muito dependente dos dados operacionais da empresa, que normalmente não são de boa qualidade.

### 2.2.1 PROBLEMAS DA BAIXA QUALIDADE DOS DADOS

O impacto da baixa qualidade dos dados nas empresas tem como reflexos principais a insatisfação dos clientes, o aumento dos custos operacionais e a falta de eficiência na tomada de decisão executiva. Em estudo realizado nos Estados Unidos, calcula-se que o custo da baixa qualidade dos dados está entre 8 – 12% do custo operacional das empresas (TAYI, 1998).

Em sistemas legados, os problemas de qualidade de dados mais comumente encontrados são (PSI, 2003): A propriedade da informação, baixo grau de correção, abrangência inadequada, falta de consistência, baixo grau de completeza, obsolescência, dados não documentados, precisão inadequada, necessidade de convergência com outras aplicações e precariedade de acesso e de disponibilidade de dados.

Na construção de uma base de dados para suporte à decisão, estes são os principais problemas que afetam as atividades dos analistas.

### 2.2.2 CRITÉRIOS DE QUALIDADE DE DADOS NA TOMADA DE DECISÃO

A qualidade dos dados na tomada de decisão apresenta características específicas, bem diferente das aplicações nos níveis operacional e gerencial. Entre estas características tem-se:

- Atributos como *precisão e atualização* normalmente não são tão relevantes quanto para aplicações financeiras, tipo contabilidade. Em valores financeiros, por exemplo, as decisões consideram a ordem de grandeza dos dados (milhares de Reais), além de ocorrências históricas ao longo dos anos, sem a necessidade da última atualização dos dados (TAYI, 1998);
- De forma antagônica, atributos como *relevância, credibilidade, entendimento, acessibilidade, manipulação*, normalmente são considerados fundamentais para os agentes decisores;
- Dados, mesmo que estejam corretos, gerados por diferentes fontes de informação, expressos em diferentes unidades ou com nível de granularidade incompatível com os

objetivos da decisão, dificultam a *acessibilidade* e a *manipulação* por parte dos agentes decisores;

- Muitos dados são coletados fora da empresa, o que dificulta a aquisição e gera incerteza sobre a qualidade. Como exemplo tem-se: dados sobre clientes, fornecedores, concorrentes, tecnologia e ambiente;
- São usados dados baseados em projeções e tendências, que por sua própria natureza geram erros de previsão. Como exemplo, tem-se: PIB estimado, crescimento demográfico, mercado consumido e vendas previstas.

A qualidade dos dados na tomada de decisão está diretamente associada à adequação ao seu uso, o que implica que o conceito de qualidade de dados é relativo. Os dados podem ser adequados a uma decisão e não ser adequados à outra (TAYI, 1998), daí a importância da participação do agente decisor na identificação e qualificação dos dados a serem manuseados.

A baixa qualidade dos dados tem outro agravante, já que a falta de “dados adequados ao uso” retarda as decisões e em consequência a implantação das alternativas de ação, implicando em custos adicionais e perda de oportunidades para a empresa (TAYI, 1998).

### 2.2.3 CONTROLE DA QUALIDADE DE DADOS NA TOMADA DE DECISÃO

Normalmente, as atividades de controle de qualidade de um produto ocorrem durante o processo produtivo do mesmo, entretanto, a grande maioria dos dados utilizados na tomada de decisão executiva é gerada por diferentes fontes de informação, internas ou externas a organização, não sendo possível a efetividade deste controle. As ações, neste caso, não são preventivas e sim corretivas, em caso de falha detectada. A atividade consiste em filtrar os dados das diferentes fontes, detectando e corrigindo as falhas, quando for possível e viável.

Entre os controles que podem ser realizados sobre os dados das diferentes fontes de informação, tem-se (PIERCE, 2002): seleção dos dados com base nos critérios de seleção e periodicidade a serem estabelecidos pelos agentes decisores, avaliação do nível de granularidade dos dados, em função da necessidade da decisão, teste de duplicidade de dados (registros), teste de consistência, envolvendo formato (numérico, alfanumérico), tamanho (número de caracteres), conteúdo (valores limites, tabela de valores), avaliação do nível de atualização dos dados, com base na data da última atualização.

### 2.2.4 CICLO TDQM (TOTAL DATA QUALITY MANAGEMENT)

O ciclo de administração da qualidade dos dados é baseado na metodologia *CQT – Controle de Qualidade Total*, que tem por objetivo o controle de qualidade dos produtos e serviços das empresas. No caso do *TDQM* o objetivo é a produção de dados com qualidade, isto é, adequados às aplicações a que se destinam. As atividades previstas no ciclo de administração dos dados são apresentadas na Figura 1 (WANG, 1998):

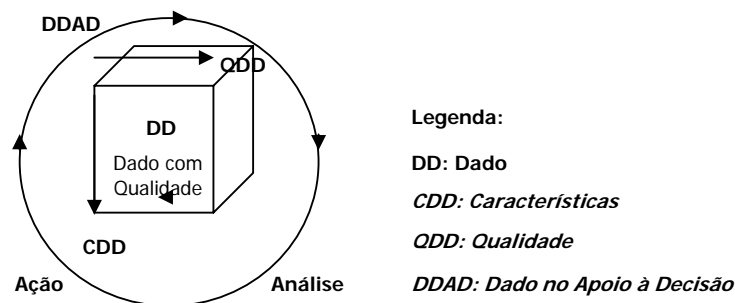


Figura 1 – Ciclo de Administração da Qualidade de Dados. Fonte: Wang (1998)

- **Definição** – a definição das características dos dados é realizada em dois níveis:
  - a) Nível Agente Decisor - identifica as fontes de informação e define parâmetros de seleção, atributos de qualidade e granularidade desejados para os dados;
  - b) Nível Analista de T.I. – identifica os dados nas fontes de informação e levanta as suas características (formato, tamanho, armazenamento), nível de granularidade, frequência de atualização, volume armazenado, série histórica;

- **Medição** – registra medições acerca dos dados levantados, tais como: taxa de erro, dispersão em relação ao valor médio;
- **Análise** – analisa a qualidade dos dados, em relação aos atributos de qualidade e acessibilidade desejados, propondo ações corretivas;
- **Ação** – desenvolve ações de melhoria da qualidade dos dados, tais como: correção de erros, conversão para unidades padronizadas, ajuste da granularidade de acordo com a decisão.

A partir dos atributos de qualidade definidos pelos agentes decisores, das características dos dados, das medições e ações de melhoria da qualidade, os analistas de TI projetam a estrutura física da base de apoio à decisão.

### 2.2.5 PROJETO TDQM

O Projeto *TDQM – Total Data Quality Management* (FISCHER, 2005), desenvolvido pelo *MIT – Massachusetts Institute of Technology*, tem como principal objetivo prover uma base formal para a qualidade de dados. Foram definidos 118 critérios, consolidados em 15 atributos, agrupados em quatro categorias: intrínseca, contextual, representacional e de acessibilidade:

- Categoria Intrínseca – são características inerentes ao dado:
  - Precisão – mede o grau de precisão com que o dado representa o mundo real;
  - Credibilidade – mede o grau de confiança no dado;
  - Imparcialidade – mede o grau de imparcialidade da fonte geradora do dado;
  - Reputação – mede a reputação do dado ao longo do tempo;
- Categoria Contextual – são características associadas ao contexto do processo que está manuseando o dado:
  - Valor Agregado – mede a extensão dos benefícios associados ao seu uso;
  - Relevância – mede a importância do dado para a decisão;
  - Atualização – mede o grau de atualização do dado. Os dados altamente voláteis precisam ser atuais;
  - Completude – mede se todos os valores de um dado estão registrados e se todos os registros de um arquivo contabilizados. Valores nulos em um dado relevante geram um nível baixo de *completude*;
  - Quantidade – refere-se ao volume armazenado de um dado;
- Categoria Representacional – associada à maneira de se representar e interpretar o dado:
  - Interpretação – mede a facilidade de interpretar, isto é, a clara definição dos termos;
  - Entendimento - facilidade de entendimento, sem ambiguidade;
  - Consistência – significa que o formato é consistente entre diferentes aplicações;
  - Representação Concisa – refere-se a uma representação breve e objetiva, isto é, à forma como o dado é representado;
- Categoria Acessibilidade – corresponde à disponibilidade e a forma de acesso aos dados:
  - Acessibilidade – mede a facilidade de acesso e a rapidez de recuperação;
  - Segurança – nível de privacidade dos dados a serem disponibilizados;
  - Manipulação – mede a facilidade de manipular os dados na base de apoio à decisão.

## 2.3. EXPLORAÇÃO DE DADOS E INFORMAÇÕES

A exploração de dados se dá pelo uso de técnicas estatísticas, rotinas OLAP e pelo uso de técnicas de mineração de dados.

### 2.3.1. TÉCNICAS ESTATÍSTICAS UTILIZADAS NA QUALIDADE DE DADOS

Diferentes técnicas estatísticas são utilizadas na seleção, correção, organização e sumarização de grandes quantidades de dados, com o objetivo de gerar informações de qualidade para os tomadores de decisão (THEARLING, 2001).

Algumas medidas são de extremo valor, como:

- a) **Histogramas** – é uma das melhores maneiras de sumarizar dados. Como exemplo, num banco de dados de clientes, é possível identificar a faixa etária da maioria dos clientes, visando direcionar campanhas de marketing.
- b) **Outras Medidas** – outras propriedades importantes dos dados medem (FISCHER, 2005): uma “tendência central” para os dados, o grau de dispersão dos dados, um padrão de distribuição dos dados;

Os estudos que avaliam a “tendência central” dos dados geram medidas do tipo:

- ◆ **Média Aritmética** - valor médio de uma variável ( $M$ ). É uma medida que sofre influência de valores extremos;
- ◆ **Mediana** - valor que divide o histograma em duas partes iguais. É uma medida que não sofre influência de valores extremos;
- ◆ **Moda** – o valor que ocorre com maior frequência;

**Variância** – valor que mede a dispersão dos dados de uma variável, em torno do valor médio. O **Desvio Padrão** é dado pela “raiz quadrada” da variância.

- ◆ **Score Z** – permite que cada ocorrência de um conjunto de dados seja representada em termos de seu desvio em relação à média, medida em termos de unidades de **Desvio Padrão (S)** (FISCHER, 2005). O **Score Z** é definido por:  $Z_j = (X_j - M) / S$

O **Score Z** mede o grau de dispersão da ocorrência, que pode estar incluída em dos quatro conjuntos de ocorrências:

- conjunto com 68% das ocorrências no intervalo:  $- 1 \leq Z_j \leq 1$ ;
- conjunto com 95% das ocorrências no intervalo:  $- 2 \leq Z_j \leq 2$ ;
- conjunto com 99% das ocorrências no intervalo:  $- 3 \leq Z_j \leq 3$ ;
- conjunto com 1% das ocorrências no intervalo:  $Z_j > 3$  ou  $Z_j < - 3$ .

Obs: as ocorrências posicionadas neste último conjunto apresentam um alto grau de dispersão.

- c) **Análise de Séries Temporais** – é utilizada quando se deseja acompanhar o comportamento de uma certa variável ao longo do tempo.
- d) **Análise de Regressão** – está técnica é utilizada na predição de valores, utilizando como base de cálculo, variáveis conhecidas. Com muita frequência, verifica-se que existe uma relação entre duas ou mais variáveis (BARBETA, 2004). Na análise de regressão clássica, há uma única variável dependente e múltiplas variáveis independentes (de predição). Por exemplo, o peso de um adulto do sexo masculino depende, em certo grau, de sua altura.

Se os pontos de coordenadas X e Y parecem estar próximos de alguma curva, uma equação não linear é apropriada para a estimação. Se não há relação indicada entre as variáveis, diz-se que não há correlação entre elas.

As técnicas estatísticas são utilizadas tanto nos procedimentos de análise e avaliação de dados, como nas etapas de enriquecimento e exploração de dados. As ferramentas de mineração de dados aplicam várias técnicas estatísticas.

### 2.3.2. TÉCNICAS OLAP E MINERAÇÃO DE DADOS

O Processo de Exploração da Base de Suporte à Decisão visa a descoberta de informações que possam subsidiar a tomada de decisão. Começa pela geração de hipóteses de padrões e relacionamentos entre as variáveis, que são verificados através da utilização de rotinas *OLAP*.

Também pode ser explorada através da utilização de técnicas de mineração de dados. O procedimento começa pela análise do escopo e objetivos da decisão, seleção de variáveis que se deseja explorar, passando pela avaliação dos modelos de mineração mais adequados, processamento dos modelos selecionados, análise das hipóteses e relacionamentos gerados e validação dos resultados da mineração. Algumas técnicas estatísticas auxiliam a exploração, tais como:

- ◆ Histogramas - é uma das melhores maneiras de sumarizar dados;
- ◆ Análise de Regressão – é utilizada na predição de valores;
- ◆ Correlação entre Variáveis – interpreta o conteúdo de registros como pontos num “espaço de dados”. Exemplo: análise de crédito para novos clientes (A, B e C), onde o novo cliente deve se enquadrar de acordo com as suas características e em relação aos seus vizinhos (THEARLING, 2001), ver Figura 2:

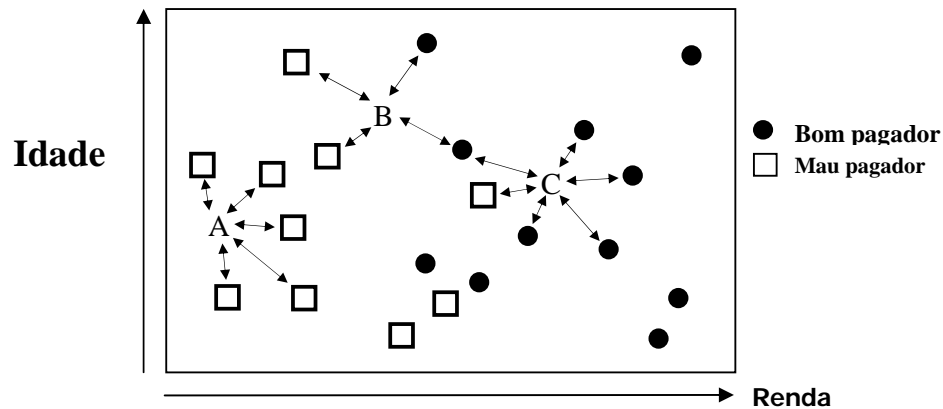


Figura 2 – Análise de Crédito para um Novo Cliente

### 3. METODOLOGIA DE PESQUISA

Para se propor os processos definidos na seção 4 deste trabalho, foi realizado um ensaio teórico sobre metodologias de apoio à decisão.

As metodologias de apoio à decisão analisadas (Murakami,2003; Betencourt, 2000, Kimball, 1988; Santos 2002; Adriaans, 1998; Marakas, 1999; Araujo,1999), ou tratam da construção e exploração de sistemas de suporte à decisão, ou de métodos de análise multicriterial em avaliações qualitativas e quantitativas de alternativas de decisão. A nossa proposta considera o efeito da qualidade de dados na tomada de decisão e integra as técnicas de construção e exploração a procedimentos que possibilitam a utilização de modelos de análise do tipo multicriterial ou não na avaliação das alternativas de decisão.

Esta constatação se dá a partir da análise de diversos trabalhos publicados nos últimos anos, entre os quais destacam-se:

- 1) Decisões que envolvem investimentos em Tecnologia da Informação e que tratam apenas de métodos de análise multicriterial;
  - a) *Decisão Estratégica em TI: um Estudo de Caso*. Milton Murakami. USP. 2003. (MURAKAMI,2003).

**Considerações:** apresenta uma metodologia de processo decisório na seleção de software aplicativo. Utiliza o método AHP e o software Expert Choise, baseado neste método. A preocupação fundamental está na escolha dos critérios de avaliação das alternativas de solução do problema;

- b) *Aplicação da Teoria da Utilidade Multiatributo no Apoio Multicritério à Decisão para Priorização de Sistemas de Informação*. M.S. Araújo. DEP/UFPE. 1999. (ARAUJO,1999).

**Considerações:** o modelo é específico para a priorização de sistemas de informação num processo de planejamento de sistemas. Tem como restrições o fato de não apresentar procedimentos de coleta de informações sobre a decisão, partindo do princípio de que todos os parâmetros já estão definidos e que o mais importante é aplicação do modelo;



- c) *Desenvolvimento de um Modelo de Análise Multicriterial para Justificativa de Investimentos em Tecnologia da Informação*. Paulo Baptista Betencourt. UFRGS. 2000. (BETENCOURT,2000);

**Considerações:** utiliza também o método AHP e o software Expert Choise, baseado neste método. A preocupação está na escolha dos critérios de avaliação, tendo sido desenvolvido um site com o objetivo de colher o maior número possível de opiniões de especialistas na área. Dá bastante ênfase na etapa de sensibilidade, na qual os pesos dos critérios sofrem ajustes;

2) Metodologias que enfatizam a geração de *DW – Data Warehouse*, sem fazer referência aos modelos de decisão multicriterial:

- a) Metodologia descrita pelo Prof. Ralph Kimball, em *The Data Warehouse LifecycleToolkit* (KIMBALL,1998).

**Considerações:** é voltada para o projeto de sistemas de apoio à decisão. Nela podemos observar a ênfase dada à modelagem dimensional, ao projeto físico e ao desenvolvimento do *data warehouse*. O ciclo de vida descrito apresenta um conjunto de atividades inter-relacionadas e seqüenciadas, que recebem uma entrada e geram um produto bem definido, assemelhando-se à construção de um sistema de informação transacional;

- b) Metodologia apresentada por Pieter Adriaans e Dolf Zantinge em *Data Mining* (ADRIAANS,1998).

**Considerações:** é uma metodologia de construção de sistemas de apoio à decisão. Nela podemos encontrar as seguintes falhas: a atividade de seleção e integração de dados é considerada pouco relevante, abordando apenas os dados que já existem na organização em sistemas legados; apresenta o mesmo problema da metodologia proposta por Kimball, ao não fazer referência ao uso de modelos de decisão;

- c) Metodologia apresentada por George M. Marakas em *Decision Support System – In The 21st Century* (MARAKAS, 1999).

**Considerações:** a metodologia apresentada tem como aspecto positivo a abordagem dada às técnicas de modelagem, baseadas em modelos de otimização, utilizados como ferramentas de suporte à tomada de decisão, tais como modelos de programação linear, planejamento financeiro, planejamento da produção. Não é abordado o uso de modelos de decisão multicriteriais.

Ao se concluir esta análise constata-se a ausência de metodologias que abordem os processos de geração de informações apoiada em critérios de qualidade de dados em todas as suas etapas e tratar os dados e informações de forma integrada, através de técnicas de armazenamento e exploração apropriadas, tais como: DW, OLAP, DM e Estatística, além de restringir em suas abordagens os modelos de decisão.

## 4. PROCESSOS DE GERAÇÃO DE INFORMAÇÕES DE QUALIDADE PARA A TOMADA DE DECISÃO

### 4.1 MODELO FUNCIONAL

O modelo funcional definido considera os seguintes instrumentos, técnicas e procedimentos distribuídos de acordo com o diagrama apresentado na Figura 3:

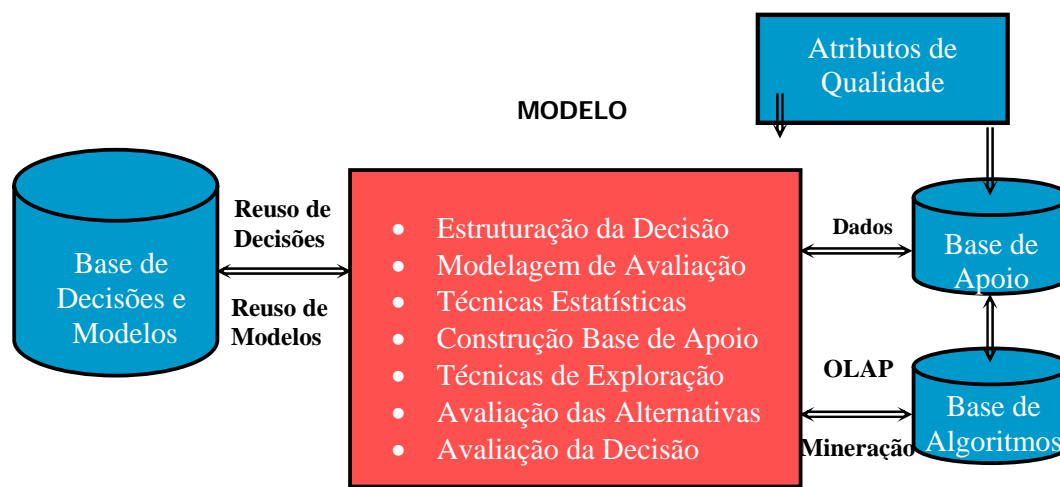


Figura 3 – Modelo Funcional Proposto

## 4.2 OS INSTRUMENTOS DE SUPORTE AO MODELO SÃO:

- *Base de Decisões e Modelos* – tem três objetivos específicos: (a) manter uma *base de decisões*, que possa servir como referência a outras decisões, prioritariamente, de mesma natureza, contendo informações de avaliação sobre os resultados obtidos; (b) manter uma *base de modelos de decisão*, associados às decisões, que possam ser facilmente configurados a uma nova decisão; e (c) servir como guia para a tomada de decisão em curso, orientando os decisores sobre as atividades a serem executadas e registrando todas as ações relevantes durante a decisão;
- *Base de Apoio à Decisão* – tem por objetivo agregar dados correlacionados com a decisão que se deseja tomar, oriundos de diversos sistemas legados e de diferentes *SGBD*, de forma a permitir a extração de informações para a tomada de decisão;
- *Base de Algoritmos* – tem por objetivo disponibilizar um conjunto de algoritmos fundamentados em técnicas estatísticas, *OLAP* e de mineração de dados, que possam ser utilizados na exploração da *base de apoio à decisão*;
- *Atributos de Qualidade* – visam estabelecer atributos de qualidade para a seleção dos dados das diferentes fontes de informação, de forma a gerar subsídios aos analistas de TI na construção da *base de apoio à decisão*;

## 4.3 AS TÉCNICAS QUE PODEM SER UTILIZADAS NO SUPORTE AO MODELO SÃO:

- *Técnicas Estatísticas* – são utilizadas em diferentes procedimentos da metodologia, tais como: (a) na análise dos dados das diferentes fontes de informação; (b) na exploração da *base de apoio à decisão*; e (c) nos procedimentos de avaliação dos critérios qualitativos;
- *Técnicas OLAP* - através da utilização de técnicas *OLAP* é possível explorar a *base de apoio à decisão*, de forma a gerar informações de suporte à decisão. O procedimento começa pela geração de hipóteses de padrões e relacionamentos entre os dados da *base*, pelos analistas de TI/decisores, que são validadas através de rotinas *OLAP*;
- *Técnicas de Mineração de Dados* - tem o objetivo de explorar a *base de apoio à decisão*, através da utilização de técnicas de mineração de dados. Os analistas de TI analisam em conjunto com os decisores os objetivos da decisão, selecionam as variáveis a serem exploradas e os modelos de mineração de dados a serem utilizados, processam os algoritmos selecionados, analisam e validam os resultados da mineração;
- *Técnicas de Modelagem de Avaliação* – visa a construção de modelos para a avaliação das decisões. Considerando que as decisões executivas utilizam, em sua grande maioria, múltiplos critérios de avaliação, quantitativos e qualitativos, tangíveis e intangíveis, é possível através de modelos que considerem diversas alternativas e critérios obter alternativas de ação eficientes, que atendam às necessidades da organização;

## 4.4 OS PROCEDIMENTOS UTILIZADAS NO MODELO SÃO:

- *Estruturação da Decisão* – define a abrangência da decisão, objetivos a serem alcançados, alternativas de ação, atributos de qualidade dos dados a serem utilizados. É de inteira responsabilidade dos decisores;
- *Construção da Base de Apoio à Decisão* – em muitas decisões os dados a serem analisados não estão estruturados em uma base de apoio à decisão, mas espalhados em vários arquivos convencionais ou em diferentes *SGBD*, de diferentes fontes de informação. Cabe aos analistas integrar esses dados de forma a facilitar a geração de informações úteis à tomada de decisão;
- *Exploração* – os procedimentos de exploração visam a descoberta e validação de informações importantes sobre a decisão. É de responsabilidade dos analistas podendo contar com o apoio dos decisores;
- *Modelagem de Avaliação* – a partir de modelos de decisão de mesma natureza e utilizando técnicas de modelagem de avaliação os analistas de decisão, juntamente com os decisores, são capazes de conceber modelos que possam avaliar as diferentes alternativas de ação;

- *Avaliação das Alternativas* – esse procedimento visa avaliar as alternativas de ação, segundo o modelo a ser definido. É de responsabilidade dos avaliadores e decisores;
- *Avaliação da Decisão* – tem por objetivo registrar informações sobre a decisão que possam úteis em decisões futuras de mesma natureza, tais como: justificativas para as alternativas de ação selecionadas, procedimentos de implantação das alternativas de ação, resultados positivos e negativos gerados.

#### 4.5 ELEMENTOS ENVOLVIDOS

A modelo proposto considera os seguintes elementos como imprescindíveis na tomada de decisão:

- Decisores – são pessoas de nível gerencial da organização, conhecedoras do problema e responsáveis diretas pela decisão;
- Analistas de TI – especialistas em banco de dados e em técnicas de exploração de dados (OLAP e mineração de dados), que são os responsáveis pelas atividades de suporte aos decisores e analistas de decisão;
- Analistas de Decisão - são especialistas em modelos de avaliação, de preferência com conhecimento na área em que a decisão está sendo tomada;
- Avaliadores – são especialistas na área em que a decisão está sendo tomada, internos ou externos à organização, que tem a responsabilidade de avaliar as alternativas de ação, segundo o modelo de decisão a ser definido. Com a utilização de múltiplos critérios qualitativos é importante se ter diferentes avaliadores, com visões distintas do mesmo problema.

#### 4.6 A BASE DE DECISÕES E MODELOS

A “base de decisões e modelos” se constitui na base de dados de referência da metodologia, onde estão armazenadas as informações sobre as decisões em andamento, os estudos de casos e os modelo de avaliação utilizados.

O diagrama de classes desta base é apresentado na Figura 4, tendo como classe principal a “*Decisão*”, em torno da qual se situam as demais classes:



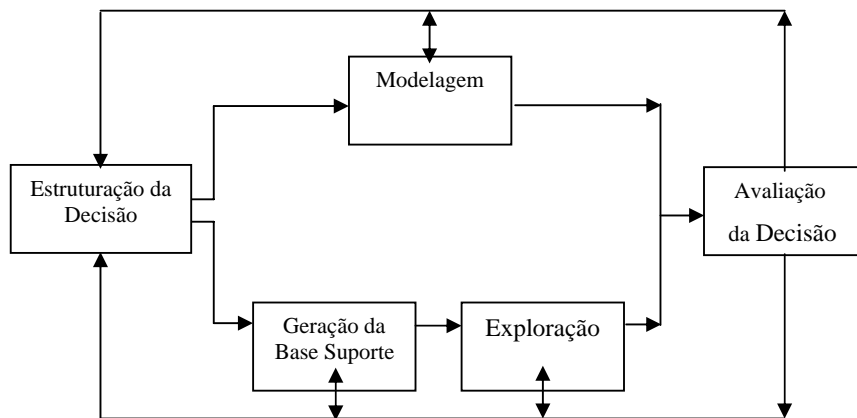
Figura 4 – Diagrama de Classes da Base de Decisões e Modelos

As classes da base de decisões contêm informações relevantes sobre as decisões, tais como:

- Decisão - contém informações que descrevem a decisão;

- Assunto - relaciona os assuntos associados à decisão;
- Fonte de informação - descreve as fontes de informação que alimentam a base de suporte à decisão;
- Alternativa - descreve as alternativas de solução para uma decisão;
- Critério - contém uma tabela genérica de critérios, estando os critérios associados a decisões de sucesso;
- Metadado - descreve os dados selecionados das fontes de informação, da plataforma de transformação e da base de apoio à decisão;
- Arquivo - descreve os arquivos selecionados das fontes de informação e da plataforma de transformação;
- Tabela - descreve as tabelas de fatos e dimensão, da base de apoio à decisão;

A proposta considera também os seguintes processos envolvidos com a tomada de decisão executiva (Figura 5):



**Figura 5 – Processos Envolvidos na Tomada de Decisão Executiva**

Este diagrama apresenta o sequenciamento dos processos, começando pelo processo de "estruturação da decisão", passando pelos processos de "geração do *base de suporte à decisão*", "exploração da *base de suporte*" e "modelagem", chegando ao processo de "avaliação da decisão". O processo "exploração da *base de suporte*" é subdividido em dois outros sub-processos: "exploração - OLAP" e "exploração - mineração de dados".

#### 4.7. PROCEDIMENTOS DE CONTROLE DE QUALIDADE DOS DADOS

Os procedimentos de controle de qualidade dos dados são definidos por processo, obedecendo a seguinte sequência:

##### 4.7.1 PROCESSO DE ESTRUTURAÇÃO DA DECISÃO

Nesse processo os decisores descrevem a decisão, seus objetivos, escopo, alternativas de ação, parâmetros de seleção e qualidade de dados". Entre as atividades associadas com a qualidade dos dados temos:

- a) Descrição da decisão – segundo Herbert Simon (SIMON, 1986) esta é a atividade de inteligência, isto é, de conhecimento do problema, de coleta de informações, sugestões, restrições ou condições que são necessárias considerar. As informações devem ser amplamente discutidas e de conhecimento de todos os envolvidos. Entre as informações a serem geradas temos: descrição da decisão – sucinta e clara, objetivos – o que se deseja atingir com a decisão, escopo – abrangência e elementos envolvidos com a decisão, alternativas – as opções de solução para o problema, requisitos e restrições – que precisam ser considerados na seleção das alternativas;
- b) Indicação das Fontes de Informação – visa definir as fontes de informação que irão alimentar os dados da *base de suporte à decisão*, tais como: fonte - setor ou órgão ou empresa responsável pelas informações, sistemas/processos - relaciona os sistemas e/ou processos sob a responsabilidade da fonte;

c) Definição de Parâmetros de Seleção dos Dados e Atributos de Qualidade - a construção da base de suporte à decisão deve considerar parâmetros de seleção dos dados (FISCHER, 2005), como: periodicidade – define o período em que os dados devem ser coletados, exemplo: de 2000 a presente data, critérios de seleção – define critérios de seleção dos dados, dentro do período especificado, granularidade – descreve o nível de agregação desejado para os dados de forma a facilitar as rotinas de exploração.

A seguir são avaliados atributos de qualidade desejados para os dados a serem selecionados nas fontes de informação: precisão – mede o grau de precisão com que os dados representam o mundo real, podendo ser: Alta ou Média, taxa de Erro (*te*) – define a taxa de erro admissível para os dados, isto é, o número de valores errados pelo número total de valores. A taxa de erro pode ser: Média (*te* < 10%) ou Baixa. (*te* < 5%), credibilidade - mede o grau de confiança nos dados, podendo ser: Alta ou Média, entendimento – mede a facilidade de entendimento dos dados, sem ambiguidade, podendo a facilidade de entendimento ser: Alta ou Média, relevância – mede a importância dos dados para a decisão, podendo ser: Alta ou Média.

A acessibilidade desejada para os dados a serem disponibilizados na *base de suporte à decisão* é avaliada através dos atributos: acessibilidade – mede a facilidade de acesso aos dados da *base*, podendo ser Alta ou Média, recuperação – mede a rapidez de recuperação dos dados da *base*, podendo ser Alta ou Média, segurança – mede a segurança dos dados da *base*, podendo ser Alta ou Média, manipulação – mede a facilidade de manipular os dados da *base de suporte à decisão*, podendo ser Alta ou Média;

Obs: o atributo de segurança é incompatível com os demais atributos de acessibilidade, quanto maior a segurança menor as facilidades de acessibilidade, recuperação e manipulação;

#### 4.7.2 PROCESSO DE CONSTRUÇÃO DA BASE DE SUPORTE À DECISÃO

As atividades desse processo são realizadas pelos analistas de TI e consideram as informações geradas pelos decisores, tais como: fontes de informação indicadas, parâmetros de seleção dos dados e a avaliação dos atributos de qualidade selecionados.

Entre as atividades associadas com a qualidade dos dados temos:

- a) Análise das Fontes de Informação – analisa e interage com as fontes de informação relacionadas, com o objetivo de avaliar a credibilidade e importância das mesmas para a tomada de decisão. Classifica as fontes de acordo com sua importância em: grande importância, média importância, baixa importância;
- b) Analisa e Seleciona Dados Legados – faz um levantamento dos dados associados às fontes de informação com grande ou média importância, de acordo com os parâmetros de seleção de dados definidos pelos decisores. Seleciona os dados com base na análise realizada e na sua relevância para a decisão. Os dados relevantes, mesmo com o nível de qualidade deficiente, são selecionados e passam por um processo de enriquecimento. As informações selecionadas são: agrupamentos de dados (tabelas ou arquivos), atributos com suas características (tipo, formato, tamanho), periodicidade de atualização;
- c) Projeta Plataforma de Transformação – conhecendo a estrutura física dos agrupamentos de dados dos sistemas legados, cujos dados tenham sido selecionados, o analista projeta uma plataforma de transformação, formada por vários tipos de arquivos (texto, sequencial, indexado), de forma a facilitar os procedimentos de transformação e enriquecimento desses dados (KIMBALL, 1998). A finalidade é não interferir na operacionalização dos sistemas;
- d) Atualiza Plataforma de Transformação – extrai os dados selecionados dos sistemas legados e atualiza a plataforma de transformação. Esse procedimento deve se repetir após o fechamento do ciclo de processamento dos sistemas legados;
- e) Avalia Dados da Plataforma de Transformação - Os dados são avaliados com base nos seguintes procedimentos:
  - ◆ Histograma - gera um histograma dos principais dados com o objetivo de medir: a “tendência central”, o grau de dispersão e o padrão de distribuição dos dados. Para os dados com alto grau de dispersão é recomendado o procedimento:
    - ◆ Calcula-se o **Escore Z** das ocorrências, definido por:  $Z_j = (X_j - M) / S$ ,
    - ◆ O conjunto das ocorrências no intervalo:  $Z_j > 3$  ou  $Z_j < -3$  é desprezado e volta-se a calcular **M', S' e Zj'**;

O conjunto das ocorrências no intervalo:  $Zj' > 2$  ou  $Zj' < -2$ , também é desprezado, gerando-se um conjunto com 95% das ocorrências, pertencente ao intervalo:

$-2 \leq Zj' \leq 2$ , com grande representatividade, eliminando-se os valores extremos;

- ◆ Taxa de Erro (**te**) – é medida por: número de valores errados / número total de valores. A taxa de erro pode ser alta, média ou baixa. Para os dados com baixa (**te < 5%**) ou média (**te < 10%**) taxa de erro, os seguintes procedimentos são recomendados:
  - ◆ despreza-se os registros com erros, ou;
  - ◆ tenta-se recuperar os erros utilizando técnicas estatísticas, como: análise de séries históricas, análise de regressão;
- ◆ Atualização – mede o grau de atualização dos dados. É uma qualidade fundamental na escolha do dado, já que um alto grau de volatilidade dificulta a utilização do mesmo no processo de decisão.

Para dados com alto grau de volatilidade deve-se buscar sumarizar valores por unidade de tempo num horizonte mais longo. Como exemplo, as vendas/diária de um grande supermercado pode-se sumarizar por: vendas/semanal, vendas/quizenal ou vendas/mensal;

f) Pré-Processamento – com base na avaliação dos dados, são realizados procedimentos de limpeza de registros ou de identificação de dados com problemas, tais como:

- ◆ Exclusão de registros duplicados;
- ◆ Exclusão de registros com dados incompatíveis e sem possibilidade de restauração;
- ◆ Identificação dos dados inconsistentes, como “não definido”:
  - ◆ ND – para dados do tipo alfanumérico;
  - ◆ 999... – para dados do tipo numérico;

g) Enriquecimento dos Dados - o enriquecimento dos dados consiste na correção dos dados “não definidos”. Muitas vezes são encontrados erros que só podem ser corrigidos com a interação com o negócio, como por exemplo: códigos impróprios, valores fora de domínio, ausência de valores. Várias técnicas estatísticas podem ser utilizadas no enriquecimento dos dados, como:

- ◆ Análise de séries temporais;
- ◆ Análise de regressão;
- ◆ Geração de números aleatórios;

h) Projeto da Base de Suporte à decisão – o projeto da base de suporte à decisão deve considerar os atributos definidos pelos decisores na *categoria acessibilidade*: acessibilidade, segurança e manipulação. Existe uma incompatibilidade entre esses atributos, já que a acessibilidade e manipulação são conflitantes com o atributo segurança. Quanto maior a segurança menor a facilidade de acesso e manipulação.

### 4.7.3 PROCESSO DE EXPLORAÇÃO

Este processo visa a exploração da *base de apoio à decisão*, com o objetivo da descoberta de informações que possam subsidiar a tomada de decisão. É composto por dois Sub-processos: Exploração Analítica – Rotinas *OLAP* e Exploração – Mineração de Dados:

**4.7.3.1 Sub-Processo de Exploração Analítica – Rotinas OLAP** – tem como objetivo explorar a *base de suporte à decisão* através da utilização de rotinas *OLAP*. Começa pela geração de hipóteses de padrões e relacionamentos entre as variáveis, finalizando com o registro do conhecimento adquirido. Atividades envolvidas:

- ◆ **Gera Hipóteses de Padrões e Relacionamentos** – ao analisar a base de suporte à decisão o analista constrói *hipóteses alternativas (Ha)* sobre inter-relações entre os dados. Para cada hipótese formulada é definido um *nível de confiança*, entre 90% e 95%, que é a probabilidade da hipótese ser aceita. Esta atividade é dependente, em parte, da intuição e habilidade do analista em propor hipóteses
- ◆ **Validação das Hipóteses** – o analista analisa o resultado do processamento das aplicações *OLAP* e verifica se a *hipótese* é verdadeira, isto é, se os resultados observados estão

compatíveis com os *níveis de confiança* definidos. Um teste de hipótese pode levar a um novo ciclo de testes, através de um procedimento de refinamento de hipóteses.

**4.7.3.2 Sub-Processo Exploração – Mineração de Dados** – visa explorar a *base de apoio à decisão* através da utilização de técnicas de mineração de dados. Começa pela análise do escopo e objetivos da decisão e seleção de variáveis que se deseja explorar, passando pela seleção dos modelos de mineração e processamento dos algoritmos selecionados, análise e validação dos resultados da mineração.

- ◆ **Analisa Escopo e Objetivos da Decisão** - o analista de TI deve conhecer bem a decisão que se deseja tomar e seus objetivos. Sem esta compreensão nenhum algoritmo, por mais sofisticado que seja, irá prover resultados satisfatórios.
- ◆ **Seleciona Modelo de Mineração** - é preciso explorar várias alternativas de modelos, a fim de encontrar aquele mais adequado à decisão. Os modelos podem ser do tipo: árvore de decisão, análise de vizinhança, regras de associação, clustering.
- ◆ **Dimensiona Modelo de Mineração** – o modelo selecionado precisa ser dimensionado para o objetivo que se deseja, por exemplo, no caso de uma árvore de decisão é preciso definir o número de nodos e a regra associada a cada nodo
- ◆ **Valida Modelo de Mineração** – a utilização de um modelo de mineração de dados requer a validação da sua precisão. Um modelo só está pronto quando todo um ciclo de teste é completado. Um modelo possui vários parâmetros que precisam ser ajustados à realidade que se tem.
- ◆ **Processa Modelo e Gera Hipóteses** – após o processamento do modelo o analista analisa e interpreta os resultados gerados. O resultado do processamento pode gerar informações que não estavam visíveis. Padrões inesperados nos dados podem gerar informações relevantes a serem utilizadas na tomada de decisão. O analista constrói *hipóteses alternativas (Ha)* sobre esses padrões e inter-relações entre os dados. Para cada hipótese formulada é definido um *nível de confiança*, entre 90% e 95%, que é a probabilidade da hipótese ser aceita.
- ◆ **Validação das Hipóteses** – as hipóteses geradas precisam ser validados, sendo esta validação feita através de rotinas *OLAP* aplicadas à *base de suporte à decisão*. O analista analisa o resultado do processamento das aplicações *OLAP* e verifica quais as *hipóteses alternativas (Ha)* são verdadeiras, isto é, aquelas cujos resultados observados estão compatíveis com os *níveis de confiança* definidos.

#### 4.7.4 PROCESSO DE MODELAGEM

Visa a definição de parâmetros que possibilitem a utilização de valores qualitativos e quantitativos na avaliação das alternativas de decisão. Entre as atividades associadas com a qualidade dos dados temos:

- a) Seleção de Critérios de Avaliação – os critérios de avaliação podem ser selecionados a partir da análise de decisões de sucesso de mesma natureza;
- b) Parametrização dos Critérios – os critérios para serem utilizados nos modelos precisam ser parametrizados, o que envolve as seguintes definições:
  - ◆ Valor máximo e mínimo;
  - ◆ Limite de indiferença – é o valor limite da diferença entre a avaliação de duas alternativas, que torna as respectivas alternativas equivalentes;
  - ◆ Limite de preferência – é o valor da diferença entre a avaliação de duas alternativas, a partir do qual uma alternativa é preferível a outra alternativa;
  - ◆ Normalização dos critérios – a normalização dos critérios tem por objetivo estabelecer uma mesma escala de valores (**0 – 1**) entre os diferentes critérios.

#### 4.7.5 PROCESSO DE AVALIAÇÃO

Tem o objetivo de avaliar e processar o modelo de decisão. Entre as atividades associadas com a qualidade dos dados temos:

- a) Avaliação das alternativas de ação – a avaliação é feita em duas etapas:
  - ◆ Primeira Etapa – para os critérios (**Ci**) quantitativos a avaliação é feita de acordo com medidas e/ou informações coletadas;



- ♦ Segunda Etapa – para os critérios (**C<sub>i</sub>**) qualitativos a avaliação é feita por vários avaliadores (**na**) que avaliam as alternativas, uma a uma, atribuindo um valor no intervalo dos limites máximo e mínimo definidos na parametrização do critério. As diversas avaliações tem por objetivo colher diferentes visões sobre o mesmo problema;
- b) Análise do Resultado da Avaliação – é feita em etapas, na seguinte ordem:
- ♦ Primeira Etapa – para cada critério (**C<sub>i</sub>**) de avaliação e considerando o conjunto das avaliações (**A<sub>ij</sub>**) dos diversos avaliadores, calcula-se os seguintes valores:
    - ♦ **M<sub>i</sub>**, **S<sub>i</sub>** e **Score Z<sub>ij</sub>**, definido por:  $Z_{ij} = (A_{ij} - M_i) / S_i$ , onde:  
 $i = 1 \dots nc$  (número de critérios) e  $j = 1 \dots na$  (número de avaliadores)
    - ♦ O conjunto das ocorrências no intervalo:  $Z_{ij} > 3$  ou  $Z_{ij} < -3$  é desprezado e volta-se a calcular **M<sub>i</sub>'**, **S<sub>i</sub>'** e **Z<sub>ij</sub>'**;
    - ♦ O conjunto das ocorrências no intervalo:  $Z_{ij}' > 2$  ou  $Z_{ij}' < -2$  também é desprezado e volta-se a calcular **M<sub>i</sub>''**, **S<sub>i</sub>''** e **Z<sub>ij</sub>''**;
    - ♦ As ocorrências no intervalo  $-1 \leq Z_{ij}'' \leq 1$  são selecionadas, gerando-se um conjunto com 68% das ocorrências, evitando-se as avaliações extremas. Novamente o valor médio é calculado (**M<sub>i</sub>'''**) e este valor é utilizado na função utilidade;
  - ♦ Segunda Etapa – calcula-se o valor da função utilidade utilizando os valores médios das avaliações (**M<sub>i</sub>'''**) obtidos na etapa anterior:
 
$$U(ak) = P1. M1(ak)''' + P2. M2(ak)''' + \dots + Pn.Mn(ak)'''$$
 onde:
 
$$K = 1 \dots I \text{ (número de alternativas)}$$
  - ♦ Terceira Etapa – as alternativas eficientes são selecionadas, isto é, aquelas que não são dominadas por outras alternativas;
  - ♦ Quarta Etapa – a seguir calcula-se o ranqueamento das alternativas eficientes, segundo a função de utilidade, obtendo-se as melhores alternativas de ação;
- c) Análise Sensibilidade dos Pesos - o estudo da análise de sensibilidade dos pesos visa evitar que uma pequena variação nos pesos dos critérios, para cima ou para baixo, possa gerar um novo ranqueamento das alternativas. O peso dos critérios é alterado e calcula-se um novo ranqueamento das alternativas. Este novo ranqueamento funciona como uma simulação da avaliação da decisão, permitindo uma análise comparativa dos resultados;

## 5. CONSIDERAÇÕES FINAIS

A tomada de decisão executiva normalmente é um processo complexo, demorado e quase sempre não estruturado. Tais características influenciam as áreas responsáveis pela tomada de decisão nas organizações.

Para contornar tais obstáculos, o processo de tomada de decisão precisa ser conduzido segundo um conjunto de procedimentos que o oriente e ordene. Esse conjunto de procedimentos, que constitui uma metodologia de suporte à decisão, deve ser formado por atividades que obedecem a uma sequência apropriada de execução, utilizem métodos e ferramentas de trabalho e tenham produtos bem definidos a serem gerados.

De uma maneira geral, os objetivos de uma metodologia de suporte à decisão são:

- Definir um roteiro básico de atividades a serem executadas em uma tomada de decisão executiva;
- Definir produtos a serem gerados em cada processo da metodologia;
- Definir padrões de documentação da decisão a ser tomada.

A nossa proposta se diferencia das demais metodologias analisadas pela ênfase do uso de critérios de qualidade dos dados nos procedimentos envolvidos com a tomada de decisão executiva. Para dar suporte a proposta apresentada está sendo implementado um framework, que será responsável também por armazenar informações sobre decisões de sucesso.

Nada pode assegurar o sucesso de uma decisão executiva, entretanto, sem a adoção de procedimentos bem definidos e de atributos de qualidade na seleção, integração, enriquecimento e exploração dos dados que constituirão a base de suporte à decisão, a probabilidade de insucesso é bem maior.



## REFERÊNCIAS

---

- ADRIAANS, P. e ZANTINGE D. (1998). *Data Mining*.
- ARAUJO, M.S. (1999). *Aplicação da Teoria da Utilidade Multiatributo no Apoio Multicritério à Decisão para Priorização de Sistemas de Informação*. DEP/UFPE.
- BARBETA, P.A. (2004). *Estatística*. Editora Atlas. 2004.
- BETENCOURT, P.R.B. (2000). *Desenvolvimento de um Modelo de Análise Multicriterial para Justificativa de Investimentos em Tecnologia da Informação*. UFRS.
- ENGLISH, L.P. (1999). *Improving Data Warehouse and Business Information Quality*. Wiley Computer Publishing.
- FISCHER, C. (2005). *Introduction to Information Quality*. MIT – Information Quality Program.
- INMON, W.H. (2001). *Data Warehousing: Como Transformar Informações em Oportunidades de Negócio*. Editora Berkeley.
- KIMBALL, R. (1997). *Dimensional Modeling Manifesto*. DBMS.
- KIMBALL, R. (1998). *The Data Warehouse Lifecycle Toolkit*. John Wiley & Sons, Inc.
- MARAKAS, G. M. (1999). *Decision Support System – In The 21st Century*.
- MURAKAMI, M. (2003). *Decisão Estratégica em TI: um Estudo de Caso*. Milton Murakami. USP
- OLSON, J.E. (2003). *Data Quality – The Accuracy Dimension*. Morgan Kaufmann Publishers.
- PIERCE, E.M. (2002). *Assessing Data Quality with Control Matrices*. Communications of the ACM Vol. 47 No. 2.
- PSI. (2003). *Programa Sociedade da Informação. Integração de Informação de Aplicações Governamentais Tecnologias de Informação*.
- PYLE, D. (1999). *Data Preparation for Data Mining*. Morgan Kaufmann Publishers.
- SANTOS, R.F. (2002). *Metodologia e Uso de Técnicas de Exploração e Análise de Dados na Construção de Data Warehouse*. Dissertação em Ciência da Computação. UFPE.
- SIMON, H.A. (1986). *Decision Making and Problem Solving*. Published by National Academy Press. Washington DC.
- STRONG, D.M. and Lee, Y.W. (1997). *Data Quality in Context*. Communications of the ACM Vol. 40 No. 5.
- TAYI, G. e BALLOW, D. (1998) *Examining Data*. Communications of the ACM Vol. 41 No. 2.
- THEARLING, K. (2001). *An Overview of Data Mining Techniques*. Morgan Kaufman Published.
- WANG, R.Y. (1998). *A Product Perspective on Total Data Quality Management*. Communications of the ACM Vol. 41 No. 2. .

**Décio Fonseca**, Doutor em Ciência da Computação/Université Paris VI. Diretor da Covest/Copset.  
E-mail: [decio.fonseca@ufpe.br](mailto:decio.fonseca@ufpe.br)

**Antonio de Vasconcellos Carneiro Campello**, Mestre em Ciência da Computação/UFPE.  
Coordenador de Graduação de Engenharia da Produção/UFPE. E-mail: [acampello@ufpe.br](mailto:acampello@ufpe.br)

**Jairo Dornelas Simião**, Doutor em Administração/ UFRGS. E-mail: [jairo@ufpe.br](mailto:jairo@ufpe.br)