

UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

**DA MATURIDADE ORGANIZACIONAL
PARA A GESTÃO AUXILIADA POR
DATA WAREHOUSE**

Lilia Arantes Corrêa

Orientador: Prof. Dr. Leonardo G. de R. Guedes

Goiânia – 2007

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

LILIA ARANTES CORRÊA

**DA MATURIDADE ORGANIZACIONAL
PARA A GESTÃO AUXILIADA POR
DATA WAREHOUSE**

Dissertação apresentada ao Curso de Mestrado em Engenharia Elétrica e de Computação da Universidade Federal de Goiás, para obtenção do título de Mestre em Engenharia de Computação.

Área de Concentração: Engenharia de Computação
Linha de Pesquisa: Sistemas Inteligentes
Orientador: Prof. Dr. Leonardo Guerra de Rezende Guedes

Goiânia – 2007

Dados Internacionais de Catalogação-na-Publicação (CIP)
(GPT/BC/UFG)

Corrêa, Lilia Arantes.

C824d Da maturidade organizacional para a gestão auxiliada por data warehouse / Lilia Arantes Corrêa. – Goiânia, 2007.

113f. : il.

Orientador: Leonardo G. de R. Guedes.

Dissertação (Mestrado) – Universidade Federal de Goiás, Escola de Engenharia Elétrica e de Computação, 2007.

Bibliografia f.108-112.

1. Armazém de dados – Gestão 2. Data warehouse – Maturidade organizacional 3. Sistema de suporte de decisão – Maturidade I. Guedes, Leonardo G. de R. II. Universidade Federal de Goiás, Escola de Engenharia Elétrica e de Computação. III. Título.

CDU : 004.65



Universidade Federal de Goiás

Escola de Engenharia Elétrica e de Computação

Coordenação do Programa de Pós Graduação em Engenharia
Elétrica e de Computação - CPPEEC

FOLHA DE APROVAÇÃO

DA MATURIDADE ORGANIZACIONAL PARA A GESTÃO AUXILIADA POR DATA WAREHOUSE

LILIA ARANTES CORRÊA

Dissertação defendida e aprovada pela Banca Examinadora constituída pelos senhores:

Prof. Dr. Leonardo Guerra de Rezende Guedes, Orientador – EEE/UFG

Prof. Dr. Reinaldo Gonçalves Nogueira – Efetivo (UFG)

Prof. Dr. Ricardo Luiz Machado – Efetivo (UCG)

Prof. Dr. José Olímpio Ferreira – Suplente (UFG)

Prof. Dr. Getúlio Antero de Deus Júnior – Suplente (UCG)

Goiânia, 27 de janeiro de 2007

Dedicatória

Ao meu esposo e grande amigo, Frede, pela compreensão e apoio
Às minhas adoradas filhas: Giovana e Beatriz
À minha mãe por tudo

Agradecimentos

Agradeço a todos aqueles que participaram da concretização deste trabalho e a todos aqueles que proporcionaram condições para que ele pudesse ser finalizado.

Ao professor Leonardo Guedes, orientador do trabalho, pelo apoio e orientação que possibilitaram acreditar na realização do trabalho.

Aos professores membros da Banca, pela contribuição valiosa de suas observações, e aos professores com quem tive o privilégio e o prazer de aprender.

Em especial, ao Frede, Giovana e Beatriz, pelo apoio e pela paciência de se privarem da minha companhia em diversos momentos, quando o tempo necessário à dedicação deste trabalho era essencial.

Sobretudo a DEUS, que me iluminou no momento de inspiração e nos momentos mais difíceis, que jamais me deixou desistir quando tudo se encaminhava para isso, e que está ao meu lado, assim como ao lado de todos aqueles que Nele crêem...E dos que não crêem também.

Índice

LISTA DE ABREVIATURAS E SIGLAS.....	9
LISTA DE FIGURAS.....	11
LISTA DE TABELAS.....	12
RESUMO.....	13
ABSTRACT	14
1. INTRODUÇÃO	15
1.1. TEMA.....	15
1.2. DELIMITAÇÃO DO TEMA	16
1.3. PROBLEMA	16
1.4. OBJETIVOS	16
1.4.1. <i>Objetivo Geral</i>	16
1.4.2. <i>Objetivos Específicos</i>	17
1.5. JUSTIFICATIVA	17
1.6. HIPÓTESES.....	18
1.6.1. <i>Hipótese Básica</i>	18
1.6.2. <i>Hipóteses Secundárias</i>	18
1.7. METODOLOGIA.....	18
1.7.1. <i>Método de Abordagem</i>	18
1.7.2. <i>Método de Procedimento</i>	19
1.7.3. <i>Técnicas</i>	19
1.7.4. <i>Delimitação do Universo</i>	20
1.7.5. <i>Amostragem</i>	20
1.8. ESTRUTURA DO TRABALHO	21
2. DATA WAREHOUSE	22
2.1 INTRODUÇÃO.....	22
2.2 CARACTERÍSTICAS DO <i>DATA WAREHOUSE</i>	26
2.3 OLAP – ONLINE ANALYTICAL PROCESSING.....	29
2.4 DATA MINING	33
3. TEORIA DE RESPOSTA AO ITEM.....	35
3.1. FUNDAMENTAÇÃO	35

3.2. CARACTERÍSTICAS.....	37
3.3. MODELOS DA TRI	40
3.3.1. <i>Modelo logístico de um parâmetro ou modelo de Rasch</i>	43
3.3.2. <i>Modelo logístico de dois parâmetros</i>	44
3.3.3. <i>Modelo logístico de três parâmetros</i>	45
3.4. ESTIMATIVA DOS PARÂMETROS.....	46
3.5. APLICAÇÃO DA TRI EM MATURIDADE.....	48
4. MATURIDADE E CMM	51
4.1. INTRODUÇÃO.....	51
4.2. CMM – <i>CAPABILITY MATURITY MODEL</i>	53
4.2.1. <i>Histórico</i>	55
4.2.2. <i>Os Cincos Níveis de Maturidade</i>	56
5. FATORES IDENTIFICADORES DA MATURIDADE ORGANIZACIONAL PARA DATA WAREHOUSE	63
5.1. INTRODUÇÃO.....	63
5.2. MAPEAMENTO DOS NÍVEIS DE MATURIDADE	66
6. AVALIAÇÃO DA MATURIDADE ORGANIZACIONAL PARA DATA WAREHOUSE E ESTUDO DE CASO.....	72
6.1. INTRODUÇÃO.....	72
6.2. FATORES.....	73
6.3. LEVANTAMENTO DO GRAU DE IMPORTÂNCIA	77
6.4. MODELO PROPOSTO PARA AVALIAÇÃO DA MATURIDADE ORGANIZACIONAL PARA <i>DATA WAREHOUSE</i>	79
6.5. ESTUDO DE CASO.....	83
6.1.1 <i>Aplicação do Modelo e Análises dos Resultados</i>	90
6.6. CONCLUSÕES DO ESTUDO DE CASO.....	102
7. CONCLUSÃO	104
7.1 CONSIDERAÇÕES FINAIS	104
7.2 SUGESTÃO PARA TRABALHOS FUTUROS.....	107
REFERÊNCIAS BIBLIOGRÁFICAS	109
APÊNDICE.....	114

Lista de Abreviaturas e Siglas

AD	Administrador de Dados
BI	<i>Business Intelligence</i> – Inteligência de Negócio
CCI	Curva Característica do Item
CMM	<i>Capability Maturity Model</i> – Modelo de Maturidade da Capabilidade
DBA	<i>Data Base Administrator</i> – Administrador de banco de dados
DW	<i>Data Warehouse</i> – Armazém de Dados
DWP	<i>Data Warehouse Process</i> – Processos de Armazém de Dados
DWPM	<i>Data Warehouse Process Maturity</i> – Maturidade de Processos de Armazém de Dados
EUA	Estados Unidos da América
ETL	<i>Extract, Transform, Load</i> – Extração, Transformação, Carga
HOLAP	<i>Hybrid On-Line Analytical Processing</i> – Processamento analítico híbrido em linha
KPA	<i>Key Process Area</i> – Áreas-Chave de Processo
MOLAP	<i>Multidimensional On-Line Analytical Processing</i> – Processamento analítico multi-dimensional em linha
OLAP	<i>On Line Analytical Processing</i> – Processamento Analítico em linha
PMI	<i>Project Management Institute</i> – Instituto de Gerenciamento de Projetos
RDBMS	<i>Relational Database Management System</i> – Sistema Gerenciador de Banco de Dados Relacional

ROI	<i>Retorn Of Investiment</i> – Retorno Do Investimento
ROLAP	<i>Relational On-Line Analytical Processing</i> Processamento analítico relacional em linha
SEI	<i>Software Engineering Institute</i> – Instituto de Engenharia de <i>Software</i>
SQL	<i>Structured Query Language</i> – Linguagem de Consulta Estruturada
TCM	Teoria Clássica de Medida
TI	Tecnologia da Informação
TRI	Teoria de Resposta ao Item

Lista de Figuras

Figura 2.1 – Diferentes Tipos de OLAP – Fonte Informatica	31
Figura 3.1 – Exemplo de uma curva característica do item	38
Figura 4.1 – Níveis de Maturidade do CMM	56
Figura 4.2 – Estrutura do CMM	58
Figura 4.3 – Estrutura do CMM.....	59
Figura 6.1 – Planilha eletrônica de auxílio aos cálculos	86
Figura 6.1 – Planilha eletrônica de auxílio na elaboração dos cálculos – Segunda Parte.....	86
Figura 6.2 – Valor Esperado – Maturidade Absoluta	87
Figura 6.3 – Valor Esperado – Maturidade Relativa	88

Lista de Tabelas

Tabela 2.1 – Diferenças entre OLTP e Data warehouse	22
Tabela 2.2 – Diferenças entre data mining e OLAP	33
Tabela 5.1 – Mapeamento entre workflows e atividades	65
Tabela 5.2 – Relação entre fatores e workflows	69
Tabela 6.1 – Grau de Importância dos Fatores Indicativos de Sucesso do Data Warehouse	75
Tabela 6.2 – Fatores Agrupados Por Implantação x Utilização	76
Tabela 6.2 – Maturidade Absoluta e Maturidade Relativa	83

Resumo

A presente dissertação apresenta um modelo para avaliar a maturidade do ambiente organizacional em uma empresa a partir de fatores indicativos de tendência ao sucesso na utilização do *data warehouse* como ferramenta de apoio à decisão gerencial. O modelo visa ser uma ferramenta de apoio na tomada de decisão específica para a implantação ou não do *data warehouse* ou ferramenta de análise para adequação do ambiente organizacional para implantação do *data warehouse*. Isto é possível porque ao identificar os requisitos com mais baixa maturidade, identifica também os insumos, artefatos ou processos de implantação e operação do *data warehouse* ou da organização que necessitam de correções e melhorias para atingir uma maturidade mais elevada e, portanto, indicativa de sucesso na implantação do sistema de apoio à decisão. São apresentados estudos feitos anteriormente que demonstram a utilidade da Teoria de Resposta ao Item (TRI) na verificação de maturidade organizacional. O modelo foi aplicado em uma empresa governamental de tecnologia oportunizando o exercício de análise crítica dos diversos fatores avaliados a partir da mensuração de sua maturidade, apresentando-se como estudo de caso da aplicação do modelo proposto. A organização pesquisada possui maturidade dos fatores indicativos de sucesso do armazém de dados equivalente à média das organizações americanas.

Palavras Chaves: maturidade, armazém de dados, tomada de decisão, Teoria de Resposta ao Item (TRI).

Abstract

The present dissertation presents a model to evaluate the maturity of the organizational environment in a company from indicative factors of trend to the success in the use of the data warehouse as a tool of support to the managerial decision. The model aims at to be a tool of support in the decision-making specific for the implantation or not of the data warehouse or tool of analysis for adequacy of the organizational environment for implantation of the data warehouse. That is possible because when identifying the requirements with lower maturity, also identify the insums, devices or process of implantation and operation of the data warehouse or the organization that needs corrections and improvements to reach a higher maturity and, therefore, indicative of success in the implantation of the system of support to the decision. Done previously studies are presented that they demonstrate the utility of the Item Response Theory (IRT) in the verification of organizational maturity. The model was applied in a governmental company of technology having oportunizando the exercise of critical analysis of the diverse factors evaluated from the mensuração of its maturity, presenting itself as study of case of the application of the considered model. The organization research possesss maturity of the indicative factors of success of the warehouse of data equivalent to the average of the American organizations.

Key words: maturity, data warehouse, decision-making, Item Response Theory (IRT).

Capítulo I

Introdução

1.1. Tema

O planejamento e a manutenção de um *data warehouse* requer tempo, dinheiro e considerável esforço gerencial. Nos dias atuais, não existe um modelo de avaliação de maturidade voltado para este ambiente. Esta avaliação torna-se de grande importância para a evolução dos processos referentes ao desenvolvimento e à operação do *data warehouse*.

A partir desta necessidade, o presente trabalho propõe inicialmente apresentar e discutir alguns métodos de estimação da Teoria de Resposta ao Item aplicados em Ribeiro, 2005 e Fleury, 2006. Posteriormente, uma vez apresentados os fatores impactantes nos processos de *data warehouse*, parametrizá-los e estabelecer seu nível de maturidade, a partir do estudo apresentado em Cordero (2004).

Assim, o tema deste trabalho é avaliar a maturidade dos fatores de um ambiente organizacional que são indicativos de tendência ao sucesso do *data warehouse* e que impactam nos seus processos de implantação e utilização, visando à tomada de decisão gerencial utilizando a Teoria de Resposta ao Item na parametrização e cálculo de seus valores.

1.2. Delimitação do Tema

O trabalho se limita à aplicação do modelo em ambientes organizacionais que tenham *data warehouse* como ferramenta de apoio aos processos decisórios ou ainda em ambientes que tenham tal intenção.

O modelo apresentado trata da maturidade de algumas características organizacionais que possam indicar a tendência de sucesso na implantação e utilização do *data warehouse* como ferramenta de apoio à decisão gerencial.

1.3. Problema

O problema deste trabalho pode ser descrito da seguinte forma: É possível a definição de um modelo para avaliação dos fatores do ambiente organizacional a partir de requisitos indicativos de sucesso da implantação e utilização de *data warehouse* como ferramenta de apoio à decisão gerencial?

1.4. Objetivos

1.4.1. Objetivo Geral

Este trabalho tem por objetivo desenvolver um modelo para avaliação de fatores do ambiente organizacional a partir de requisitos indicativos de sucesso da implantação e utilização de *data warehouse* como ferramenta de apoio à decisão gerencial.

1.4.2. Objetivos Específicos

- Levantar, a partir de revisão bibliográfica, os fatores que possam indicar sucesso na implantação e utilização de *data warehouse* como ferramenta de apoio à decisão gerencial;
- Aplicar a TRI nos fatores apresentados para identificar seu nível de maturidade presente no ambiente organizacional.
- Apresentar um modelo de maturidade organizacional para atendimento do objetivo geral alinhado as práticas apresentadas no modelo de maturidade para processos de *software* CMM (Capablity Maturity Model) CMM, 2001.

1.5. Justificativa

Os trabalhos de Sen (2006) apontam para a necessidade de se elaborar um modelo de maturidade que seja específico aos processos de *data warehouse*. Estes trabalhos sugerem que a aplicação de um modelo de maturidade pertinente a *data warehouse* permite:

- Identificar os fatores que influenciam a maturidade;
- Identificar o nível de maturidade da organização nestes processos;

- Possibilitar melhorias nos processos da organização a partir da identificação do nível de maturidade atual da organização.

1.6. Hipóteses

1.6.1. Hipótese Básica

A hipótese básica deste trabalho é: Um modelo para avaliação da maturidade organizacional a partir dos fatores indicativos de tendência ao sucesso do *data warehouse* como uma ferramenta de apoio a decisão gerencial pode ser proposto.

1.6.2. Hipóteses Secundárias

As hipóteses secundárias são:

- Apóia o processo de tomada de decisão na organização para a implantação e operação do *data warehouse*.
- Apresenta requisitos indicativos de tendência ao sucesso do *data warehouse* como uma ferramenta de apoio a decisão gerencial.
- Identifica o nível de maturidade destes requisitos na organização.

1.7. Metodologia

1.7.1. Método de Abordagem

Neste trabalho utilizou-se o método de abordagem **hipotético-dedutivo**. Esse método consiste em partir de um problema, testar uma hipótese objetivando criticar sua solução e

eliminar o erro. Ao final do teste, o problema é resolvido e/ou novos problemas surgem, e o ciclo é continuado.

1.7.2. Método de Procedimento

Esta pesquisa emprega o método **estruturalista**. Parte-se da investigação dos aspectos inerentes aos processos constantes na elaboração e manutenção de um *data warehouse*. Por intermédio da aplicação de teorias de maturidade, tais como CMM, propõe um modelo de maturidade dos fatores que impactam nos processos de *data warehouse* como uma ferramenta de apoio a decisão gerencial.

A pesquisa empregou ainda o método **fenomenológico**, através da realização de um estudo de caso. Este estudo utilizou um protocolo composto do questionário apresentado no apêndice. O questionário foi respondido pelo gerente do *data warehouse* da empresa estudada. As informações obtidas foram então utilizadas para efetuar os cálculos de maturidade apresentados na conclusão. O protocolo realizado para realização da pesquisa foi: Inicialmente a pesquisadora entrou em contato com a gerente através de telefone informando sobre o envio do questionário por email. Após o recebimento do questionário, o respondente telefonou para a pesquisadora no intuito de sanar algumas dúvidas menores que surgiram. Diante deste fato, o questionário sofre alteração na redação das perguntas, e foi reenviado à empresa. Dentro de uma semana, o questionário foi devolvido à pesquisadora completamente respondido.

1.7.3. Técnicas

Este estudo utiliza as seguintes técnicas de pesquisa:

Pesquisa Bibliográfica: necessária à fundamentação teórica do trabalho e utilizada durante todo o seu desenvolvimento.

Observação Direta Extensiva: empregada no exame dos fatores que caracterizam os processos de *data warehouse*, sendo aplicada durante a fase de coleta dos dados.

Entrevista: corresponde aos diálogos com o indivíduo da organização estudada, objetivando a coleta de dados que possam auxiliar no diagnóstico do problema.

1.7.4. Delimitação do Universo

Por se tratar de um estudo de caso, a pesquisa está relacionada a apenas uma organização, sendo que a mesma possui o *data warehouse* instalado em seu parque tecnológico. A pesquisa foi efetuada em novembro de 2006.

A organização pesquisada atua na formulação e execução de políticas tributárias, arrecadação, gestão de recursos e provisão financeira aos órgãos e entidades da Administração Pública Estadual. Possui 2.473 servidores efetivos e foi fundada em janeiro de 1945.

1.7.5. Amostragem

A aquisição das amostras tem uma forte dependência da disponibilidade da organização em cooperar com a pesquisa. Condição esta que resulta em pesquisa realizada por conveniência ou de acordo com o acesso permitido nas organizações. Assim a quantidade de dados poderá sofrer uma limitação.

A avaliação das hipóteses, portanto, será baseada em amostras de conveniências (não probabilistas) de poucas organizações.

1.8. Estrutura do Trabalho

O primeiro capítulo apresenta uma visão geral do trabalho com os fundamentos metodológicos da pesquisa: problema, hipóteses, variáveis, métodos, técnicas e delimitação do universo. Após esta introdução, o texto é organizado em mais seis capítulos.

O segundo capítulo apresenta uma revisão bibliográfica de *data warehouse*, a ferramenta alvo do estudo.

O terceiro capítulo descreve a Teoria de Resposta ao Item - TRI, demonstrando sua utilização nos últimos anos em assuntos referentes à maturidade.

No quarto capítulo aborda-se o assunto maturidade especificamente. Neste capítulo, uma revisão bibliográfica do CMM é apresentada, dentre outros tópicos afetos à maturidade de processos.

O quinto capítulo apresenta discussões a respeito dos fatores que influenciam a percepção da maturidade de *data warehouse* como uma ferramenta de apoio a decisão gerencial.

O modelo proposto e o estudo de caso são apresentados no sexto capítulo.

No sétimo capítulo apresenta-se a conclusão deste trabalho juntamente com sugestões para futuros trabalhos.

Capítulo II

Data warehouse

2.1 Introdução

No atual cenário dos negócios, observa-se que as organizações têm investido altos recursos financeiros em tecnologia, e nem sempre conseguem obter o retorno planejado e esperado, tais como aquisição de maior fatia de mercado ou redução de custos. Assim, é possível inferir que a detenção de um enorme volume de dados em uma empresa não significa muito. O mais importante e necessário é transformar os dados brutos em informação.

As aplicações *On-Line Transaction Processing* – OLTP atendem, apenas a manipulação de dados operacionais, o que não é suficiente para apoiar o usuário na tomada de decisões estratégicas. Para suprir essas e outras carências dos sistemas de banco de dados no auxílio à tomada de decisões, as empresas passaram a construir os *Data warehouse* – depósitos unificados de dados.

O conceito de *data warehouse* baseia-se na aplicação de antigas idéias que somente agora puderam ser viabilizadas pela conjunção de diferentes tecnologias. O principal objetivo do *data warehouse* é satisfazer as necessidades dos usuários – executivos, gerentes e analistas de negócios – quanto ao armazenamento dos dados que servirão para se realizar as consultas e análises necessárias para o gerenciamento dos negócios.

Os sistemas de apoio à decisão são interativos e têm como principal objetivo ajudar as pessoas com poder de decisão em um negócio - usuários finais - a utilizarem dados e modelos para identificar e resolver problemas, assim como definir resoluções. Os grandes bancos de dados resultantes dos sistemas de processamento de transações são verdadeiras fontes de informações para o processo decisório, embora alguns pontos importantes da recuperação de dados para aplicações de suporte à decisão mostrem que a tecnologia dos bancos de dados, isoladamente, não é adequada às suas necessidades específicas. As diferenças básicas entre as informações armazenadas a partir destes sistemas de transações (OLTP) e as informações disponibilizadas pelo *data warehouse* são mostradas na tabela 2.1.

Característica	OLTP	Data Warehouse
Objetivo	Controle dos processos operacionais	Tomada de decisão
Cliente	Usuários dos sistemas básicos	Gestores do negócio
Dados	Atômicos, atualizados e dinâmicos	Atômicos, consolidados, históricos e estáveis
Estrutura	Normalizada	Dimensional
Tempo de Resposta	Segundos	Minutos
Foco	Orientado à aplicação	Orientado à informação
Acesso	Alto	Baixo a moderado
Atualização	Contínua	Periódica
Aplicação	Estruturada e processamento repetitivo	Não estruturada e processamento analíticos

Tabela 2.1 – Diferenças entre OLTP e *Data warehouse* fonte: Adaptado de INFORMATICA, 2000

Dentre as várias razões para se construir um *data warehouse*, pode-se citar:

- Algumas consultas de suporte à decisão não podem ser expressas ou não são facilmente expressas em *Structured Query Language* (SQL);
- Empresas de grande porte possuem diversas fontes de dados armazenados sob diferentes esquemas, e, muitas vezes, partes da empresa não têm acesso a elas por questões de desempenho ou segurança, ou porque para consultar tais dados esparsos deve-se acessar individualmente cada uma dessas fontes, o que é pouco eficiente;
- A partir de um grande volume de dados, extrair informações relevantes, classificá-las e associá-las – tarefas imprescindíveis para o processo de tomada de decisão – tornam-se difíceis e trabalhosas de serem realizadas.

Segundo Inmon (1997) *data warehouse* é uma coleção de dados dividida por assuntos, integrada, não volátil, que varia com o tempo e que suporta decisões gerenciais.

O *data warehouse* é um grande banco (ou armazém) de dados, normalmente centralizado, que armazena todas as informações sobre as operações da organização, formando uma base histórica. O custo de sua implantação é geralmente elevado, mas se for bem planejado, proporciona um retorno altamente vantajoso para empresa, visto que faz cruzamentos de dados anteriormente desconexos, gerando respostas rápidas, com flexibilidade, a perguntas de difícil resposta. É um meio de prover no tempo apropriado informações completas e corretas, em um formato

compreensível, para a eficaz tomada de decisões em todos os níveis: estratégico, tático e operacional.

O ambiente *data warehouse* é um repositório de dados derivados de sistemas diversos, armazenados sob um esquema único e organizados de forma a otimizar as consultas de apoio à tomada de decisão – Inmon (1997).

A correta utilização do *data warehouse* é vista como algo que traz dois reais benefícios para a empresa: integração de dados e acesso aos dados. Ele remove uma boa parte do processamento e dos dados que são desnecessários e indesejáveis.

Data warehousing – termo bastante empregado pelos usuários – é o processamento dos dados dos bancos de dados fontes até o armazenamento final no *data warehouse*, como também é o processamento dos dados durante as consultas e análises. É um processamento informacional que visa dar suporte ao processo de tomada de decisão, complementando o processamento transacional existente, que suporta a operação do negócio – UNICAMP, 2000.

O armazém de dados pode guardar informações e conhecimentos necessários à tomada de decisão, mas que nem sempre são fáceis de vislumbrar. Torna-se necessário, então, a utilização de ferramentas que organizem e façam buscas nesses dados, tais como **OLAP** e **data mining** - sistema que faz verdadeiras minerações atrás de correlações entre itens aparentemente desconexos, que se ocultam em grandes quantidades de dados.

Um *data warehouse* é um sistema complexo que integra muitos componentes: diversos tipos de *software* e de hardware, redes de computadores, sistemas de comunicações de dados, servidores, mainframes e sistemas de administração de banco de

dados, como também muitas pessoas de diferentes unidades organizacionais, com objetivos diferentes - LAMBERT, 2000.

Tendo em vista a grande complexidade e elevados valores envolvidos, o *data warehouse* pode ser implantado em partes – módulos – denominadas *data marts*. Este tipo de desenvolvimento consiste em agregar a informação por setores. Inicialmente, além de melhorar a performance torna o acesso mais simples para o usuário final.

Data Marts trata-se de um banco de dados específico de uma área, que pode servir para adquirir conhecimentos sobre o funcionamento da ferramenta em sua totalidade, a fim de proporcionar maior segurança em sua implantação, ou para melhorar o fluxo de caixa, por começar a gerar retorno em um módulo, antes de investir nos demais. Portanto, é necessário que a organização decida primeiramente o enfoque da implantação: *data warehouse* ou *data marts*. A organização ainda tem a opção de implantar em um processo sucessivo, ou seja, implantar o *data warehouse* através da implantação de sucessivos *data marts* – Inmon (1997). A implantação em processo sucessivo deve ser feita através do planejamento do todo - *data warehouse* e da implementação em partes - os *data marts* - que vão se somando.

2.2 Características do *Data Warehouse*

A implantação de um *data warehouse* deve respeitar a forma de operar e a missão organizacional da empresa, e ainda prever possibilidades de expansão - escalabilidade e os níveis de detalhes com que os dados serão consultados - granularidade. As consultas aos dados podem começar de diferentes pontos e seguir

caminhos diversos, conforme a curiosidade de quem as faz (análise multidimensional), apresentando facilidade de uso, flexibilidade e grande rapidez (o mais próximo possível do tempo real). Para que se chegue a tão grande rapidez, os maiores processamentos (totalizações) são feitos durante a noite, de modo que as consultas dizem respeito ao período que se encerra com o dia anterior – INMON, 1997.

É necessário se atentar que o importante para os usuários é que a ferramenta consiga proporcionar o suporte necessário aos seus processos decisórios, e não seus conceitos ou definições.

William H. Inmon, um dos precursores, enumera as seguintes características do *data warehouse* - INMON, 1997:

- orientada a negócios - principais assuntos - da empresa como clientes, vendas, produtos, apólices, tratamentos, seguros, viagens, etc., enquanto os sistemas de informações tradicionais são orientados a processos como estoques, entradas e saídas de materiais, compras e vendas, faturamento, contabilidade, etc.;
- integrada, ou seja, trabalha de forma a padronizar os termos e as estruturas técnicas que são utilizados nos sistemas de informações tradicionais, por exemplo, nestes diversos sistemas o sexo pode ser armazenados como: "m" ou "f", "0" ou "1", "x" ou "y", "macho" ou "fêmea", "homem" ou "mulher", "dama" ou "cavalheiro" e outras formas; no *data warehouse* apenas uma destas formas poderá aparecer padronizando esta referência;
- não volátil, ou seja, no ambiente operacional, os dados sofrem as alterações necessárias como: incluir, alterar ou excluir dados; porém, no *data warehouse* os dados permitem apenas

duas atividades: a sua carga para o banco de dados e as consultas; os dados nunca podem ser alterados;

- variável no tempo, ou seja, a estrutura dos dados do *data warehouse* sempre contém algum elemento de tempo, enquanto nos sistemas de informações tradicionais isso não ocorre obrigatoriamente; nestes sistemas, o horizonte de tempo é normalmente de 2 a 3 meses, enquanto no *data warehouse* este horizonte é de 5 a 10 anos.

Para que o *data warehouse* alcance o desempenho planejado e seja operado com um alto grau de sucesso, é necessário a utilização de várias ferramentas. As existentes no mercado atualmente são divididas nas seguintes classes – UNICAMP, 2000:

- de Extração, Limpeza e Transformação (ETL);
- de Banco de Dados;
- de Administração: Automatização de procedimentos administrativos do *Data warehouse* e Administração dos Metadados; e
- de Acesso;
- de Consulta e Relatórios (*Query & Reporting*)
 - Orientada para Questões Estruturadas
 - Visão Tabular dos Dados
 - Resultado em Formato de Relatório
- de OLAP (OnLine Analytical Processing)
 - Orientada para Questões Não Estruturadas
 - Visão Multidimensional dos Dados
 - Resultado em Formato de *Grid's* ou *CrossTab's*
- de *Data Mining*
 - Orientada para Conhecimento Escondido – relacionamentos, padrões e tendências

- de Resultados:
 - Análises Estatísticas
 - Árvores de Decisão
 - Redes Neurais

2.3 OLAP – OnLine Analytical Processing

Organizações com numerosas e complexas informações de negócios, armazenam tais dados em sofisticados ambientes de *data warehouse*, tornando-os poderosas ferramentas de processamento analítico, para que consigam processar os dados o mais rápido possível. As ferramentas OLAP transformam os dados armazenados em resultados estratégicos utilizados para a tomada de decisões de negócios críticas.

O DW é utilizado para armazenar informações e o OLAP para recuperá-las, ambos são especializados para exercer suas funções de forma eficiente. As duas tecnologias são complementares de modo que um bom DW é planejado com produção de relatórios em mente. Desta forma, para explorar o DW completamente é necessário o OLAP que irá extrair e alavancar totalmente as informações nele contidas.

A ferramenta OLAP oferece visões multidimensionais de dados agregados, que podem ser analisados para prover valiosos insight (critérios) de acordo com as necessidades da empresa. As ferramentas OLAP podem ter diferentes métodos de armazenamento de dados: MOLAP, ROLAP e HOLAP.

- MOLAP - Multidimensional On-Line Analytical Processing:

Dados armazenados em forma multidimensional e implementação variando de acordo com a ferramenta. Frequentemente é implementado em um banco de dados relacional, porém não na terceira forma normal. O acesso aos dados ocorre diretamente no banco de dados do servidor multidimensional. As restrições de armazenamento e desempenho limitarão o tamanho do banco de dados, não esquecendo o limite das dimensões que também restringem sua manipulação. A complexidade que existe no processo de carga de um banco de dados multidimensional, pode acarretar a demora no processo. O processo de carga é complexo devido à série de cálculos que devem ser realizados para agregar os dados às dimensões e preencher as estruturas do banco. Depois do processo concluído, ainda é realizado uma série de mecanismos para melhorar a capacidade de pesquisa.

- ROLAP - Relational On-Line Analytical Processing:

Dados são armazenados no modelo relacional como também suas consultas são processadas pelo gerenciador do banco relacional.

- HOLAP - Hybrid On-Line Analytical Processing:

Arquitetura mais recente, na qual ocorre uma combinação entre ROLAP e MOLAP. A vantagem é que com a mistura de tecnologias pode-se extrair o que há de melhor de cada uma:

- alta performance do MOLAP
- escalabilidade do ROLAP

A figura abaixo exemplifica a diferença entre os diferentes tipos de OLAP's existentes.

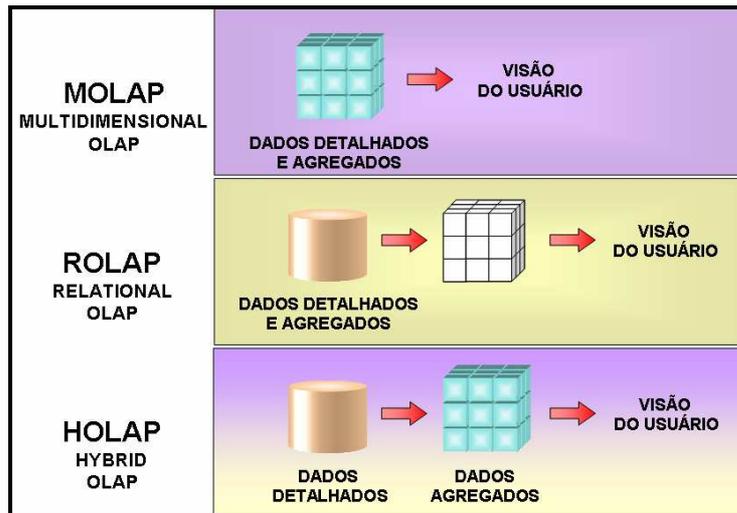


Figura 2.1 – Diferentes Tipos de OLAP – Fonte Informatica

A visão multidimensional consiste de consultas que fornecem dados a respeito de medidas de desempenho, decompostas por uma ou mais dimensões dessas medidas. Podendo também serem filtradas pela dimensão e/ou pelo valor da medida. As visões multidimensionais fornecem as técnicas básicas para cálculo e análise requeridos pelas aplicações de BI. Para se obter a visão multidimensional é necessário compreender outras características:

- **Cubo** é uma estrutura que armazena os dados de negócio em formato multidimensional, tornando-os mais fácil de analisar.
- **Dimensão** é uma unidade de análise que agrupa dados de negócio relacionados. As dimensões se tornam cabeçalho de colunas e linhas.
- **Membro** é um subconjunto de uma dimensão. Cada nível hierárquico tem membros apropriados aquele nível.
- **Hierarquia** é composta por todos os níveis de uma dimensão, podendo ser balanceada ou não. Na hierarquia balanceada os níveis mais baixo são equivalentes, porém, isto não ocorre nas

hierarquias não balanceadas onde a equivalência hierárquica não existe.

- **Medida** é uma dimensão especial utilizada para realizar comparações. Ela inclui membros tais como: custos, lucros ou taxas.

Atualmente, existem muitas ferramentas de OLAP no mercado. Na maioria das ferramentas observa-se a existência de dois componentes: a ferramenta do administrador e a ferramenta do usuário final. O componente do administrador é usado para administrar e gerar os cubos de dados a serem acessados. O componente do usuário final é usado para acessar os dados, extraindo-os de suas bases, com os quais geram relatórios capazes de responder as suas questões gerenciais. Algumas das características destas ferramentas:

- **Consultas ad-hoc:** geradas pelos usuários finais de acordo com suas necessidades de cruzar informações de uma forma não vista e que os levem a descoberta do que procuram. Segundo Inmon "são consultas com acesso casual único e tratamento de dados segundo parâmetros nunca antes utilizados de forma iterativa e heurística" - INMON, 1997.
- **Slice and Dice:** possibilita a alteração da perspectiva de visão. Serve para modificar a posição de uma informação, trocar linhas por colunas de maneira facilitar a compreensão dos usuários e girar o cubo sempre que houver necessidade.
- **Drill down/up:** consiste em realizar exploração em diferentes níveis de detalhes da informação. Com drill down dividi-se um item de resumo em seus componentes detalhados.

2.4 Data Mining

Fazer previsões, detectar tendências futuras, visualizar novas oportunidades de negócios, são algumas das vantagens possíveis através das ferramentas de *data mining*. Essa tecnologia de mineração dos dados ganhou atenção com o surgimento do DW, o qual possui bases de dados bem organizadas e consolidadas. A premissa do *data mining* é uma argumentação **ativa**, ou seja, não é o usuário quem define o problema e seleciona a ferramenta para analisá-lo, mas sim as próprias ferramentas de *data mining* que se encarregam de pesquisar e procurar possíveis relacionamentos entre dados.

O OLAP e *data mining* são partes integrantes de todo e qualquer processo de suporte à decisão. Ainda, nos dias de hoje, a maioria dos sistemas de OLAP tem o foco no provimento de acesso aos dados multidimensionais, enquanto os sistemas de *data mining* lidam com a análise de influência para os dados de uma única dimensão. Grandes empresas como IBM e Oracle contêm versões de seus RDBMS que possuem ferramentas de OLAP e *data mining*. Quando os usuários possuem ferramentas de OLAP e não de mineração de dados, eles gastam boa parte de seu tempo fazendo as tarefas pertinentes a um DM, como classificações e predições das informações recebidas.

O *data mining* trata de técnicas de análise de dados destinadas a descobrir e entender tendências, comportamentos, anomalias e outras relações não óbvias, presentes em grupos – clientes, produtos, empresas, lojas, agências, processos, regiões, etc – que sejam relevantes para um determinado negócio ou atividade. Usa complexas pesquisas em grandes volumes de dados para explorar e identificar relacionamentos entre variáveis que antes eram

previamente independentes. Isso permite que as organizações gerenciem relacionamentos de causa e efeito mais eficientemente, e antecipem políticas para isso.

Permitem diagnosticar o comportamento dos negócios, requerendo para isso a mínima participação do usuário. As ferramentas de *data mining* são baseadas em algoritmos que formam a construção de blocos de inteligência artificial, redes neurais, regras de indução e lógica de predicados que facilitam o trabalho dos analistas de negócios a visualizarem novas perspectivas para velhas questões. Essas ferramentas são extremamente úteis para detecção de fraudes e para previsões de comportamentos. Mas são também soluções complexas que requerem dos profissionais boa capacidade de análise e *expertise*.

As diferenças entre *data mining* e OLAP podem ser vistas na tabela 2.2.

Data Mining	OLAP
Método alternativo – trabalha contra todos os dados de um grupo. Método automático – estuda os dados e apresenta o <i>achado</i> . Visa descobrir padrões nos dados	Trabalho com alguns dados e tem que perfurar (<i>drill down</i>) em busca de detalhes. Explorados manualmente.
Questão desconhecida Ferramenta utilizada para busca de conhecimentos	Exploração através da verificação – o analista elabora uma hipótese para confirmá-la ou refutá-la.

Tabela 2.2 – Diferenças entre *data mining* e OLAP

Capítulo III

Teoria de Resposta ao Item

3.1. Fundamentação

A Teoria da Resposta ao Item – TRI é um poderoso instrumento de avaliação, que vem sendo progressivamente aplicado em processos quantitativos em diversas áreas do conhecimento, tais como: psico-social, onde Granger (1998) estuda o padrão de proficiência do profissional no cenário norte-americano; área médica, onde Deroos (1998) apresentam um estudo sobre as causas da depressão entre crianças americanas de origem africana e de cor branca; área de marketing, onde Bayley (2001) investiga o nível de satisfação dos clientes em uma empresa pública australiana.

A Teoria da Resposta ao Item – TRI é uma ferramenta estatística, existente desde 1950, que possui grandes vantagens sobre outros métodos tradicionais de análise de itens. Porém, sua

utilização é bastante recente no Brasil (década de 90), devido às dificuldades das análises estatísticas envolvidas, que exigem grande poder de processamento e avançados *softwares*.

Trata-se de uma reunião de modelos estatísticos usados para fazer previsões, estimativas ou inferências sobre as habilidades (ou competências) medidas em um teste. Através dos modelos estatísticos é possível prever tais habilidades por meio de correspondências entre a pontuação obtida por um estudante em uma situação de teste e os itens a ele fornecido Hambleton (1991).

Ela tem utilizado duas funções matemáticas para caracterizar os parâmetros métricos dos itens componentes de um teste: a logística e a normal padronizada Muniz (1992). Ambas fornecem informações sobre os parâmetros dos itens através das suas curvas características (CCIs), ressaltando-se que as CCIs não são iguais, apesar de compartilharem alguns aspectos na sua forma geral. Também é possível a obtenção de informações sobre o teste através da sua curva característica (CCT), que fornece o erro padrão de medida, ou seja, a quantidade de erros presente no teste ao avaliar determinada magnitude da variável medida.

A TRI propõe modelos probabilísticos de variáveis que não são medidas diretamente – variáveis latentes – para representar a relação entre a probabilidade de um indivíduo apresentar determinada resposta a um item e seus traços latentes ou habilidades na área do conhecimento avaliada, os quais não podem ser observados diretamente.

Tem como característica principal o item, podendo-se entender por item, tarefas ou ações empíricas que constituem a

representação do traço latente, ou seja, a habilidade que se pretende medir.

Do ponto de vista prático, questões de resposta livre são de difícil uso na TRI (exceto se a resposta for caracterizada como certo ou errada ou, ainda, com algum tipo de graduação). Como resultado, a maioria dos testes usados na TRI são de múltipla escolha e os itens podem ser dicotômicos (certo ou errado) ou politômicos (incorporam variáveis categóricas em suas respostas). O modelo proposto utiliza itens politômicos.

Diante do exposto, percebe-se que a TRI é muito útil ao permitir avaliar o quão preparada uma empresa está para implantar e operar um *data warehouse* como ferramenta de apoio a decisão gerencial, pois, cada representante da organização responderá itens referentes à sua habilidade, tornando a avaliação personalizada.

3.2. Características

A Teoria Clássica de Medida é uma abordagem onde o primeiro interesse é obter uma descrição das características dos parâmetros estatísticos de cada item individualmente. Entretanto, a análise de cada item tem como interesse somente seu efeito no escore total do teste. Também, as características estatísticas do teste dependem inteiramente das características estatísticas dos itens utilizados para construí-lo.

As vantagens advindas da utilização da Teoria da Resposta ao Item sobre a Teoria Clássica de Medida são – CARNEIRO, 2002:

1. Mudança do foco de análise do questionário como um todo para a análise de cada item (questão);

2. Utilização de modelos matemáticos que relacionam um ou mais traços latentes (não observados) de um indivíduo com a probabilidade deste dar certa resposta a um item;
3. Utilização de traço latente, ou seja, habilidade ou proficiência;
4. Possibilidade de, a partir de respostas dadas por indivíduos a um conjunto de itens:
 - estimar os parâmetros dos itens – calibração
 - estimar a habilidade do respondente
 - estimar a proficiência média de um grupo de respondentes
5. Modelagem da probabilidade de resposta a um item como função da habilidade e de parâmetros que expressam certas propriedades dos itens;
6. Permissão de escolha da origem e escala dos dados, tornando os parâmetros dos itens e as proficiências dos indivíduos invariantes.

A TRI formaliza a relação entre os elementos essenciais da situação na qual uma pessoa responde a um problema. Nessa situação, quanto maior a habilidade da pessoa na modalidade requerida pelo problema, maior será a probabilidade que ela responda corretamente. Por outro lado, sendo a habilidade constante, quanto maior for a dificuldade do problema, menor será a probabilidade de que ela o acerte.

A partir da escolha de um modelo de resposta para o ajuste de itens, obtém-se uma distribuição da habilidade de acordo com os parâmetros de cada item, o que oferece informações sobre o

comportamento do mesmo. Essa distribuição é chamada de Curva Característica ao Item - modelo matemático que representa essa situação – figura 01. Em todos os modelos de resposta há uma distribuição com suas particularidades, que indicam a probabilidade de acerto em função da habilidade das pessoas que o respondem e da dificuldade do problema.

Dependendo do modelo, podem ser incluídas outras características do item, como a discriminação e as probabilidades de acerto ao acaso. Em geral, procura-se reunir itens para os quais se supõe que certa proficiência ou habilidade é dominante.

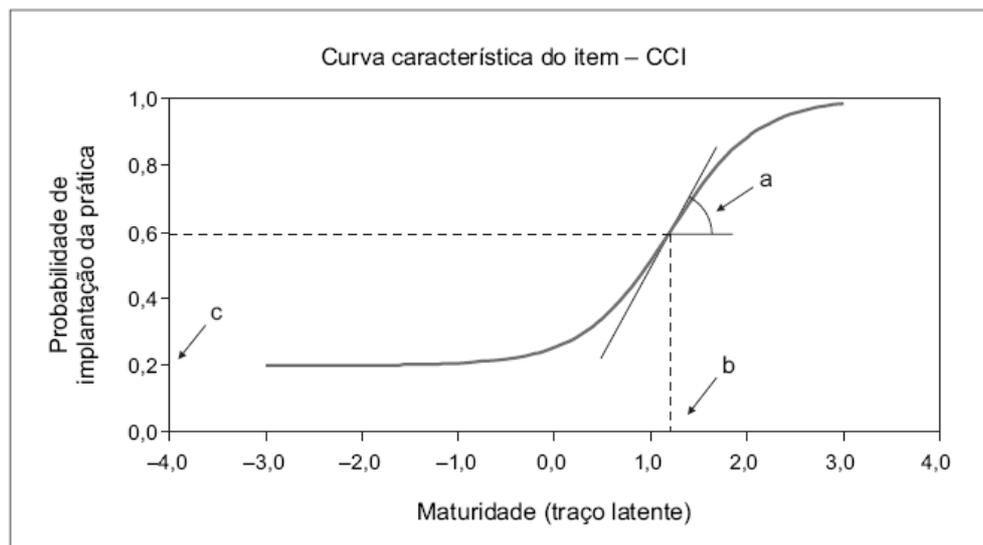


Figura 3.1 – Exemplo de uma curva característica do item
Fonte: Adaptado de CARNEIRO et al (2002)

A Teoria da Resposta ao Item trabalha com traços latentes (aptidões ou habilidades) e se baseia em dois axiomas:

- O desempenho de um examinando em um item do teste é função de um conjunto de traços latentes;

- A relação entre desempenho e habilidades pode ser descrita por uma equação monotônica crescente, chamada de Curva Característica do Item.

Existem ainda dois pressupostos básicos:

- Unidimensionalidade - admite que haja apenas uma habilidade dominante (um fator dominante) na realização de um conjunto de itens; e
- Independência local - indica que para examinandos com uma habilidade dada, a probabilidade de resposta a um conjunto de itens é igual aos produtos das probabilidades das respostas do examinando a cada item individual.

Segundo BAKER, 2001, a independência local implica unidimensionalidade, pois se houver correlação entre o comportamento de um examinando em dois itens de um teste, esta se deve à influência de fatores que não o dominante.

Controlados estes fatores, o fator dominante será a única fonte de variação, tornando as respostas independentes. Uma propriedade importante da TRI é a invariabilidade dos parâmetros, isto é, os parâmetros dos itens obtidos de grupos diferentes de respondentes testados e os parâmetros de proficiência baseados em grupos diferentes de itens são invariantes, exceto pela escola de origem e escala.

3.3. Modelos da TRI

As informações contidas nas CCIs – curva característica do item – sobre os parâmetros métricos dos itens dependem do modelo teórico escolhido. Há vários modelos matemáticos propostos na TRI, que envolvem basicamente a determinação dos níveis de

discriminação, dificuldade dos itens e a resposta aleatória, que se diferenciam em termos do número de parâmetros que avaliam, constituindo modelos logísticos de um, dois ou três parâmetros. Todos os modelos trabalham com traços latentes, o que significa que os sistemas latentes possuem dimensões com propriedades mensuráveis – BAKER, 1992.

A análise da dificuldade, discriminação e validade destes itens, são realizadas através de dados coletados de uma amostra de sujeitos representativa da população em estudo, utilizando-se análises estatísticas. Para a área da educação, por exemplo, a TRI é utilizada na avaliação do desempenho acadêmico dos alunos, através de instrumentos de medição – provas, aplicadas a alunos de Instituições de Ensino selecionadas.

Os modelos logísticos utilizados na TRI dependem principalmente de três fatores:

- Quantidade de variáveis inerentes às habilidades medidas – uma ou mais habilidades;
- Natureza do item – modelo de forma dicotômica ou não dicotômica;
- Número de populações envolvidas – uma, duas ou mais.

Para Swaminathan (1985), deve-se observar as seguintes considerações sobre os modelos de resposta ao item:

- São modelos que supõem que o desempenho de um respondente em um questionário pode ser predito em termos de uma ou mais habilidades pessoais;

- O modelo de resposta de um item especifica a relação entre a pontuação observada de um respondente e as habilidades esperadas ou assumidas que fundamentam os testes;
- O desempenho de um respondente em um teste deve ser estimado a partir da pontuação observada em um conjunto de itens de teste.

A Teoria de Resposta ao Item possui vários modelos a serem estudados. A principal diferença entre os modelos matemáticos são justamente as suas expressões e conseqüentemente suas curvas características e a forma de pontuação empregada. Como o trabalho será feito com uma avaliação (teste) onde as respostas são objetivas, uma descrição dos principais modelos para esse tipo de resposta será relatada a seguir.

Os três modelos mais utilizados são - FLEURY,2006:

1) modelo logístico de um parâmetro: avalia somente a dificuldade do item b ;

2) modelo logístico de dois parâmetros: avalia a dificuldade do item b_i e a discriminação do item a ;

3) modelo logístico de três parâmetros: considera a dificuldade do item b , a discriminação do item a e a resposta correta dada ao acaso c .

Os valores do item de discriminação a são definidos no intervalo de $-\infty$ a $+\infty$. Os itens que possuem o valor de a negativos são descartados, pois não possuem boa capacidade de discriminação, e os valores de a muito elevados não são usuais, resultam em curvas características de item muito acentuadas. Assim, valores típicos do índice de discriminação situam-se entre 0 e -2 . Já

os valores de dificuldade b_i variam de -2 a 2 , sendo que os valores próximos a -2 são considerados itens fáceis e os valores próximos a 2 são considerados itens difíceis BAKER, 2001. E os valores de θ variam neste modelo de -3 a 3 , sendo que os valores próximos a -3 denotam fraca habilidade para os itens em questão e os valores próximos a 3 denotam mais alta habilidade para os mesmos itens BAKER, 2001.

3.3.1. Modelo logístico de um parâmetro ou modelo de Rasch

Rasch começou seu trabalho em medidas educacionais e psicométricas por volta de 1940. Por volta da década de 50, Rasch usando a função de Poisson desenvolveu dois modelos, um para leitura de testes e um modelo para aproveitamento e inteligência de testes, com a finalidade de produção de scores. Com base em seus trabalhos a motivação de Rasch foi representar a probabilidade de resposta como função da habilidade do respondente e a característica do item.

Considerando θ o parâmetro de habilidade do respondente; e b_i a dificuldade do item i . O sucesso do respondente é a razão entre sua habilidade e a soma, da habilidade com a dificuldade do item. Dessa forma, Rasch construiu a seguinte expressão:

$$P_i(\Phi) = \frac{1}{1 + \exp^{-1(\Phi - b_i)}} \quad (3.1)$$

Onde:

$P_i(\theta)$ é a probabilidade de um respondente com habilidade θ , responder a um item i corretamente;

b_i é o índice de dificuldade do item i ;

θ é a habilidade.

Este modelo é um caso especial do modelo logístico de três parâmetros, onde todos os parâmetros assumem uma capacidade semelhante de discriminação, ou seja, possuem o mesmo valor do parâmetro a , e a mínima possibilidade de adivinhação sendo que ocorrerá de se ter valores nulos ou muito baixos para o parâmetro c . Portanto, o único parâmetro que o modelo utiliza é o índice de dificuldade b .

Para o parâmetro a é fixado o valor 1 para todos os itens. Como possui apenas um parâmetro, este modelo se torna atrativo, pois, facilita o trabalho de administração de itens, e também torna mais rápida a tarefa de estimativa do parâmetro em relação aos outros modelos discutidos anteriormente.

3.3.2. Modelo logístico de dois parâmetros

Esse modelo, proposto por Birnbaum em 1968, a partir da substituição da função de distribuição normal, proposta no modelo de Lord (1952) pela função logística, pressupõe a relação monótona entre o valor da variável latente (a ser estimada) do indivíduo e a sua probabilidade de escolha por uma das duas alternativas segundo uma função de distribuição logística parametrizada por coeficientes que representam determinadas características do item.

É muito natural, no contexto de avaliação educacional, onde a variável latente é identificada com a habilidade cognitiva do aluno, e as possibilidades de escolha são acertar ou não o item. O modelo de dois parâmetros expressa a relação entre a variável latente θ e a resposta dada ao item da seguinte forma:

$$P_i(\Phi) = \frac{1}{1 + \exp^{-D \cdot a_i \cdot (\Phi - b_i)}} \quad (3.2)$$

Onde:

a é o parâmetro de discriminação (ou de inclinação) do item i ;

b_i é o índice de dificuldade;

i representa o número do item;

θ é a habilidade;

D é um fator de escala constante e igual a 1.

Utiliza-se o valor de D igual a 1,7 quando se deseja que a função logística forneça resultados semelhantes ao da função ogiva normal, diz-se, então, que o modelo está na métrica normal.

Muito embora, a motivação principal na proposição desse modelo tenha sido o de sua utilização em avaliação educacional, pode-se empregá-lo com finalidade diversa. É o caso de, por exemplo, utilizá-lo para a construção de um índice que mede a condição sócio-econômica de indivíduos de uma população. Podem ser consideradas como variáveis indicadoras da condição sócio-econômica, a posse de determinados bens tais como, automóveis, eletrodomésticos, imóveis e outros.

3.3.3. Modelo logístico de três parâmetros

Obtido a partir do modelo descrito anteriormente, no modelo logístico de três parâmetros há um parâmetro até agora não estudado: o parâmetro c que é chamado índice de adivinhação.

Nenhum dos modelos anteriores levava o fator adivinhação em consideração. O parâmetro c é a probabilidade de acertar um item somente pela adivinhação. É importante observar que por definição o valor de c não varia em função do nível de habilidade. Assim, para a mais baixa e mais alta habilidade do respondente, tem-se a mesma probabilidade de acertar um item pelo fator adivinhação. Para o modelo logístico de três parâmetros tem-se a seguinte expressão:

$$P_i(\Phi) = c_i + \frac{(1 - c_i)}{1 + \exp^{-D \cdot a_i \cdot (\Phi - b_i)}} \quad (3.3)$$

Onde:

a é o parâmetro de discriminação (ou de inclinação) do item i ;

b_i é o índice de dificuldade;

c é o índice de adivinhação;

i representa o número do item;

θ é a habilidade;

D é um fator de escala constante e igual a 1.

3.4. Estimativa dos Parâmetros

No modelo da TRI, uma das etapas mais importantes é a estimativa dos parâmetros dos itens e das habilidades dos respondentes, que em geral são desconhecidos, considerando que apenas têm-se as respostas aos itens do questionário - RIBEIRO, 2005.

Porém, do ponto de vista teórico, pode-se dividir o problema em três situações:

- Conhecimento dos parâmetros dos itens → estima-se apenas as habilidades;
- Conhecimento as habilidades dos respondentes → estima-se os parâmetros dos itens;
- Estima-se os parâmetros dos itens e as habilidades dos respondentes simultaneamente – situação mais comum.

E para que se possa estimar os parâmetros dos itens ou habilidades, é necessário fazer a equalização dos parâmetros dos itens. Entende-se por equalizar tornar comparável, ou colocar os parâmetros provenientes de diferentes provas ou habilidades de respondentes de populações distintas em uma mesma métrica, ou em uma mesma escala.

A equalização pode ser feita de duas maneiras:

- via população – quando um único grupo de respondentes é submetido a provas distintas;
- via itens comuns – através de itens comuns entre os testes, servindo de ligações entre eles.

Existem vários métodos utilizados para a estimativa de parâmetros de itens de provas, os quais dependem do número de grupos e de tipos de prova. Na TRI, este processo de estimativa é conhecido como calibração. Normalmente a estimativa é realizada pelo método da Máxima Verossimilhança, aplicando-se algum processo iterativo como o algoritmo Newton-Raphson e Scoring de Fisher, ou através dos procedimentos Bayesianos, utilizados com bastante frequência.

A tarefa de estimativa dos parâmetros consiste em descobrir o valor de uma determinada variável que mais se aproxima de seu valor real, dado para a execução do cálculo, uma amostra de valores aleatórios regidos por uma função ou modelo de distribuição.

3.5. Aplicação da TRI em Maturidade

A abordagem mais utilizada para verificar o grau de maturidade de uma organização para implantação de um *data warehouse* – ou do *data warehouse* implantado – é a aplicação de um questionário a ser respondido pelas organizações. Com base nas respostas obtém-se um escore que indica o nível de maturidade das organizações nestas implantações. De acordo com a Teoria Clássica de Medida – TCM, o escore observado é composto de um escore verdadeiro (o nível real de maturidade) e um erro de medida – CARNEIRO, 2002.

Na TCM são avaliadas as propriedades psicométricas dos itens (questões), especificamente, a confiabilidade (consistência internada escala), a análise do item e a validade do construto. Contudo, uma característica dessa teoria consiste no fato da análise psicométrica do construto – Construção puramente mental, criada a partir de elementos mais simples, para ser parte de uma teoria – ter ênfase no instrumento de medição como um todo, não avaliando as diferenças de item para item – CARNEIRO, 2002.

Assim, tem-se as seguintes características da avaliação feita através da Teoria Clássica de Medida - TCM:

- Os resultados dependem do particular conjunto de questões do formulário e dos indivíduos que a fizeram, ou seja, as análises e

interpretações estão sempre associadas à prova / formulário como um todo e ao grupo de indivíduos; e

- A comparação entre organizações ou grupos de indivíduos somente é possível quando eles são submetidos às mesmas provas ou, pelo menos, ao que se denomina de provas paralelas, quase sempre difíceis de serem construídas.

Diante das limitações da TCM, Carneiro (2002) sugeriu a criação de um modelo de análise dos construtos da GQT com base de Teoria da Resposta ao Item – TRI, com ênfase na definição e na interpretação dos parâmetros do modelo sugerido, no estudo de sua adequabilidade e na discussão de uma aplicação. Assim, foi demonstrado que a TRI pode ser uma poderosa ferramenta na avaliação da maturidade de GQT. Os trabalhos desenvolvidos, visando investigar os fatores críticos de sucesso da GQT nas organizações, levantaram vários questionamentos que só obtiveram respostas com o uso da TRI.

Em “Um Método para Avaliação de Maturidade Gerencial em Empresas de Tecnologia através da Teoria de Resposta ao Item”, RIBEIRO, 2005 também utilizou-se da TRI para avaliar a maturidade. Seu trabalho propõe um método para avaliar o quão aptos estão os gerentes, bem como delimitar, avaliar grau de maturidade conceitual referente aos aspectos gerenciais relativos às habilidades administrativas, técnicas e pessoais dos gerentes em empresas de tecnologia.

Tem-se ainda a utilização da TRI aplicada à avaliação da maturidade no trabalho de Fleury (2006). Neste trabalho, a proposta do autor é possibilitar delimitar e avaliar o grau de maturidade conceitual referente às ações específicas relativas às habilidades

administrativas, técnicas e pessoais para cada nível de gerente, diretor e supervisor em organizações de Engenharia Elétrica.

Capítulo IV

Maturidade e CMM

4.1. Introdução

Maturidade é a extensão para a qual um processo específico é explicitamente definido, gerenciado, medido, controlado e efetivado – GONÇALVES, 2001.

Atingir um grau de maturidade significa alcançar um patamar de melhoria. Este patamar só é atingido pela organização, através da implementação, de maneira bem-sucedida, de um conjunto predeterminado de processos - quando se obtém o aumento da capacidade de um conjunto finito de processos prescritos por um determinado nível. Uma vez atingidos estes patamares, progressivamente levará a organização a resultados mais efetivos em termos de qualidade, custos e prazos – GONÇALVES, 2001.

A maturidade representa o potencial de crescimento de capacidade e indica a riqueza do processo da organização e a

consistência com que o mesmo é aplicado em todos os seus projetos. Em uma organização madura, o processo é bem compreendido, geralmente feito através de documentação e treinamento, e está sendo continuamente monitorado e melhorado pelos seus usuários – GONÇALVES, 2001.

A maturidade de processo implica que a produtividade e a **qualidade** – resultantes do processo – possam ser continuamente melhoradas através de ganhos consistentes na disciplina alcançada com a sua utilização.

Quando uma organização obtém ganhos na maturidade de um processo, ela o institucionaliza por meio de políticas, padrões e estruturas organizacionais. A institucionalização exige a construção de uma infra-estrutura e de uma cultura corporativa que possa dar suporte aos métodos, práticas e procedimentos de negócio que perdurem após possíveis afastamentos daqueles que originalmente os definiram – GONÇALVES, 2001.

Sem uma estratégia organizada, a gerência e a equipe de projetos terão dificuldades de chegarem a um consenso sobre a prioridade das atividades envolvidas em um processo. Para se conseguir resultados duradouros a partir de esforços em melhoria de processos, é necessário projetar um caminho evolutivo que incremente, em estágios, a maturidade do processo de *software* da organização.

Uma estrutura de maturidade ordena esses estágios pré-estabelecidos, onde os resultados positivos alcançados em cada estágio são utilizados como embasamento para o próximo, objetivando melhorias no processo em sua totalidade - HUMPHREYS,1987. Dessa forma, uma estratégia de melhoria

projetada a partir de uma estrutura de maturidade orienta o caminho a ser seguido para a contínua melhoria do mesmo, e seus avanços.

A definição de uma estrutura de maturidade também permite identificar deficiências na organização, não objetivando fornecer soluções rápidas para projetos em dificuldades.

Por um *data warehouse* ser construído para satisfazer os usuários, é importante que os usuários considerem o processo de DW **maduro** e dinâmico, e não caótico. Devem estar aptos a ver os frutos da maturidade dos processos de *data warehouse* através do alcance a dados de alta qualidade, da criação dos tipos corretos de cubos de usuários e da geração de relatórios úteis.

Um processo de *data warehouse* maduro com suas propriedades controladas, repetíveis, e otimizadas pode resultar em economias de custo significativas.

Para satisfazer a comunidade de usuário, é importante que o departamento de *data warehouse* da organização determine o retorno no investimento (ROI) para seus projetos, o que não é, geralmente, uma tarefa fácil.

4.2.CMM – *Capability Maturity Model*

O CMM – Modelo de Maturidade da Capabilidade – é um *framework* (uma estrutura) voltado à maturidade, que descreve princípios e práticas relacionados ao processo de desenvolvimento de produtos e serviços tecnológicos. Objetiva auxiliar organizações a melhorar a capacidade de seus processos, através de um caminho evolucionário.

O CMM fornece às organizações um guia de como obter controle em seus processos para desenvolver e manter *software* e como evoluir em direção a ter uma cultura de engenharia de *software* e alcançar excelência de gestão – GONÇALVES, 2001.

O CMM foi projetado para guiar as organizações de *software* no processo de seleção das estratégias de melhoria contínua, determinando a maturidade atual do processo e identificando as questões mais críticas para a qualidade e a melhoria do processo de *software*. Assim, quando uma organização atinge um nível de maturidade, considera-se que seus processos alcançaram uma determinada capacidade, ou seja, tem mecanismos que garantem a repetição sucessiva de bons resultados (inclusive futuros) relacionados principalmente à qualidade, custos e prazos – GONÇALVES, 2001.

Focando em um conjunto limitado de atividades e trabalhando agressivamente para concluí-las com êxito, a organização pode melhorar o processo em toda a sua estrutura, possibilitando ganhos contínuos e duradouros na capacidade do processo de *software* – CMM, 2002.

Os principais objetivos do CMM são – CMM, 2002:

- Auxiliar o gerenciamento e mudança de processo;
- Auxiliar a melhoria do processo interno de *software*;
- Prover um guia já comprovado mundialmente para a melhoria contínua dos processos das organizações de forma gradativa e disciplinada;
- Prover uma estrutura básica para a realização de avaliações e comparações eficientes e eficazes.

4.2.1. Histórico

A estrutura em estágios do CMM é baseada em princípios de qualidade de produto dos últimos sessenta anos. Nos anos 30, Walter Shewhart promulgou os princípios de controle estatístico da qualidade. Seus princípios foram desenvolvidos e demonstrados com sucesso no trabalho Deming (1986) e Juran (1988). Esses princípios foram adaptados pelo Software Engineering Institute - SEI dentro da estrutura de maturidade que estabelece a gestão de projeto e os fundamentos de engenharia para o controle quantitativo do processo de *software*, que é a base para a contínua melhoria do processo.

O Software Engineering Institute é uma entidade mantida pelo governo norte-americano e patrocinada pelo Departamento de Defesa dos Estados Unidos - US DoD, com o propósito central de auxiliar as organizações a realizar melhorias mensuráveis em suas capacidades relativas à engenharia de *software*. O SEI - Instituto de Engenharia de *Software* - surgiu com a finalidade de ajudar o US DoD a selecionar fornecedores cada vez mais qualificados e capazes. Os modelos de maturidade e métodos de avaliação criados por esta entidade se tornaram referência mundial em melhores práticas e melhoria de processos.

A estrutura de maturidade da qual esses princípios de qualidade foram adaptados foi primeiramente inspirada por Philip Crosby em seu livro *Quality is Free* CMM,2002. O quadro de maturidade de gestão da qualidade de Crosby descreve cinco etapas para a adoção das práticas de qualidade. Essa estrutura de maturidade foi adaptada para o processo de *software* por Ron Radice e seus colegas, trabalhando sob a direção de Watts Humphrey da IBM

RADICE,1985. Humphey trouxe essa estrutura de maturidade para o SEI - Instituto de Engenharia de *Software* em 1986 e acrescentou o conceito de níveis de maturidade e desenvolveu o fundamento para o seu uso atual através da indústria de *software*. Desde 1990, o SEI, com a ajuda de muitas pessoas do governo e da indústria, tem expandido e refinado o modelo baseado em vários anos de experiência na aplicação deste modelo na melhoria do processo de *software*.

4.2.2. Os Cincos Níveis de Maturidade

A melhoria contínua dos processos baseia-se em muitas e pequenas etapas evolutivas, e não em inovações revolucionárias. O CMM fornece uma estrutura para organizar essas etapas evolutivas em cinco níveis de maturidade que estabelecem fundamentos sucessivos para a contínua melhoria do processo – CMM, 2002.

Esses cinco níveis de maturidade definem uma escala ordinal para medir a maturidade de um processo de *software* da organização e para avaliar a sua capacidade de processo de *software*. Os níveis também ajudam uma organização a priorizar seus esforços de melhorias, CMM - 2002.

Nível de maturidade é um estágio evolutivo bem definido em busca de um processo de *software* maduro. Trata-se de uma base evolucionária bem definida, na qual se obtém um processo maduro – CMM, 2002.

Cada nível consiste de uma série de metas de processo que, quando satisfeitas, estabilizam um componente importante do processo. Cada nível de maturidade fornece uma gama de fundamentos para a melhoria contínua do processo. Cada nível

compreende um conjunto de objetivos de processos que, quando satisfeitos, estabilizam um componente importante do processo de *software*. Alcançando cada nível da estrutura de maturidade, estabelece-se diferentes componentes no processo de *software*, resultando em um crescimento na capacidade de processo da organização - GONÇALVES, 2001.

A utilização de um modelo de maturidade faz com que a organização priorize ações de melhoria para o crescimento da maturidade do processo. A estrutura do CMM é mostrada na figura 4.1, onde as setas indicam o tipo de capacidade de processo que está sendo institucionalizado pela organização a cada etapa da estrutura de maturidade.

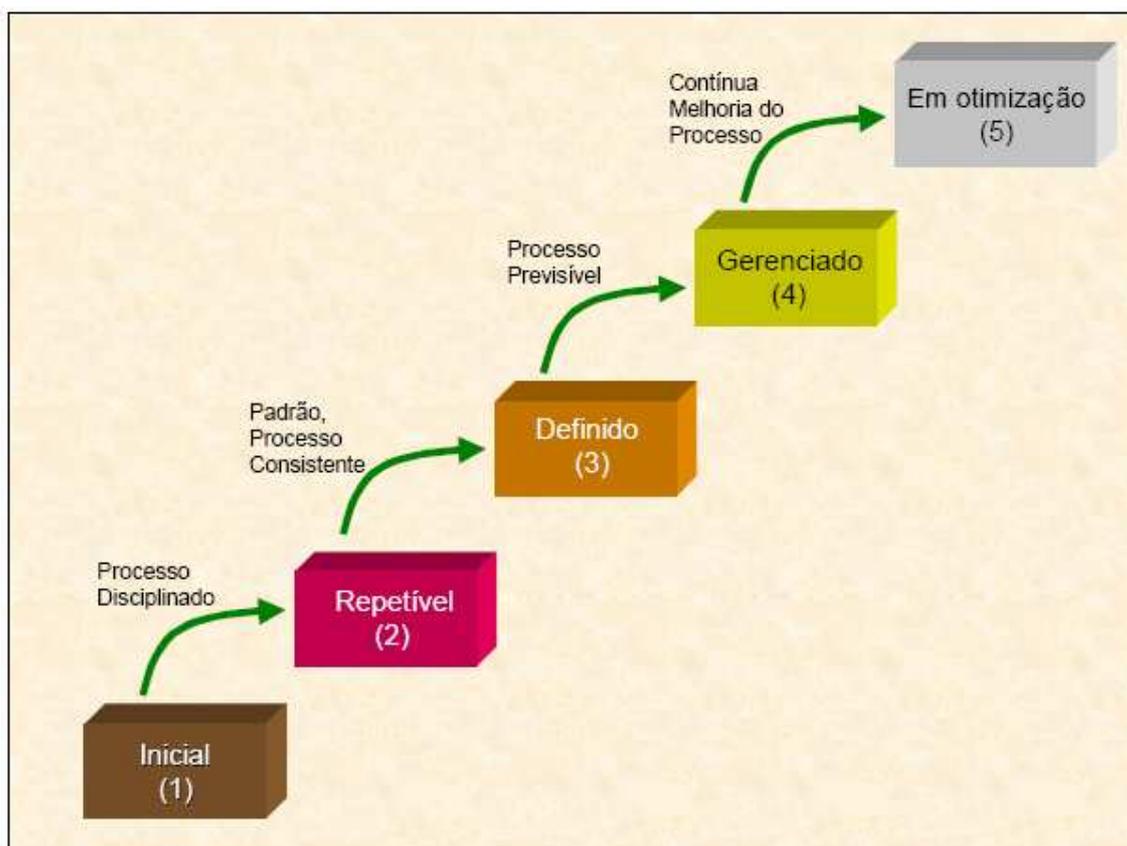


Figura 4.1 – Níveis de Maturidade do CMM - Fonte: Gonçalves (2001)

Os níveis de maturidade constantes no CMM são – CMM, 2002:

- **Nível Inicial:** O processo de *software* é caracterizado como ad hoc e até mesmo caótico. Poucos processos são definidos e o sucesso depende de esforço individual de todos os membros da organização.
- **Nível Repetível:** Os processos básicos de gestão de projeto são estabelecidos para acompanhar custo, cronograma e funcionalidade. A necessária disciplina do processo existe para repetir sucessos anteriores em projetos similares.
- **Nível Definido:** O processo de *software* para as atividades de gestão e engenharia é documentado, padronizado e integrado em um processo de *software* padrão para a organização. Todos os projetos utilizam uma versão aprovada do processo de *software* padrão para desenvolver e manter *software*.
- **Nível Gerenciado:** Medidas detalhadas do processo de *software* e da qualidade do produto são realizadas. O processo e os produtos de *software* são quantitativamente compreendidos e controlados.
- **Nível Em Otimização:** A melhoria contínua do processo é realizada com base no *feedback* quantitativo do processo e do produto.

Cada nível de maturidade indica o nível de capacidade do processo de desenvolvimento de *software* da organização.

Com exceção do Nível 1, cada nível de maturidade é decomposto em áreas-chave de processo (KPA – *Key Process Area*), conforme ilustrado na figura 4.2, que conduzem ao alcance de metas

de melhoria do processo, indicando os pontos que as organizações devem focar para melhorar seu processo de *software*.

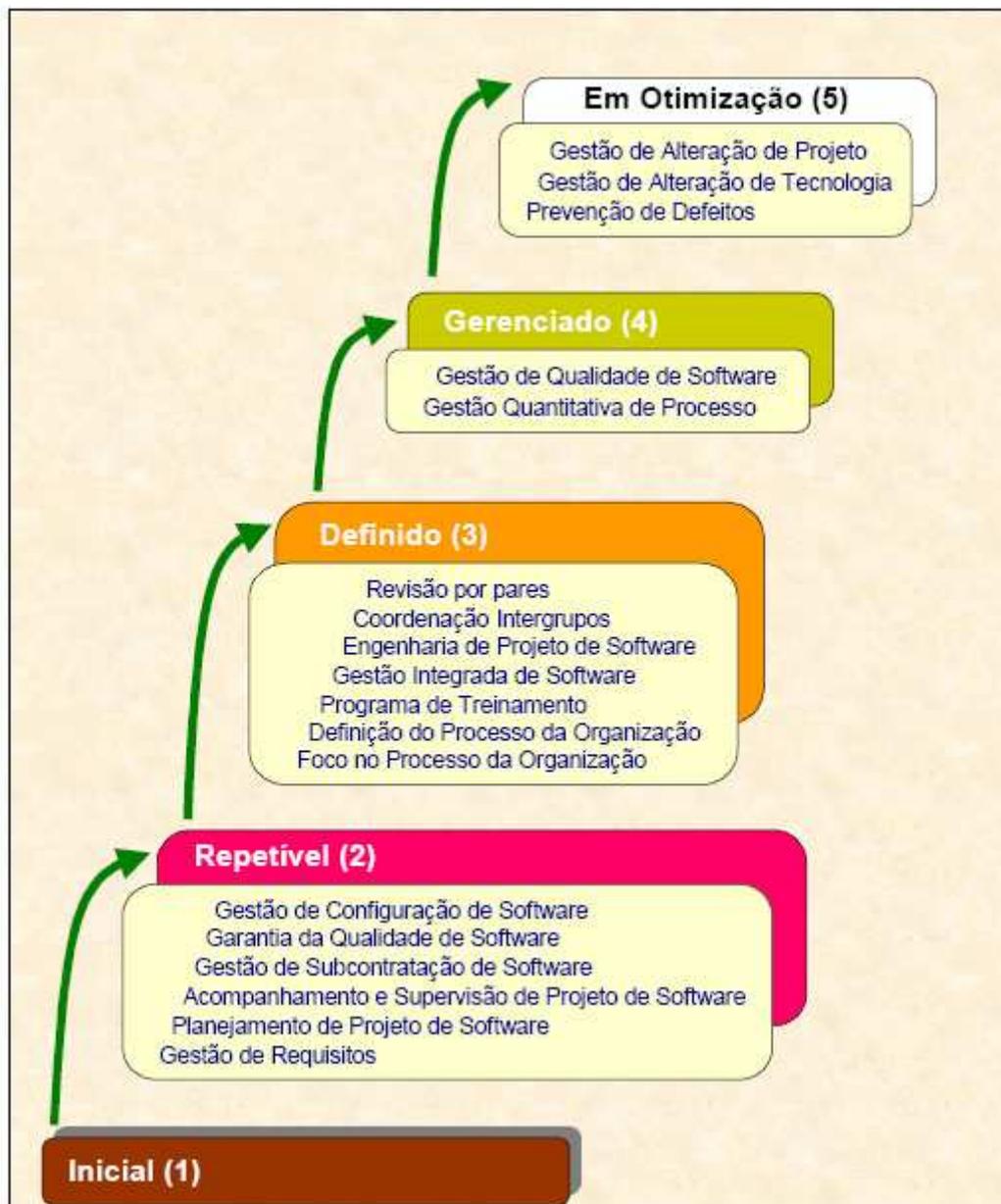


Figura 4.2 – Estrutura do CMM - Fonte: Gonçalves (2001)

As áreas-chave estão organizadas em cinco seções, chamadas de características comuns, que por sua vez, contêm práticas-chave. As características comuns determinam as características de institucionalização ou de implementação das

práticas-chave, enquanto que as práticas-chave descrevem as atividades e/ou infra-estrutura necessárias para satisfazer as metas de uma KPA – área de processo chave. Esta estrutura é ilustrada na figura 4.3

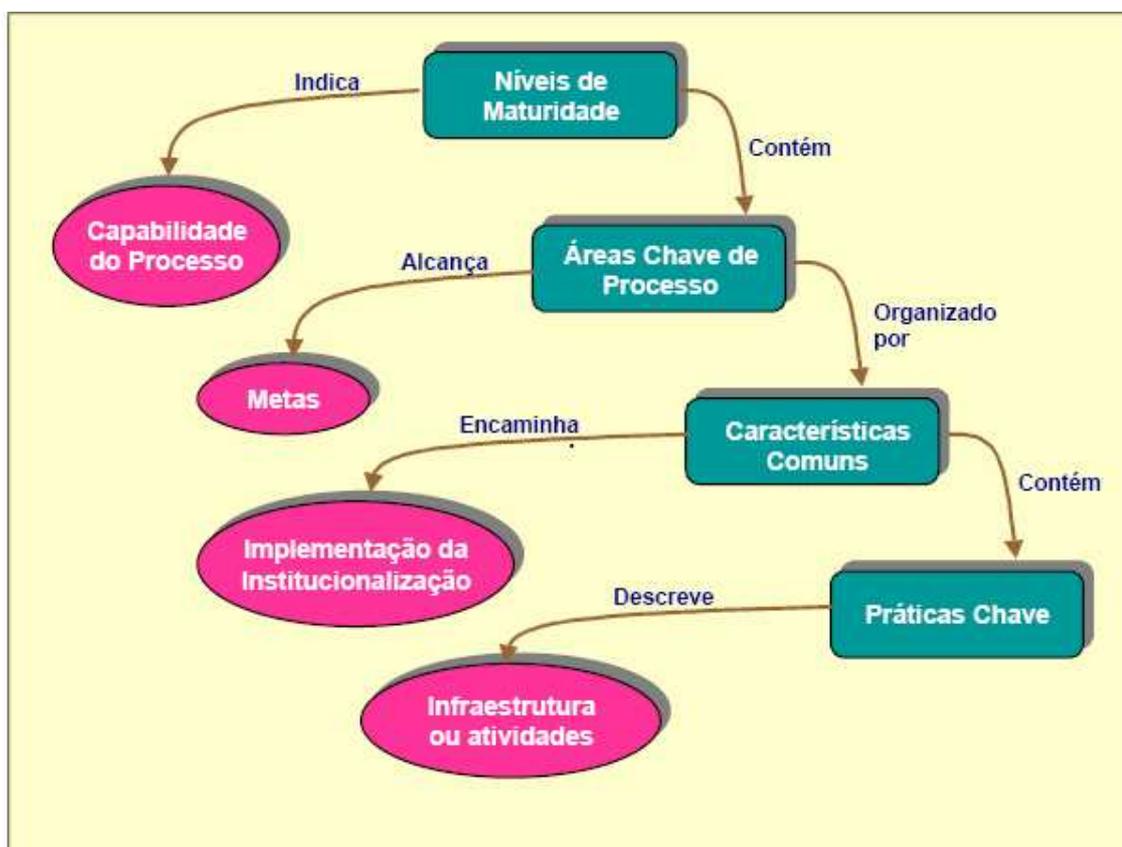


Figura 4.3 – Estrutura do CMM - Fonte: Gonçalves (2001)

Segundo Santander (2000), apesar de grande parte das organizações envolvidas com o desenvolvimento de *software* possuírem processos imaturos, os quais são enquadrados no nível Inicial do CMM, existe a preocupação por parte destas empresas em melhorar e disciplinar seus processos de *software*, objetivando atingir os requisitos mínimos exigidos para alcançar o próximo nível do CMM. Fiorino (1998) observa que a dificuldade de uma empresa em sair do nível Inicial, que é a realidade da grande maioria das empresas, reside em iniciar um processo de mudança cultural com ênfase em processos e foco gerencial para atingir o nível Repetível.

O CMM é um modelo **descritivo**, no sentido de descrever atributos essenciais (chaves) que seriam esperados para caracterizar uma organização em um nível particular de maturidade – GONÇALVES, 2001.

É um modelo **normativo**, no sentido de que as práticas detalhadas caracterizam os tipos normais de comportamento que seriam esperados em uma organização que desenvolve projetos em larga escala num contexto de contratação governamental. A intenção é que o CMM tenha um nível suficiente de abstração que não restrinja a maneira que o processo de *software* é implementado pela organização; ele simplesmente descreve o que normalmente seria esperado dos atributos essenciais do processo de *software* – GONÇALVES, 2001.

O CMM **não é prescritivo**; ele não diz à organização como melhorar. O CMM descreve a organização em cada nível de maturidade sem prescrever os meios específicos para consegui-lo. Pode levar vários anos para se passar do Nível Inicial para o Nível Repetível; a movimentação entre os outros níveis geralmente levará cerca de dois anos. A melhoria de processo de *software* ocorre dentro do contexto dos planos estratégicos e dos objetivos de negócio da organização, da sua estrutura organizacional, das tecnologias em uso, da sua cultura social e sistema de gestão – GONÇALVES, 2001.

O CMM está voltado para os aspectos da Gestão da Qualidade Total; a melhoria de processo bem sucedida implica que os aspectos fora do escopo do processo de *software* também sejam encaminhados (por exemplo: as questões pessoais envolvidas nas mudanças da cultura organizacional que possibilitem a implementação e a institucionalização das melhorias de processo) – GONÇALVES, 2001.

Os aprimoramentos na previsão dos resultados de um projeto assumem que os resultados do projeto de *software* se tornam mais previsíveis à medida que a interferência, na forma de re-trabalho, é eliminada.

Sistemas sem precedentes complicam o quadro, uma vez que as novas tecnologias e as novas aplicações reduzem a capacidade do processo através do aumento da variabilidade. Mesmo no caso de sistemas sem precedentes, as características das práticas de gestão e de engenharia de muitas organizações maduras ajudam a identificar e encaminhar os problemas no ciclo de desenvolvimento antes mesmo de serem detectados em organizações menos maduras. A detecção antecipada de defeitos contribui para a estabilidade e o desempenho do projeto, eliminando o re-trabalho durante as fases posteriores.

A gestão de riscos é uma parte integrante da gestão de projeto num processo maduro. Em alguns casos, um processo maduro pode significar que projetos fadados ao fracasso sejam identificados mais cedo no ciclo de vida de *software* e o investimento em causa perdida seja minimizado.

Capítulo V

Fatores Identificadores da Maturidade Organizacional para *Data Warehouse*

5.1. Introdução

O objetivo do trabalho desenvolvido por Sen (2006) foi identificar os fatores que influenciam a percepção da maturidade dos processos de *data warehouse* como ferramenta de apoio à decisão gerencial. O foco do artigo está na percepção gerencial dos Processos de *Data Warehouse* (DWP). O resultado da pesquisa exploratória indicou que alguns fatores têm impacto na maturidade dos DWP, como já fora percebido por diversos profissionais de Tecnologia da Informação (TI) – Sen (2006).

A elaboração do trabalho partiu da identificação do significado de processo: trata simplesmente de uma seqüência de

passos feita para um dado propósito. Assim, os autores levantaram que *data warehouse*, apesar de ser um processo não muito novo, não tem recebido a devida atenção como um processo. Percebeu-se que CMM não cobre os processos relativos à gerência de dados. Como a maioria das empresas segue padrões e práticas direcionadas ao desenvolvimento de *software*, não alcançam níveis de maturidades desejados em *data warehouse*. Assim, focaram no *data warehouse* como um processo e estudaram suas implicações de maturidade pelo desenvolvimento de um modelo similar ao CMM.

O segundo passo para elaboração do trabalho foi a identificação das tarefas pertinentes aos processos de *data warehouse*. Para tal, os autores conduziram uma sessão de *brainstorm*¹ com vários profissionais da área, de várias companhias tais como: Dell, Travelocity, BNSF, Railways, BP, e outras. Os participantes tinham de 5 a 7 anos de experiência. Percebeu-se que gerenciar os processos de DW é um desafio pois envolve a resolução criativa de problemas e a coordenação de equipes de diferentes áreas ou departamentos durante seu desenvolvimento e as operação de ETL (*Extract, Transform, Load* – Extração, Transformação e Carga). As atividades, ou tarefas identificadas foram:

- Compilação e Análise dos Requisitos de Negócios;
- Modelagem lógica dos dados;
- Mapeamento dos dados;
- Modelagem e implementação física;
- Projeto do ETL;
- Projeto dos relatórios e cubos;

¹ procedimento utilizado para relatar o maior número possível de idéias

- Planejamento do projeto;
- Gerenciamento da qualidade dos dados;
- Validação e testes;
- Arquivamento do DW;
- Backup e recuperação do DW;
- Gerenciamento de mudança;
- Determinação do ROI.

Posteriormente, os autores identificaram os *workflows* (fluxos de trabalho) relativos à *data warehouse* através do seqüenciamento das tarefas levantadas. Os *workflows* identificados e seus objetivos são listados abaixo:

- Modelagem de negócios: Lidar com os problemas de comunicação entre usuários;
- Requisitos: Focar em quais as funções o DW deve prover aos usuários;
- Análise e Projeto: Descrever as funções a serem desenvolvidas – o centro é a arquitetura do DW;
- Implementação: Focar na implementação do modelo projetado;
- Teste: Focar na verificação dos componentes desenvolvidos;
- Desenvolvimento: Produzir as versões do produto e entregar a tecnologia ao usuário final;
- Gerenciamento do Projeto: Balancear objetivos, gerenciar riscos e restrições;
- Gerenciamento de Configuração e Mudança: Controlar os vários artefatos produzidos no projeto do DW, inclusive versões;
- Ambiente: Fornece o ambiente apropriado à equipe de DW.

A tabela 5.1 apresenta o mapeamento entre as atividades e os *workflows* identificados por Sen (2006).

WorkFlow	Atividades
Modelagem de Negócios	Projeto do ETL
Requisitos	Compilação e Análise dos Requisitos de Negócios
Análise e Projeto	Modelagem lógica dos dados
Implementação	Mapeamento dos dados Modelagem e implementação física
Teste	Validação e testes
Desenvolvimento	Projeto dos relatórios e cubos
Gerenciamento de Projeto	Planejamento do projeto Gerenciamento da qualidade dos dados Determinação do ROI
Gerenciamento de Configuração e Mudança	Arquivamento do DW Gerenciamento de mudança
Ambiente	Backup e recuperação do DW

Tabela 5.1 – Mapeamento entre *workflows* e atividades levantados por Sen (2006)

5.2. Mapeamento dos Níveis de Maturidade

Reconhecendo o fato de que CMM não trata de maturidade de uma organização em relação à gerência de dados, Mullins (1997) mapeou níveis de maturidade com base nos níveis do CMM e em como os dados são controlados dentro da organização. Dataflux (2005) e Laney (2000) emitiram idéias similares. Marco (2002), em uma série de artigos, enfatizou a necessidade de ter um modelo de maturidade como o CMM para o processo de *data warehouse*. Enfatizando a necessidade de ter um modelo de maturidade como o CMM para *data warehouse*, ele aplicou CMM aos processos de DW e desenvolveu seis níveis de maturidade de DW, do nível 0 (não executado) ao nível 5 (continuamente melhorando).

Neste trabalho foi definido um modelo de maturidade de DWP baseado nas idéias de Mullins para a gerência de dados e Marco (2002) para armazenamento de dados – *data warehouse*. Este modelo foi desenvolvido seguindo os seguintes passos:

- Identificação das tarefas de DWP através de sessões de brainstorm;
- Associação das tarefas identificadas às tarefas de desenvolvimento de *software* constantes no CMM;
- Identificação e definição dos *workflows* – fluxos de trabalhos;
- Distribuição das tarefas identificadas nos devidos *workflows*.

Assim, os níveis de maturidade de DWP levantados por Sen (2006) e aplicados neste trabalho, foram:

Nível 1: Neste nível não se tem nenhum procedimento ou regulamento para gerenciar *datawarehouse's*. Geralmente os dados estão dispersos em diferentes arquivos e os banco de dados utilizam formatos diferentes. Neste nível têm-se um alto nível de redundância. Alterações são feitas sem antes fazer nenhuma análise de impacto destas alterações. A maioria dos *data mart's* existentes possuem redundância também. Os projetos tendem a ser mais caros, sendo que poucos têm sucesso e a maioria falham.

Nível 2: As organizações possuem políticas de gerenciamento de dados que dizem como e quando as estruturas destes dados são criadas, alteradas e gerenciadas. Apesar de possuir as políticas, estas não são institucionalizadas. Neste estágio, geralmente tem-se um DBA (administrador de banco de dados) designado. Algumas práticas padrões são feitas, tais como: gerenciamento de mudanças dos esquemas, monitoramento do desempenho e ajustes do banco de dados. Neste nível, algumas

organizações têm feito algumas atividades de *data warehouse*. Os processos e políticas existentes tendem a ser seguidos por grupos e não por toda a organização.

Nível 3: Organizações neste nível possuem políticas determinadas para tratamento dos dados. As melhores práticas de desenvolvimento, manutenção e operação do *data warehouse* são documentadas e usadas em toda empresa. As políticas de gerenciamento de dados tornam-se componentes chaves do ciclo de vida do desenvolvimento de aplicações. As políticas são executadas e testadas para garantir que os requisitos de qualidade estão sendo atendidos. Neste nível as organizações percebem a importância do gerenciamento de dados para o negócio e geralmente criam a função de AD – Administrador de Dados, além da função de DBA já existente. Normalmente existe uma boa interação entre as duas funções e utilização de ferramentas apropriadas de *data warehouse*. Neste nível, normalmente existem poucos *data mart's* independentes e a maioria dos projetos tende a ser de sucessos e não de falhas.

Nível 4: Uma organização neste nível insere um ambiente de metadados gerenciável. Isto permite à equipe de gerenciamento de dados, catalogar e manter os metadados como estruturas de dados da corporação. Neste nível inicia-se o processo de auditoria dos dados visando medir a qualidade dos mesmos. Todos os processos de DW existentes devem (é o objetivo do nível) ser mensuráveis. Técnicas estatísticas são usadas para analisar as medidas coletadas. Os projetos de DW tem sucesso persistente e as organizações podem prever desempenho futuro com razoável precisão.

Nível 5: Organizações neste nível utilizam as melhores práticas aprendidas nos demais níveis para melhorar continuamente seu acesso aos dados, a qualidade dos dados e o desempenho de seu

data warehouse. Existe um nível muito baixo de redundância em dados, processos ou tecnologias. A pouca redundância remanescente é documentada e compreendida pela organização. Os processos são alinhados com os objetivos estratégicos da empresa e existe uma tentativa de otimizar os investimentos em *data warehouse*.

Para levantar os fatores que influenciam a percepção da maturidade dos processos de *data warehouse*, os autores se basearam na percepção relatada por gerentes das organizações estudadas. De acordo com Sen (2006) são as percepções dos gerentes do ambiente externo que dirigem as ações organizacionais. Posteriormente, discutiu-se como a maturidade de DWP percebida pelos gerentes de TI podem influenciar decisões de investimento e iniciativas relacionadas a dados dentro de suas organizações. As características foram levantadas com base na literatura de *data warehouse* e de maturidade. Para um armazém de dados, os estágios da maturidade são definidos por um número de aspectos ou de características, incluindo: escopo, estrutura analítica, percepções executivas, tipos analíticos, administração, capital de giro, plataforma tecnológica, gerenciamento e administração de mudança.

As características, ou fatores, identificadas pelos autores são as constantes na tabela 5.2. Note que a tabela demonstra o mapeamento existente com os *workflows* identificados.

Fatores do DW	WorkFlow
Tipo de Arquitetura	Análise e Projeto
Tamanho do Data Warehouse	Requisitos Implementação Desenvolvimento
Cultura Analítica	Ambiente
Gerenciamento de Projeto	Gerenciamento de Projeto
Alinhamento da Arquitetura	Modelagem de Negócios Análise e Projeto
Qualidade dos Dados	Análise e Projeto Teste Gerenciamento de Configuração e Mudança
Prontidão Organizacional	Modelagem de Negócios
Folga Organizacional	Gerenciamento de Projeto
Gerenciamento de Mudanças	Gerenciamento de Configuração e Mudança

Tabela 5.2 – Relação entre fatores e *workflows*, levantados por Sen (2006)

Após o levantamento dos nove fatores influentes da maturidade dos processos de *data warehouse*, os autores conduziram um estudo de campo para validar a influência destes fatores na maturidade dos processos de *data warehouse*.

A pesquisa foi feita com base em questionário que se utilizou de múltiplos **itens** para medir a maioria dos fatores levantados.

Posteriormente, conduziu-se o envio destes questionários a uma lista de profissionais e diversas empresas e que continha uma breve descrição de cada nível de maturidade delineado por Sen (2006). Os respondentes indicavam qual o nível de maturidade acreditavam que sua organização estava.

Para validar os resultados obtidos, os autores testaram as escalas de medida para as propriedades de validade e confiabilidade. A validade mede o grau a que uma escala mede exatamente as variáveis sob investigação, e a confiabilidade mede a estabilidade da escala. A validade da variável é avaliada normalmente usando três tipos: índice, convergente, e diferenciada. A característica validade sendo satisfeita demonstra que todos os aspectos da variável estão sendo medidos. O conteúdo da validade foi estabelecido identificando os fatores indicadores e refinando-os através da completa análise da literatura.

Após a validação dos resultados, feita através de análises estatísticas, os autores concluíram que os fatores levantados auxiliam na percepção da maturidade dos processos de *data warehouse*.

Capítulo VI

Avaliação da Maturidade Organizacional para *Data Warehouse* e Estudo de Caso

6.1. Introdução

Data warehouses fornecem dados exatos e adequadamente gerenciados, permitindo às corporações reagir mais rapidamente às novas oportunidades ou declínios do mercado. Permitem a geração de dados integrados e históricos auxiliando os diretores decidirem embasados em fatos e não em intuições ou especulações, o que reduz a probabilidade de erros aumentando a velocidade na hora da decisão

Conhecer mais sobre essa tecnologia permitirá aos administradores descobrir novas maneiras de diferenciar sua empresa numa economia globalizada, deixando-os mais seguros para

definirem as metas e adotarem diferentes estratégias em sua organização, conseguindo assim visualizarem antes de seus concorrentes novos mercados e oportunidades atuando de maneiras diferentes conforme o perfil de seus consumidores. Pesquisas do IDC (1996) indicam que os resultados da utilização de *data warehouse* são irrefutáveis: como investimento, dão às organizações a capacidade de obter benefícios astronômicos, cumprindo muitas das promessas que sempre se fizeram com relação às tecnologias.

Diante da grande importância adquirida por esta ferramenta, a avaliação de sua maturidade, na busca pela melhoria contínua, torna-se uma ação de cunho altamente estratégico.

6.2. Fatores

O modelo proposto tem por base características macros de *data warehouses*, aqui definidos como Fatores Principais, extraído do nome da técnica de mineração de dados de análise por componente principal. Estas características são traços do *data warehouse* facilmente perceptíveis pelos membros da organização que lidam direta, ou indiretamente, com o *data warehouse*. Estes fatores, conforme levantado em Sen (2006), são:

- Alinhamento da Arquitetura: trata da infra-estrutura de TI para o armazém de dados. Deve ser capaz de compartilhar todos os dados necessários às várias aplicações OLAP e às aplicações de BI;
- Prontidão Organizacional: refere-se à habilidade de reconhecer o valor da informação (externa ou interna) nova, de assimilá-la, e de aplicá-la eficazmente para realizar benefícios econômicos; sugere-se ser crítico às inovações organizacionais. A prontidão

de uma organização reflete sua potencialidade de absorver informações a respeito das inovações apropriadas de modo que aquelas inovações possam ser aplicadas na sustentação das atividades operacionais ou estratégicas (reduzindo, se não completamente, eliminando, as barreiras do conhecimento). Sen (2006);

- Cultura de Decisão Analítica: refere-se à característica de estimular, incentivar e recompensar a tomada de decisão baseada em fatos. Tal cultura fornece um ambiente para tomada de decisão em uma organização. Sen (2006);
- Folga Organizacional: trata da robustês da infra-estrutura existente de TI para acomodar e alavancar a inovação tecnológica. É um fator que influencia a oportunidade do uso dentro do ambiente de trabalho Sen (2006);
- Qualidade dos Dados: trata da qualidade dos dados armazenados no *data warehouse*, ou seja, depende do processo de transformação e carga dos dados de origem para o armazém de dados. Para se aumentar esta qualidade, é necessário introduzir políticas de gerência de melhoria dos dados Sen (2006).
- Gerenciamento de Projetos: refere-se à habilidade da organização em gerenciar seu armazém de dados de maneira a manter prazos e custos, levantar possível problemas antecipadamente, e fornecer sustentação aos usuários durante todo o desenvolvimento e a execução do projeto;
- Gerenciamento de Mudanças: refere-se à habilidade do gerente do *data warehouse* em rastrear e controlar as mudanças à tempo, ou seja, de maneira que as mudanças não impactam

negativamente no desempenho e no resultado do *data warehouse*.

A identificação de cada um destes Fatores Principais é feita a partir de vários Fatores Discriminantes. O fator discriminante é a característica do *data warehouse* que a organização possui – em menor ou menor grau – e que, de acordo com a pesquisa de Sen (2006), influencia na maturidade dos processos de *data warehouse*.

A tabela 6.1 exibe a listagem dos Fatores Principais detalhadas nos Fatores Discriminantes. A tabela mostra também o grau de importância b_i (detalhado no item 6.3) que foi utilizado nos cálculos da maturidade fatores organizacionais indicativos de sucesso do *data warehouse*. Note que o Fator Gerenciamento de Mudança não se encontra nesta tabela, por não ter nenhum fator discriminante para identificá-lo, conforme constante na pesquisa de Sen (2006).

Fatores Principais	Fatores Discriminantes	Grau de Importância
Alinhamento da arquitetura		
	Existência de infra-estrutura robusta de TI – hardware, telecomunicações, redes, sistemas operacionais, dados, etc dentro da empresa	0.767
	Existência de um grande número de aplicações de sistemas de informações dentro da empresa que podem ser utilizados e beneficiados pelo DW	0.762
	Complexidade dos dados disponíveis e usados dentro da empresa	0.694
Prontidão Organizacional		
	Familiaridade, visão e entendimento do que o DW pode fazer pela empresa – pelos usuários chaves de DW	0.759
	Necessidade de extensivo treinamento dos usuários chaves de DW para desenvolver as habilidades necessárias para o entendimento e apropriado uso do DW.	0.407
	Não existência de barreira de conhecimento ao DW. Usuários chaves do DW terem conhecimentos técnicos necessários para explorar o DW	0.752
Cultura de Decisão Analítica		
	Confiança em fatos, gráficos e diagramas no momento da tomada de decisão	0.799
	Compartilhamento de informações pela organização	0.619

	Incentivo e recompensa de tomadas de decisões baseadas em fatos	0.772
	Requisitos de dados em formatos customizados e em níveis variados de detalhes	0.779
	Uso de relatórios atualizados para análises de tendência	0.779
Folga Organizacional		
	Recursos financeiros totalmente comprometidos com o DW	0.698
	Comprometimento de todos os recursos técnicos necessários ao DW, tais como: servidores de banco de dados, ferramentas de OLAP, ferramentas dirigidas ao usuário final, e outras	0.816
Qualidade dos Dados		
	Qualidade dos dados disponíveis	0.888
	Confiabilidade dos dados disponíveis	0.887
	Falta de acordos claros na definição de dados e nas regras de negócio da empresa neste momento	0.544
Gerenciamento de Projetos de DW		
	Lida com toda e qualquer resistência (política) que possa ocorrer durante a implementação do DW	0.835
	Gerencia as mudanças (e conseqüências) que possam ocorrer durante a implementação do DW	0.877
	Fornecer suporte a todo o processo de implementação do DW empresa	0.665

Tabela 6.1 – Grau de Importância dos Fatores Indicativos de Sucesso do *Data Warehouse*.
Fonte: adaptado de Sen (2006)

No intuito de ser uma ferramenta de apoio aos processos decisórios da organização, a proposta apresenta os fatores levantados agrupados por funcionalidade, ou seja, fatores indicativos de implantação ou de utilização. Assim, espera-se demonstrar à organização, qual tipo de fator discriminante está com mais baixa maturidade. Isso possibilita a uma organização que pretende ainda implantar o *data warehouse*, identificar se seu ambiente tecnológico está apto a ter o *data warehouse* instalado, antes mesmo de sua implantação, possibilitando correções e alterações previamente.

A tabela 6.2 mostra este agrupamento diferenciado

Fatores Discriminantes Indicativos de Tendência ao Sucesso do DW	
Implantação	Utilização
P – Familiaridade, visão e entendimento do que o DW pode fazer pela empresa – pelos usuários chaves de DW	A – Existência de infraestrutura robusta de TI – hardware, telecomunicações, redes, sistemas operacionais, dados, etc dentro da empresa
C – Confiança em fatos, gráficos e diagramas no momento da tomada de decisão	A – Existência de um grande número de aplicações de sistemas de informações dentro da empresa que podem ser utilizados e beneficiados pelo DW
C – Compartilhamento de informações pela organização	A – Complexidade dos dados disponíveis e usados dentro da empresa
C – Incentivo e recompensa de tomadas de decisões baseadas em fatos	P – Necessidade de extensivo treinamento dos usuários chaves de DW para desenvolver as habilidades necessárias para o entendimento e apropriado uso do DW.
C – Uso de relatórios atualizados para análises de tendência	P – Não existência de barreira de conhecimento ao DW. Usuários chaves do DW terem conhecimentos técnicos necessários para explorar o DW
	C – Requisitos de dados em formatos customizados e em níveis variados de detalhes
	F – Recursos financeiros totalmente comprometidos com o DW
	F – Comprometimento de todos os recursos técnicos necessários ao DW, tais como: servidores de banco de dados, ferramentas de OLAP, ferramentas dirigidas ao usuário final, e outras
	Q – Qualidade dos dados disponíveis
	Q – Confiabilidade dos dados disponíveis
	Q – Falta de acordos claros na definição de dados e nas regras de negócio da empresa neste momento
	G – Gerente lida com toda e qualquer resistência (política) que possa ocorrer durante a implementação do DW
	G – Gerente gerencia as mudanças (e consequências) que possam ocorrer durante a implementação do DW
	G – Fornece suporte a todo o processo de implementação do DW empresa

Tabela 6.2 – Fatores Agrupados Por Implantação x Utilização

6.3. Levantamento do grau de importância

O grau de importância, a ser utilizado nos cálculos desta proposta, advém do fator de carregamento levantado por Sen (2006). O fator de carregamento indica o quanto cada fator discriminante impacta na existência do Fator Principal no ambiente organizacional. Para levantar este fator, Sen (2006) conduziu uma análise multivariada (ou multidimensional) nas características levantadas, usando o nível de maturidade como um fator. As

características foram avaliadas através da escala de *Likert* de 7 pontos. Alguns indicadores foram expressados negativamente para minimizar problemas de respostas mecânicas. Depois as escalas de medidas foram testadas nas propriedades de validade e confiabilidade. A validade mede o grau a que uma escala mede exatamente as variáveis sob investigação, e a confiabilidade mede a estabilidade da escala. A validade da variável é avaliada normalmente usando três formulários de validade: índice, convergente, e diferenciada. A validade satisfeita avalia se todos os aspectos da variável estão sendo medidos. Estabeleceu-se o conteúdo da validade identificando os itens indicadores e refinando-os através da completa análise da literatura e pelo processo de pré-teste e de piloto, feitos durante a coleta das características, através de entrevistas.

Posteriormente, efetuou-se uma análise de fator conjunto de todos os itens multi-indicadores que medem os múltiplos construtores para poderem revelar se os itens carregados na pré-determinação teórica definem construtores e se havia qualquer carregamento cruzado de itens em outros construtores. Uma análise de fator do componente principal conjunta (de 20 indicadores que representam seis variáveis de pesquisa, excluindo o gerenciamento de mudança) com a técnica de rotação do fator varimax foi usada testando a validade convergente e diferenciada. Esta técnica testa se os indicadores teorizados convergem em construtores apropriados e diferenciam variáveis múltiplas com carregamento transversal mínimo entre fatores. O critério padrão do Valor de *Eigen* mais de 1.0, fator de carregamento maior que 0.4, e, como explicado, estruturas de fatores simples foram usados na análise. Neste ponto, alguns indicadores foram excluídos da análise por ter carregado em um fator não prognosticado pela pesquisa prévia. A confiabilidade da medida é avaliada normalmente através do *Alfa de Cronbach*, que

testa a consistência interna da escala. Um ponto inicial mínimo de 0.6 é considerado normalmente como adequado para os estudos exploratórios empíricos. Todas as variáveis tiveram valores que atestaram uma confiabilidade satisfatória.

6.4. Modelo Proposto para Avaliação da Maturidade Organizacional para *Data Warehouse*

A proposta considerou que a maturidade dos fatores indicativos de tendência ao sucesso do *data warehouse*, impactantes nos seus processos é avaliada a partir do grau de presença de determinadas características no ambiente organizacional. O fator de carregamento – pesos destas características, extraídos por Sen (2006), também é utilizado para avaliar esta maturidade de acordo como também fora utilizado em Ribeiro (2005) e Fleury (2006).

A proposta também considera o método proposto por Ribeiro (2005) na elaboração dos cálculos e calibrações de valores para obtenção desta maturidade.

Para implementação do modelo proposto utilizou-se a Teoria de Resposta ao Item com o modelo logístico de um parâmetro, que relaciona a probabilidade de acerto de um item P , seu índice de importância, ou peso, b_i e a habilidade θ da organização no fator em questão dada pela equação (3.1) reescrita abaixo:

$$P_i(\Phi) = \frac{1}{1 + \exp^{-1(\Phi - b_i)}} \quad (6.1)$$

- Interpretação do Parâmetro b

De acordo com Carneiro (2002), este parâmetro é medido na mesma escala da maturidade θ . No contexto da qualidade, ele

está associado ao grau de dificuldade de uma organização em implantar a prática da GQT: à medida que b_i cresce, aumenta o grau de dificuldade de implantação das práticas (o que significa que poucas organizações a implantam), e vice-versa. A correspondência entre o parâmetro b_i e a maturidade organizacional θ consiste na premissa de que as organizações que ainda não despertaram efetivamente para o modelo de gestão com base na GQT, isto é, com pouca maturidade, terão grandes dificuldades na implantação dessas práticas. Essa afirmação torna-se possível porque os parâmetros de dificuldade e maturidade estão na mesma escala. Para estimar o parâmetro b_i utilizamos os graus de importância descritos na Tabela 6.1. Porém, para trabalhar na mesma faixa da maturidade, normalizamos os valores, que variam de 0 a 100% para variar de -3 a 3, portanto, os novos valores para o nível de dificuldade das práticas ficou definido como:

$$b_i = (6 \times b_i' - 3) \quad (6.2)$$

onde:

$b_i =$ é o índice de dificuldade de implantação da i -ésima prática;

b_i' é o fator de carregamento (0 a 100%)

- Parâmetro de Maturidade Organizacional θ

O parâmetro θ representa a maturidade organizacional na implantação das práticas da GQT. Teoricamente este parâmetro pode assumir valores de $-\infty$ e $+\infty$. Para que mantivesse a mesma escala do nível de dificuldade b , assumiu-se novamente a variação de -3 a 3. Após tabular os dados, calculamos a média das respostas obtidas para cada item. Normalizando para a faixa requerida, o parâmetro de maturidade ficou definido como:

$$\theta_i = \frac{3\theta_i' - 6}{2} \quad (6.3)$$

onde:

θ_i é maturidade organizacional para aplicação da i-ésima prática;

θ_i' é a média das respostas obtidas na pesquisa (0 a 4)

Neste trabalho, a maturidade organizacional significa o quanto a empresa analisada possui de cada quesito.

- Cálculo da Probabilidade $P(\theta)$

Aplicando então a TRI utilizando o Modelo de Rasch, alcançamos a probabilidade que um fator discriminante tem de, dado seu grau de importância θ_i , impactar na maturidade do Fator Principal.

- 6.2.4. Cálculo da Nota

Para obter a pontuação da empresa partiremos da definição de Esperança, que é uma variável aleatória que fornece a média de todos os valores que esperaríamos, ou seja, é o valor médio que resultaria se observássemos muitas vezes uma variável aleatória, também chamado de Valor Esperado. No caso deste estudo, devemos calcular a Esperança de atendimento a um Fator Discriminante com grau de importância b_i , sendo $P(\theta_i)$ a probabilidade de alcance do fator principal de maturidade θ_i :

$$E(x) = b_i \times P(\theta_i) \quad (6.4)$$

Como se deseja obter o Valor Esperado do alcance de n fatores, deve-se dividir a somatória dos valores esperados de todos os itens pela somatória das importâncias de todos os itens. Assim, a nota de cada um dos Fatores Principais é calculada por:

$$\pi = \frac{\sum_{i=1}^n E_j(x)}{\sum_{i=1}^n P(\theta_i)} = \frac{\sum_{i=1}^n b_i \cdot P(\theta_i)}{\sum_{i=1}^n P(\theta_i)} \quad (6.5)$$

Posteriormente, efetuou-se o cálculo da maturidade da organização, especificamente neste trabalho, com relação aos processos de *data warehouse* como ferramenta de apoio à decisão gerencial. Este cálculo é feito através da mesma expressão, porém considerando todos os fatores discriminantes sem o agrupamento em fatores principais.

O valor da maturidade encontrado deve, então, ser aplicado na expressão abaixo para que possamos inferir uma comparação como padrão americano – que é o padrão constante na pesquisa de Sen (2006).

O trabalho considera os valores de b_i (grau de importância) tendendo a 50%, o que seria a média. Isto significa que, assumiu-se que o grau de importância pode ser considerado como uma média dos valores das organizações. Para saber o quão uma organização brasileira é mais ou menos madura que a média das empresas americanas – compara-se com estas empresas porque o grau de importância é retirado da pesquisa de Sen (2006) feita nos E.U.A –, aplica-se o valor da maturidade encontrado na seguinte expressão:

$$P' = \frac{P - 50\%}{50\%} \quad (6.6)$$

Assim é possível avaliar a maturidade relativa à maturidade da média daquele país. Para se efetuar esta avaliação, considera-se que os valores denominados graus de importância,

advindos da pesquisa feita nos E.U.A., corresponde à média nacional daquele país.

6.5. Estudo de Caso

No intuito de validar o modelo proposto, um estudo de caso foi proposto e desenvolvido em uma grande empresa governamental do Estado de Goiás que atua na administração e execução de políticas tributárias, captação de recursos, gestão das finanças públicas e fomento ao desenvolvimento social e econômico. Possui 2.473 servidores efetivos, e foi fundada em janeiro de 1.945. A missão da empresa é: Administração e execução de políticas tributárias, captação de recursos, gestão das finanças públicas e fomento ao desenvolvimento social e econômico.

A empresa foi escolhida através da análise de características administrativas – ser uma das maiores empresas do governo estadual, possuir *data warehouse* instalado há aproximadamente 5 anos – e também pela maior facilidade de acesso.

Inicialmente, foi feita uma entrevista com o responsável pela Coordenação de Informática da empresa informando-o sobre o procedimento científico a ser desenvolvido e solicitando sua permissão para efetuar a pesquisa. Em seguida, o formulário – constante no apêndice – foi enviado à analista responsável pelo *data warehouse* da organização.

Posteriormente, de posse das respostas, apresentadas no apêndice, foi iniciada a etapa de calibração dos valores e elaboração dos cálculos conforme as expressões apresentadas anteriormente.

A elaboração dos cálculos necessários à proposta, foi feita com o auxílio de uma planilha eletrônica, constante na figura 3. Os resultados da maturidade absoluta e da maturidade relativa dos fatores organizacionais indicativos de sucesso do *data warehouse* são mostrados na tabela 6.2.

	Maturidade (%)	Maturidade Relativa (%)
Da Organização - GERAL	65,59%	31,18%
Alinhamento da Arquitetura	77,56%	55,12%
Prontidão Organizacional	70,88%	41,76%
Cultura de Decisão Analítica	59,77%	19,55%
Folga Organizacional	47,70%	-4,61%
Qualidade dos Dados	63,89%	27,77%
Gerenciamento de Projetos	51,49%	2,98%
Gerenciamento de Mudanças	26,89%	-46,21%

Tabela 6.2 – Maturidade Absoluta e Maturidade Relativa

A figura 6.2 mostra o gráfico da maturidade absoluta da organização e de seus fatores principais, enquanto a figura 6.3 exibe um gráfico com os valores da maturidade relativa (da organização e de seus fatores principais) que se trata do valor percentual comparativo com o padrão das empresas avaliadas por Sen (2006) a partir das médias dos fatores extraídos em sua pesquisa.

	Entre na coluna abaixo os Loading Factors de cada Item do Fator 1 do PAPER	Aqui cada Loading Factor é transformado para a escala de -3 a 3	Entre na coluna abaixo a média dos valores obtidos no questionário para cada Item do Fator 1	Aqui cada Habilidade é transformado para a escala de -3 a 3	Aqui é calculado a probabilidade de cada ITEM ser atendido dado a habilidade obtida no questionário	Aqui é calculado o Valor Esperado de cada ITEM dado a habilidade obtida no questionário
	Índice de dificuldade (b)		Habilidade (q)		Cálculo da Esperança	
1 - Arquitetura	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b
Infra-estrutura	0,77	1,60	7,00	3,00	0,80	1,28
Número de aplicações	0,76	1,57	7,00	3,00	0,81	1,27
Dados complexos	0,69	1,16	6,00	2,00	0,70	0,81
Somatório de b =		4,34			Somatório da Esperança =	3,36
		MATURIDADE :		0,78		
	Índice de dificuldade (b)		Habilidade (q)		Cálculo da Esperança	
2 - Prontidão	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b
Familiaridade	0,76	1,55	6,00	2,00	0,61	0,95
Treinamento	0,41	-0,56	2,00	-2,00	0,19	-0,11
Barreira de conhecimento	0,75	1,51	6,00	2,00	0,62	0,94
Somatório de b =		2,51			Somatório da Esperança =	1,78
		MATURIDADE :		0,71		
	Índice de dificuldade (b)		Habilidade (q)		Cálculo da Esperança	
3 - Cultura Decisão	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b
Confiança em fatos	0,80	1,79	7,00	3,00	0,77	1,38
Compartilhamento de Informações	0,62	0,71	5,00	1,00	0,57	0,41
Incentivo e recompensa	0,77	1,63	2,00	-2,00	0,03	0,04
Dados formatados	0,78	1,67	7,00	3,00	0,79	1,32

Uso de relatórios p/ análises	0,78	1,67	7,00	3,00	0,79	1,32
Somatório de b =		7,49			Somatório da Esperança =	4,48
			MATURIDADE :		0,60	
4 - Folga Organizacional	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b
Recursos comprometidos	0,70	1,19	2,00	-2,00	0,04	0,05
Comprometimento de todos os recursos técnicos	0,82	1,90	7,00	3,00	0,75	1,42
Somatório de b =		3,08			Somatório da Esperança =	1,47
			MATURIDADE :		0,48	
5 - Qualidade dos Dados	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b
Alta qualidade dos dados	0,89	2,33	7,00	3,00	0,66	1,54
Confiabilidade dos dados	0,89	2,32	7,00	3,00	0,66	1,54
Falta de acordos e regras	0,54	0,26	3,00	-1,00	0,22	0,06
Somatório de b =		4,91			Somatório da Esperança =	3,14
			MATURIDADE :		0,64	
6 - Gerenciamento de Projetos	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b
Lida com resistências	0,84	2,01	6,00	2,00	0,50	1,00
Gerencia mudanças	0,88	2,26	6,00	2,00	0,43	0,98
Fornecer suporte ao processo	0,67	0,99	6,00	2,00	0,73	0,73
Somatório de b =		5,26			Somatório da Esperança =	2,71
			MATURIDADE :		0,51	
7 - Gerenciamento de Mudanças	Grau de Importância	Conversão para b (-3 a 3)	Respostas (1 a 7)	Conversão para q (-3 a 3)	P(q)	E(q) = P(q) x b

Gerenciamento efetivo	0,00	-3,00	0,00	-4,00	0,27	-0,81
	Somatório de b =				Somatório da Esperança =	
		-3,00				-0,81
			MATURIDADE :			0,27

Figura 6.1 – Planilha eletrônica de auxílio na elaboração dos cálculos

MATURIDADE GERAL = SOMA (TODAS AS ESPERANÇAS) / SOMA (TODAS AS DIFICULDADES)			
MATURIDADE GERAL:	0,66	65,59%	
MATURIDADE REFERENCIADA:	0,31	31,18%	

	Maturidade	Maturidade (%)	Maturidade Referenciada	Maturidade Referenciada (%)
GERAL	0,66	65,59%	0,31	31,18%
Alinhamento da Arquitetura	0,78	77,56%	0,55	55,12%
Prontidão Organizacional	0,71	70,88%	0,42	41,76%
Cultura de Decisão Analítica	0,60	59,77%	0,20	19,55%
Folga Organizacional	0,48	47,70%	-0,05	-4,61%
Qualidade dos Dados	0,64	63,89%	0,28	27,77%
Gerenciamento de Projetos	0,51	51,49%	0,03	2,98%
Gerenciamento de Mudanças	0,27	26,89%	-0,46	-46,21%

Figura 6.1 – Planilha eletrônica de auxílio na elaboração dos cálculos – Segunda Parte

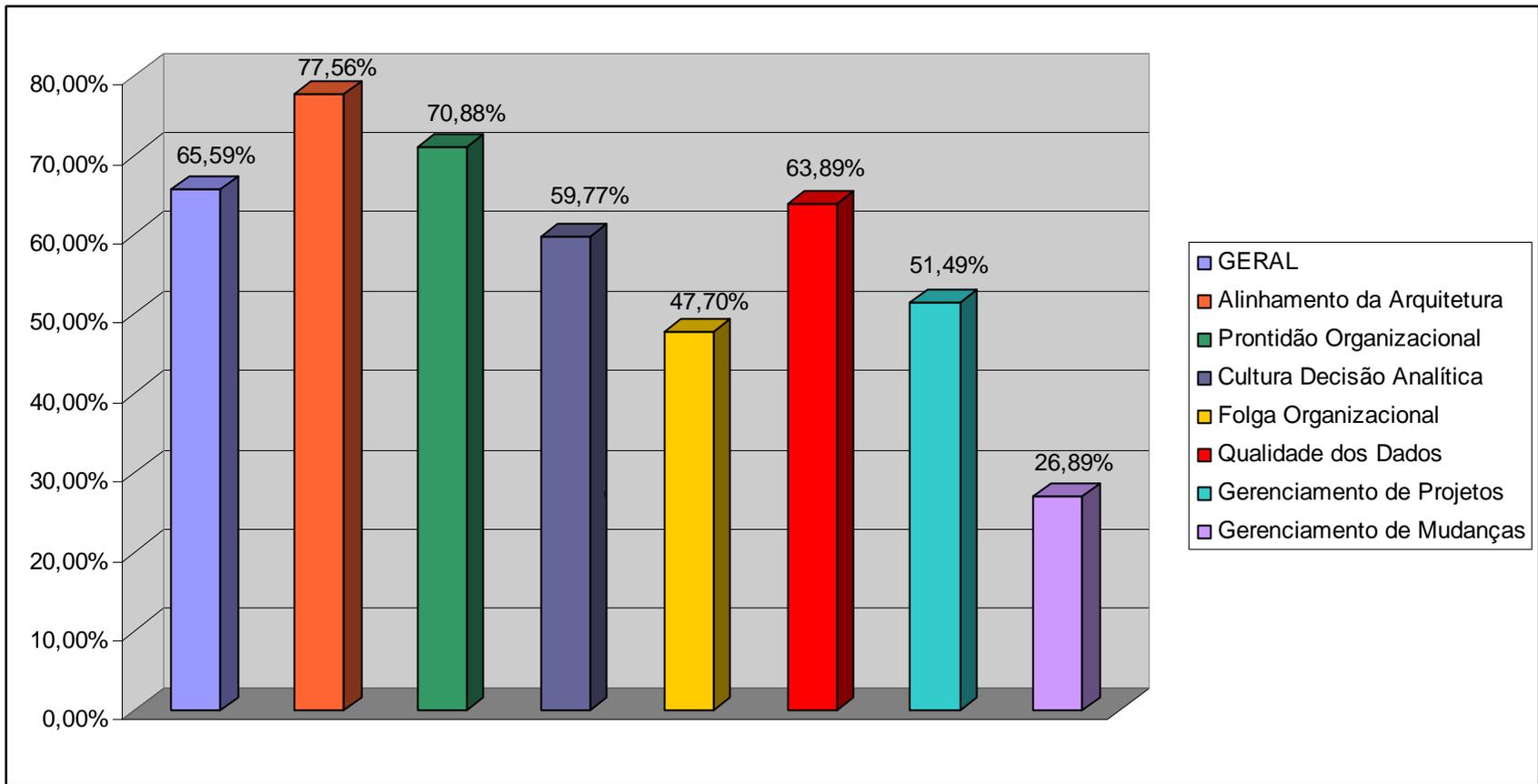


Figura 6.2 – Valor Esperado – Maturidade Absoluta

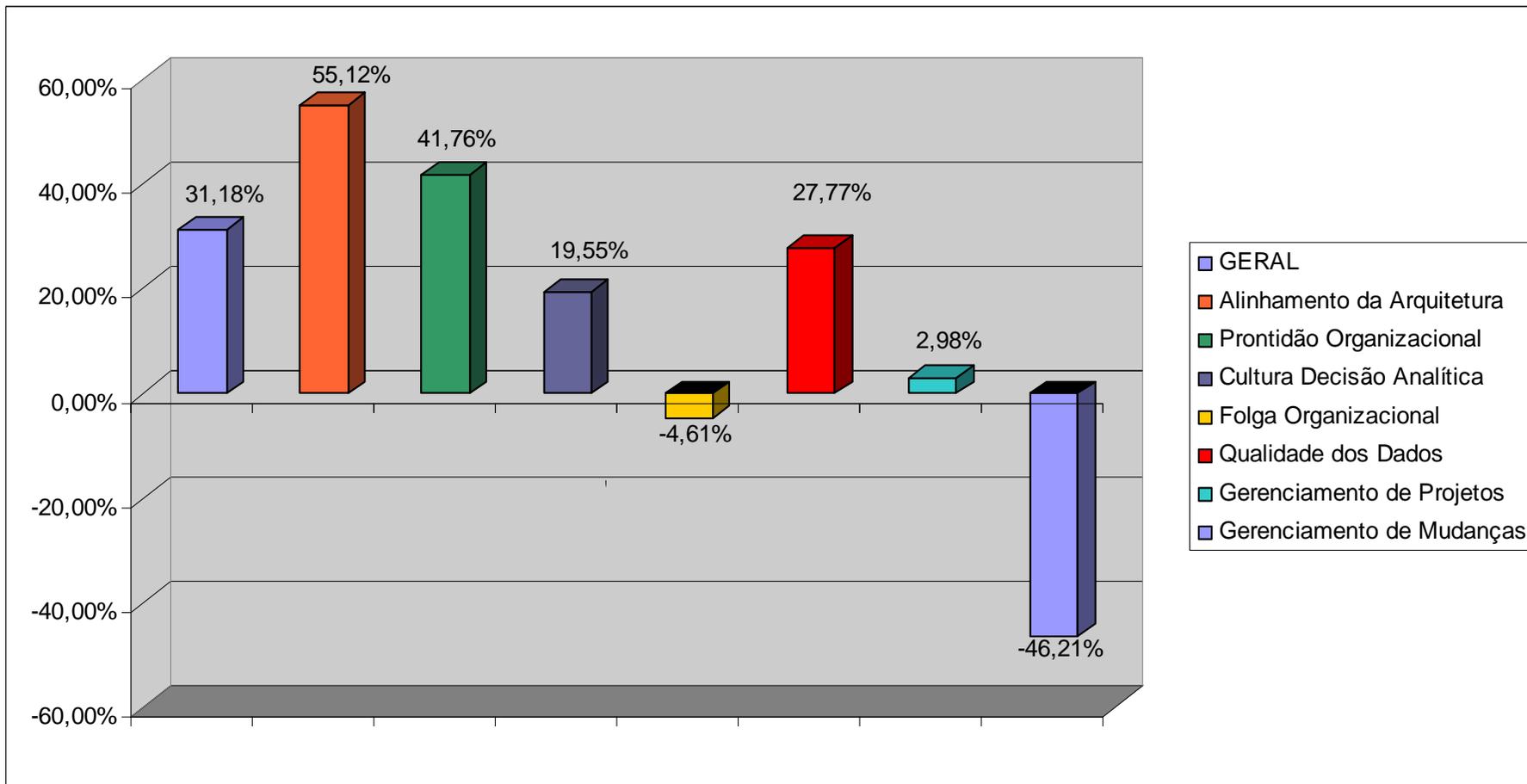


Figura 6.3 – Valor Esperado – Maturidade Relativa

6.1.1 Aplicação do Modelo e Análises dos Resultados

Os resultados advindos dos cálculos nos forneceram um valor da maturidade da organização de **65,59%**, enquanto o valor encontrado para a maturidade relativa foi de **31,18%**.

A análise dos resultados calculados nos indica que a organização estudada está 31,18% **mais** madura que o padrão de Sen (2006). Em termos mais amplos, pode-se considerar que a empresa está dentro dos padrões americanos de maturidade de *data warehouse*.

Abalizamos a seguir cada um dos Fatores individualmente:

Fator 1 - Alinhamento da Arquitetura

Como percebe-se ao observar a figura 6.1, o valor da maturidade relativa para o fator Alinhamento da Arquitetura da organização pesquisada é de 55,12%. A análise do resultado deste fator nos mostra que a organização estudada possui uma infraestrutura TI robusta o suficiente para fornecer uma boa estrutura para o *data warehouse*, que existem aplicações de sistemas de informações que o utilizam e se beneficiam do mesmo. Mostra ainda que os dados disponíveis e utilizados são, provavelmente, complexos. Para que a organização mantenha uma boa maturidade neste fator é necessário que se mantenha a boa infraestrutura de TI e a disponibilidade dos dados.

Analisando este valor, observamos que no fator Alinhamento da Arquitetura, a organização encontra-se mais madura que o padrão das empresas americanas.

Conforme levantado por Sen (2006) em sua pesquisa, a maioria das empresas com um maior grau de maturidade de processos de *data warehouse* possui uma maior integração de sua arquitetura de TI – que engloba a arquitetura dos dados. Estas empresas também costumam ter um alinhamento entre a estratégia da organização e a arquitetura de TI.

Este fator tem uma correspondência com a KPA – área de processo chave – **Gestão de Alteração de Tecnologia** do CMM, tendo em vista que este fator trata da existência de uma infraestrutura de tecnologia robusta e esta KPA se refere à identificação de novas tecnologias que possam trazer vantagens estratégicas para a organização.

Assim, fica claro que o modelo auxilia na tomada de decisão baseada na estratégia da organização. Isto significa que, ao identificar uma fraca maturidade no fator Alinhamento da Arquitetura, a organização deve planejar alterações da arquitetura de TI fazendo-a alinhada com a estratégia da organização.

Fator 2 - Prontidão Organizacional

Pelos valores mostrados na tabela 6.2, a organização pesquisada está 41,76% mais madura que o padrão das empresas americanas. A análise do resultado deste fator indica que os usuários da organização avaliada possui familiaridade, visão e entendimento do que o *data warehouse* pode fazer pela empresa. Indica também que o treinamento dos usuários chaves do *data warehouse* é necessário para manter a atualização das novas tecnologias porém não urgente ou essencial. Esta análise mostra também que não existem grandes barreiras ao *data warehouse*, pois os usuários

chaves possuem conhecimentos técnicos necessários para explorar as potencialidades do *data warehouse*.

Para que a organização mantenha uma boa maturidade neste fator é necessário que se mantenha a boa infra-estrutura de TI e a disponibilidade dos dados.

No Fator Prontidão Organizacional, um baixo valor demonstra uma necessidade de executar treinamentos tanto à equipe de usuários quanto à equipe de técnicos. Isto pode ser corroborado com as explicações dada por Sen (2006) em sua pesquisa, onde ele cita que a familiaridade dos membros da organização, a compreensão a cerca do que o *data warehouse* pode fazer por eles e pela empresa, é de suma importância para que o mesmo tenha o sucesso planejado e esperado.

Sergio Lozinsky diz em seu artigo "*Prontidão: um atributo-chave das empresas competitivas*", que prontidão organizacional significa manter um processo permanente de "investigação" de novas soluções de mercado – tecnologias, parcerias estratégicas, técnicas e métodos – que podem agregar valor a produtos ou serviços, ou mesmo criar novos negócios. Saber "filtrar" aquelas soluções que melhor suportam as estratégias da empresa" - Sen (2006).

De acordo com Sen (2006) prontidão significa encontrar os melhores meios para crescer de forma sustentável, e introduzir controles que privilegiem a eficiência operacional da organização, e não se limitem a uma atividade de policiamento dos procedimentos. É desenhar um modelo de organização que tem capacidade de absorver possíveis riscos, gerando uma empresa mais preparada posteriormente.

Este fator tem uma correspondência direta com a KPA – área de processo chave – “**Programa de Treinamento**” do CMM, pois para garantir a prontidão organizacional, é necessário que a empresa mantenha programas tais como seminários, fóruns e outros treinamentos que visem mostrar aos gerentes e executivos o real valor do *Data Warehouse* para que os mesmos o utilizem de forma eficaz para conseguir os objetivos e metas planejados.

Assim, a utilização do modelo de maturidade melhora a tomada de decisão consciente da organização – baseada em fatos, tendo em vista a visualização de possíveis deficiências relacionadas à prontidão organizacional. O modelo, ao mostrar estas deficiências, leva a organização a aumentar esta prontidão para, então, redirecionar suas operações para ocorrências de mercado que não podem ser previstas.

Fator 3 – Cultura de Decisão Analítica

O valor constante na tabela 6.2 para o fator Cultura de Decisão Analítica nos indica que a organização estudada está 19,55% mais madura que o padrão americano. A análise deste resultado mostra que os usuários, no momento da tomada de decisão, confiam em fatos, gráficos e diagramas gerados pelo *data warehouse*. Indica ainda que há um bom compartilhamento de informações e incentivo e recompensa de tomadas de decisões baseadas em fatos. Indica também que os requisitos de dados estão em formatos customizados e detalhados. Há uma boa utilização de relatórios atualizados para análises de tendência.

Para manter uma boa maturidade neste fator, a organização deve continuar incentivando a tomada de decisão baseada em fatos.

A cultura de decisão analítica corresponde à tendência da organização de tomar decisões baseadas em fatos. Claudia Hazan HAZAN,2004 corrobora a necessidade de se ter os fatos como base para a tomada de decisão em todos os níveis da organização: "... é a análise de fatos e dados gerados em cada um de seus processos, bem como os obtidos externamente, incluindo-se os referenciais comparativos pertinentes. O conhecimento adquirido por meio das informações é retido pela organização para que funcione de maneira mais ágil e independente. Para o processo de tomada de decisões ser efetivo, a organização deve dispor de sistemas estruturados de informação adequados ao seu negócio e desenvolver formas de obtenção e uso sistemático de informações comparativas."

A tomada de decisão mostra-se como mais um passo importante a ser seguido por uma organização que pretende atingir o nível de excelência de seus processos. Para que esta tomada de decisão seja realmente baseada em fatos, o auxílio de uma ferramenta de *BI*, como o *data warehouse*, é de vital importância.

Este fator se relaciona com a KPA "**Foco nos Processos da Organização**" do CMM, pois a decisão analítica depende da responsabilidade organizacional pelas atividades dos processos do *data warehouse*. Este relacionamento existe porque a cultura de decisão analítica é voltada aos fatos, enquanto, os processos da organização fornece parte destes fatos.

O modelo proposto demonstra o nível de maturidade deste fator dentro de uma organização, indicando assim, a necessidades de possíveis melhorias. Um ambiente analítico torna-se um ponto de alavanca para que as organizações identifiquem, planejem, e julguem o sucesso da inovação, conforme citado por ECKERSON,2004.

Fator 4 – Folga Organizacional

Como pode-se observar na tabela 6.2, o valor da maturidade relativa para a Folga Organizacional é de -4,61%. A análise deste número diferencia-se dos demais devido ao índice ser negativo. A análise do resultado deste fator nos mostra que a organização avaliada não está disponibilizando a ideal folga organizacional para permitir reservas a serem utilizadas no momento oportuno. A análise indica ainda que os recursos financeiros estão totalmente – ou quase totalmente – comprometidos com o *data warehouse*, assim como os recursos técnicos, tais como: servidores de banco de dados, ferramentas de OLAP e ferramentas dirigidas ao usuário final.

Para que a organização eleve a maturidade deste fator é necessário que a mesma disponibilize recursos técnicos que tenham uma folga adequada para ser utilizada em momentos críticos. Faz-se necessário ainda que a mesma aumente os recursos financeiros do *data warehouse* não permitindo que tais valores fiquem totalmente comprometidos, é necessário que deixe-se uma folga emergencial a ser utilizada em momentos de extrema necessidade.

O conceito de folga organizacional dado por Hirschmann Hirschmann, 1970 é o de “reservas potenciais que podem dispor as organizações, utilizadas abaixo da sua capacidade em condições normais, mas que podem ser mobilizadas para enfrentar situações imprevisíveis do mercado”. Este conceito demonstra claramente que, uma organização com baixa maturidade neste fator necessita incrementar seus recursos técnicos, para que as reservas não estejam indisponíveis em um provável momento de urgência.

A organização estudada pode estar disponibilizando quase sua totalidade de recursos financeiros e técnicos, sendo necessário planejar melhor estes recursos.

Aumentar o nível de maturidade exige um investimento considerável de recursos financeiros, pessoal, tempo – para desenvolver, documentar, administrar, e executar vários procedimentos formalizados –, ferramentas de projeto e em tecnologias (por exemplo, servidores de banco de dados, ferramentas de OLAP, etc.), ou ainda treinar a equipe de TI existente ou recrutar uma nova equipe. Estas necessidades, ao surgirem repentinamente, exigem uma disponibilização imediata de recurso, e a ideal folga organizacional pode possibilitar uma reação mais rápida a uma demanda de mercado.

O fator “Folga Organizacional” se relaciona com a KPA “**Foco nos Processos da Organização**” do CMM, porque ao estabelecer a responsabilidade organizacional pelas atividades dos seus processos, a empresa deve determinar também os recursos que devem ficar disponíveis para garantir a folga organizacional adequada.

Assim, percebe-se que a aplicação do modelo, ao perceber deficiências neste fator, aumenta as possibilidades da organização aumentar a velocidade da sua *tomada de decisão*, devido o processo reativo ao mercado ser mais rápido.

Fator 5 – Qualidade dos Dados

O valor constante na tabela 6.2 para o fator Qualidade dos Dados nos mostra que a organização pesquisada está 27,77% mais madura que o padrão americano. A análise do resultado deste fator mostra que a organização estudada possui uma boa qualidade dos

dados no data warehouse, ou seja, a carga e transformação dos dados são atividades que estão sendo executadas corretamente. A análise indica ainda que existe uma confiabilidade nos dados disponíveis e claros acordos ao definir regras de negócio da empresa.

A empresa pode melhorar a maturidade deste fator através da melhoria contínua dos dados do *data warehouse*. Esta melhoria pode ser alcançada através de ideal e correto transferência dos dados dos sistemas origem para a base do *data warehouse*.

De acordo com Mullins (1997) podem haver diferentes níveis de qualidade dos dados em uma organização, variando de alguma consciência que estes problemas afetam a tomada de decisão do negócio à concretização de que a informação é um recurso da empresa e que a qualidade dos dados é uma iniciativa estratégica e contínua da empresa como demonstrado pelo ROI.

Diante das exposições de Mullins (1997) percebe-se a grande importância da qualidade dos dados armazenados – através dos RDBMS – para a obtenção de um *data warehouse* maduro.

Um dos maiores impactos da baixa qualidade dos dados, dentre outros, é a demora na entrega de dados para tomadores de decisão, o que pode ocasionar perda de clientes e outras deficiências organizacionais.

Organizações com baixa maturidade no fator Qualidade dos Dados possuem um alto risco de incorrer em erros, ao se basearem em dados “falseados” para a tomada de decisão.

Este fator tem correspondência com as KPA's “**Garantia da Qualidade de Software**” e “**Gestão da Qualidade de Software**” do CMM que tratam de prover a equipe com a introspecção objetiva em

processos e de obter metas específicas de qualidade, respectivamente.

Assim, o modelo proposto, ao identificar a baixa maturidade deste fator, mostra à organização que a mesma pode estar se baseando em valores parciais ou não-verdadeiros e que é necessário desenvolver um plano de ação para correção destes dados para que a tomada de decisão seja baseada em fatos realmente reais.

Fator 6 – Gerenciamento de Projetos

Como pode-se observar nos gráficos e tabela apresentados, o valor da maturidade relativa da organização no fator Gerenciamento de Projetos é de 2,98%. Este número demonstra que a organização está praticamente no mesmo nível de maturidade encontrado na maioria das empresas americanas.

A análise dos resultados deste fator mostra claramente a facilidade do gerente em lidar com eventuais resistências que possam ocorrer durante a implementação e utilização do *data warehouse* e em fornecer todo o suporte necessário ao processo de implementação do *data warehouse* na organização. Para se manter este nível de maturidade, ou ainda aumentar tal nível, a organização deve sempre estar focada em treinamento do gerente nas melhores práticas, tais como as difundidas pelo Instituto de Gerenciamento de Projetos – PMI.

O gerenciamento tradicional de projetos permitiu a realização de ações impressionantes, como a ida do homem à Lua, a construção de plataformas de petróleo no mar do norte, a construção do Túnel do Canal da Mancha entre outras.

O domínio dos conhecimentos em gerenciamento de projetos é extremamente útil, porém não é suficiente para garantir resultados satisfatórios, pois ainda ocorrem muitos problemas de atraso, de custos e de escopo. Essa realidade é freqüentemente constatada por várias organizações. Porém, as mesmas organizações também verificam a melhoria observada após a implantação de metodologia de gerenciamento de projeto. Esta melhoria é percebida na redução de custos, atrasos e na diminuição de alterações de escopo.

O nível de maturidade do *data warehouse* é também uma função de como os projetos são controlados. Para se ter um processo de *data warehouse* maduro, é importante ter um bom desempenho da gerência superior, tal como fins de prazos críticos, custos dentro do orçamento, suporte às mais importantes funcionalidades e outros. Entretanto, para alcançar estes objetivos, é extremamente crítico que o processo de gerenciamento de projetos seja planejado o mais cedo possível porém, permitindo flexibilidade.

A gerência de projeto de *data warehouse* é uma tarefa difícil. Conforme observado nas pesquisas de Sen (2006), nas organizações com níveis mais baixos de maturidade, os projetos grandes e complexos são empreendidos sem uma consciência de seu impacto, tendo resultados tais como: cancelamento do projeto, estouros de orçamento, ou falha do *data warehouse*. Porém, em organizações com níveis mais elevados de maturidade, os gerentes de projeto especificam, medem, analisam, e otimizam os objetivos dos processos para entregar projetos bem sucedidos.

O modelo proposto permite a identificação de falhas no processo de gerenciamento do projeto, ao exibir um nível de maturidade baixo neste fator. A partir desta identificação, a

organização tem consciência da necessidade de pesquisar melhores práticas e metodologias para implantá-las. Assim, após a mudança de estratégia do gerenciamento de projetos, a organização pode alcançar o sucesso almejado.

Este fator tem correspondência com as KPA's "**Gestão Quantitativa do Processo**" e "**Planejamento de Projeto de Software**" do CMM, pois o gerenciamento do projeto é área responsável pelo controle do desempenho do processo, e ainda por planejar as atividades de todo o projeto.

Fator 7 – Gerenciamento de Mudanças

No fator Gerenciamento de Mudanças, a maturidade relativa da organização pesquisada é de -46,21%. Aqui tem-se o ponto problemático mais crítico da organização. O índice negativo indica uma maturidade menor que o padrão das empresas americanas. O valor alto -46,21 indica que esta diferença é bastante acentuada.

A análise do resultado deste fator mostra que o gerente do *data warehouse* está com dificuldades em lidar e gerenciar as mudanças – e suas conseqüências – que ocorrem durante a implementação e utilização do *data warehouse*. Como a maturidade do fator Gerenciamento de Projetos não é negativa, nem tampouco baixa, provavelmente a organização não estipulou um sistema de gerenciamento de mudanças, ou seja, a organização permite que os usuários do *data warehouse* façam as solicitações a qualquer momento e exige que as mesmas sejam cumpridas o mais rápido possível.

Para que a organização eleve a maturidade deste fator, é necessário que a mesma instale um sistema de gerenciamento de mudanças e o divulgue na organização. Isso faz com que os usuários entendam que toda e qualquer mudança terá um prazo para ter a viabilidade analisada e para ser efetuada, caso seja aprovada. O gerente, por si só, não tem autonomia para efetuar uma grande mudança neste aspecto, é necessário o total apoio da organização para implantar um sistema que realmente permita a adequada análise de viabilidade de mudanças e, conseqüentemente, um aumento do tempo de atendimento de tais mudanças.

O gerenciamento de mudanças é mais complicado quando o formato dos dados de origem é empurrado para o armazém, em vez de ser controlado por administradores de *data warehouse*. Conseqüentemente, controlar as mudanças e rastreá-las no tempo - controle efetivo e gerenciamento de configuração - é uma tarefa muito difícil.

De acordo com Eckerson, o gerenciamento de mudanças é uma característica chave da maturidade e deve ter um forte programa de administração para guiar o desenvolvimento e a expansão do *data warehouse*. Enquanto as mudanças são feitas "a toda hora e de qualquer modo" por organizações de nível 1 em resposta às requisições, as organizações de nível 5 têm políticas de gerenciamento de mudanças estritas ECKERSON,2004.

Assim, é facilmente perceptível que a aplicação do modelo proposto permite identificar necessidades de implantação de políticas de gerenciamento de mudanças, ao identificar o nível de maturidade da organização, pois processos de dw mais maduros são suportados por habilidades de gerenciamento de mudanças mais elevadas.

Este fator tem correspondência com a KPA “**Acompanhamento e Supervisão de Projeto de Software**” do CMM, pois o gerenciamento de mudanças só é efetivo se houver acompanhamento e supervisão de todo o processo do projeto de *data warehouse*. Este relacionamento se deve porque esta área de processo fornece insumos para o correto e ideal gerenciamento de mudanças.

6.6. Conclusões do Estudo de Caso

O estudo de caso realizado demonstrou que a organização pesquisada possui uma maturidade de *data warehouse* equivalente à média das organizações americanas. Pode-se ter esta afirmação a partir do valor encontrado para a maturidade relativa: **31,18%**.

Analisando os resultados do estudo de caso, percebe-se que a organização estudada está 31,18% **mais** madura que o padrão americano.

Assim, pode-se afirmar que o modelo proposto demonstra claramente que a maturidade de um “Fator Principal” tem impacto na nota da maturidade da organização. Isto se deve ao fato de que a maturidade da organização é feita através do somatório de todos os fatores discriminantes, e, o somatório de alguns destes fatores discriminantes nos dá a nota do “Fator Principal”.

A grande vantagem do modelo proposto - composto de dois diferentes tipos de fatores - é que além da organização ter conhecimento da sua maturidade organizacional dos processos de *data warehouse*, é também possível saber a característica do *data warehouse* onde ocorrem as maiores deficiências. Assim, é possível indicar a característica que necessita de mais urgência em correções,

adaptações ou alterações. Ou seja, o modelo ajuda a tomada de decisão das organizações, pois indica qual a característica – fator principal – está com o mais baixo índice de maturidade

Capítulo VII

Conclusão

7.1 Considerações Finais

O intuito deste capítulo é apresentar as conclusões obtidas após a revisão da bibliografia e a análise dos resultados pós estudo de caso.

O principal objetivo de uma organização, ao implantar um *data warehouse* é prover suporte à tomada de decisão de gerentes executivos e de negócio. A verdadeira, e valiosa, saída de um *data warehouse* é o conjunto de decisões de negócio que são feitas baseadas nos dados apresentados por esta ferramenta. Estas decisões podem ser úteis como subsídios à equipe de *data warehouse* na justificativa dos dispêndios financeiros necessários à adequada implantação e manutenção de um *data warehouse*.

Apesar do forte crescimento do mercado de *data warehouse* nos últimos anos, pesquisas do Garther Group mostram

que aproximadamente um quarto destas iniciativas falham. Dentre as organizações que investem nesta ferramenta, poucas conseguem alcançar a habilidade de avaliar, integrar e analisar os diversos dados disponíveis em seu parque informático. Assim, esta estatística sugere esforços para descobrir o que aflige o gerenciamento e a implementação dos projetos de *data warehouse*.

A proposta deste trabalho vem de encontro à esta necessidade, pois ao identificar os fatores organizacionais com baixa maturidade, indica as características de *data warehouse* que necessitam de maior urgência em melhorias ou correções.

A TRI é utilizada para efetuar medidas através de modelos que priorizam o item e não o instrumento de medida como um todo, buscando assim a efetiva comparação entre as organizações quanto ao *data warehouse*, como também a análise dos itens (requisitos ou fatores indicativos de tendência ao sucesso) que compõem o instrumento de medida, pela estimação do grau de maturidade das organizações em processos de *data warehouse* e dos parâmetros dos itens, em uma mesma métrica.

A maturidade de *data warehouse* segue alguns dos parâmetros do CMM, porém deve ser focada em processos de administração e manutenção de dados. O modelo proposto utiliza o CMM como um guia para definir níveis de maturidade dos fatores indicativos de sucessos do *data warehouse*. A identificação deste nível de maturidade, auxilia a organização na tomada de decisão nos assuntos relacionados ao *data warehouse*.

Após a verificação da aplicação do modelo através do estudo de caso, foi observado que o modelo proposto permite

analisar os requisitos de tendência ao sucesso que vão impactar nos processos de *data warehouse*.

Foi verificado que é possível identificar o nível de maturidade do ambiente organizacional nos requisitos que impactam na boa implantação, operação e utilização (visando a gestão, a solução) do *data warehouse*. A partir da maturidade destes fatores, permite ainda elencar possíveis problemas do ambiente organizacional.

O modelo possibilita ainda o fornecimento de diretrizes para auxiliar na resolução destes problemas. Estas diretrizes possibilitam à organização melhorar seus processos e, conseqüentemente, alcançar níveis mais elevados de maturidade.

Deste modo, o modelo proposto mostra-se como uma poderosa ferramenta de suporte à tomada de decisão, pois, à partir do momento que identifica quais os fatores estão com mais baixa maturidade, direciona a gerência do *data warehouse* aos pontos críticos que necessitam de correções e alterações.

Portanto, o problema apresentado nesta dissertação “*É possível a definição de um modelo para avaliação dos fatores do ambiente organizacional a partir de requisitos indicativos de sucesso da implantação e utilização de data warehouse como ferramenta de apoio à decisão gerencial?*” pode ser resolvido através da aplicação do modelo proposto, corroborando sua hipótese básica onde, a aplicação da TRI e a retirada de conceitos do CMM permitem a criação de um modelo de maturidade que seja aplicável à área de gerenciamento de dados, mais especificamente ao ambiente *data warehouse*.

A proposta aqui descrita objetiva auxiliar o processo de tomada de decisão das organizações através da identificação dos fatores organizacionais – indicativos de tendência de sucesso do *data warehouse* – que necessitam maior urgência em correções. Assim, é possível elevar estes fatores a níveis mais altos de maturidade.

A tomada de decisão baseada em fatos é um processo estratégico para qualquer organização. Este processo é auxiliado pelo modelo através da identificação dos aspectos organizacionais que necessitam de correções ou melhorias para que a adoção do *data warehouse* cumpra completamente seu papel: disponibilizar fatos corretos e coerentes para auxiliar a tomada de decisão de gerentes e administradores da organização.

7.2 Sugestão para Trabalhos Futuros

Diante do fato de que qualquer trabalho não esgota todas as possibilidades que permeiam seu tema, torna-se importante a continuidade do estudo, visando o aprimoramento dos resultados e ainda, a inferência de novos resultados.

A maior limitação do trabalho foi o estudo de caso ter sido feito em apenas uma empresa, assim, algumas sugestões e recomendações para futuros trabalhos são:

- Estender a aplicação do modelo a outras organizações;
- Examinar as diferenças entre organizações maduras e imaturas de DW com respeito a fatores tecnológicos, organizacionais ou culturais;
- Elaborar outros fatores que poderiam influenciar a maturidade dos processos de *data warehouse*;

- Criar uma ferramenta que efetue a modelagem dos processos de *data warehouse* visando facilitar o desenvolvimento de sua maturidade.
- Efetuar uma pesquisa em empresas brasileiras que verifique se o fator de carregamento (grau de importância) dos fatores discriminantes dentro de cada Fator Principal no ambiente organizacional equivale ao das empresas americanas.

Referências Bibliográficas

BAKER, F.B. The Basics of Item Response Theory. Second Edition. ERIC Clearinghouse on Assessment and Evaluation, 2001.

BAKER, F.B. Item Response Theory Parameter Estimation Techniques. Marcel Dekker, Inc, 1992.

BAYLEY, S. Measuring customer satisfaction. Evaluation Journal of Australasia, v. 1 (new series), n. 1, 2001.

CARNEIRO, A. J. W & Andrade, F. D. & Vasconcelos, P.A. & ARAÚJO, S. A. Uma Proposta de Análise de um Construto para Medição dos Fatores Críticos da Gestão pela Qualidade por Intermédio da Teoria da Resposta ao Item. Revista Gestão & Produção, Vol. 9, Nº 2, 2002

CMM, Capability Maturity Model® Integration (CMMISM), Version 1.1 CMU/SEI-2002-TR-029, ESC-TR-2002-029, 2002

CROSBY P.B. Quality is Free, McGraw-Hill, New York, NY, 1979.

DATAFLUX, 2005, Enterprise Data Management Maturity Model Online. www.dataflux.com, 2005

DEMING, W. E. Out of the Crisis, MIT Center for Advanced Engineering Study, Cambridge, MA, 1986.

DEROOS, Y. & Meares, P. A. Application of Rasch analysis: exploring differences in depression between african-american and white children. Journal of Social Service Research, v. 23, n. 3/4, p. 93-107, 1998

ECKERSON, W. "Gauge your data warehouse maturity," DM Rev. Nov. 2004 Online. www.DMReview.com

FIORINO, S.T. & STAA A.V. & BAPTISTA R.M., Engenharia de *Software* com CMM, Brasport, 1998.

FLEURY, Nélio Benedito. Das Habilidades Gerenciais Técnicas, Pessoais e Administrativas Para Direção, Supervisão e Gerência em Empresas de Engenharia Elétrica. Dissertação de Mestrado, EEEC, UFG, Brasil, 2005

GRANGER, C. V. & DEUTSCH, A. & Linn, R. T. Rasch analysis of the functional independence measure (FIMTM) mastery test. Arch Phys med Rehabil, v. 79, p. 52-57, 1998.

GONÇALVES, J. M. & BOAS A. V. Modelo de Maturidade de Capabilidade de *Software* (CMM) Versão 1.2 – Traduzido, Fundação CPqD. Online. www.cpqd.com.br, 2001

GUEDES, L. G. R. & RIBEIRO, L. G. Avaliação da Maturidade Gerencial em Projetos de Tecnologia. Anais do I Congresso Brasileiro de Gerenciamento de Projetos, 2004

HAMBLETON, R. K. & SWAMINATHAN, H. & ROGERS, H. J. Fundamentals of Item Response Theory. North Caroline. Sage Publications, 1991

HAZAN, Claudia. Uma ferramenta na busca da excelência Medições de *Software*. Online. <http://www.serpro.gov.br/publicacao/tematec/tematec/2004/ttec75>, 2004.

HUMPHREYS, Watts S. Characterizing the *Software* Process: A Maturity Framework, *Software* Engineering Institute, CMU/SEI-87-TR-11, ADA182895, June 1987.

IDC. The Financial Impact of Data Warehousing, 1996

INFORMATICA, Informatica Cursos Microsoft. Curso "Implementação do Processo de Data Warehousing", 2000.

INMON, W.H. Como Usar o Data Warehouse. ISBN: 8535201416
Editora IBPI, 1997

JURAN, J.M. Juran on Planning for Quality, Macmillan New York,
NY,1988.

LAMBERT, B. Data warehousing fundamentals. What you need to
know to succeed. Data Management Review. Março, 2000

LAKATOS, E. M. & MARCONI M. A. Fundamentos de Metodologia
Científica. São Paulo: Atlas, 2003

LANEY, D. Laney. Ascending the Information Maturity Model: Part
1—Data Quality Meta Group Online. [www.metagroup.com/us/
displayArticle.do?oid=29868](http://www.metagroup.com/us/displayArticle.do?oid=29868). Mar. 2000

LOZINSKY,2004 Lozinsky, Sérgio. Prontidão: um atributo-chave das
Empresas Competitivas, B2B Magazine, 2004. Online:
[http://www.ibm.com/br/services/articles/2004/04/a31m004t11s0220
04.shtml](http://www.ibm.com/br/services/articles/2004/04/a31m004t11s022004.shtml)

MARCO, D. "Meta data & knowledge management: capability maturity
model: an introduction," DM Rev. Online. [www.dmreview.com/
article_sub.cfm?articleId=5567](http://www.dmreview.com/article_sub.cfm?articleId=5567), Aug. 2002

MUÑNIZ, J. & HAMBLETON, R.K. Melo siglo de Teoría de Respuestas a
los Ítems. Anuario de Psicología, 52, 41-66, 1992

MULLINS, C .“The capability maturity Model - from a data
perspective, ” The Data Administration Newsletter. Online.
www.tdan.com, Dec. 1997

RADICE, R. A. & HARDING, J.T. & MUNNIS, P.E. & PHILIPS, R.W. A
Programming Process Study, IBM System Journal Vol.24,No.2,1985.

RIBEIRO, Lucília Gomes. Um Método para Avaliação de Maturidade Gerencial em Empresas de Tecnologia através da Teoria de Resposta ao Item. Dissertação de Mestrado, EEEC, UFG, Brasil, 2005

RIBEIRO, L. G. & GUEDES, L. G. R. Avaliação de Maturidade Gerencial em Empresas de Tecnologia - Estudo de Caso: COMDATA. In: III Fórum Associação Brasileira de Engenheiros Eletricistas - ABEE, Goiânia. 2005.

RIBEIRO, L. G. & GUEDES, L. G. R. . Método 2L de Quantificação da Maturidade Operacional e Conceitual. In: XII SIMPEP - Simpósio de Engenharia de Produção. Bauru.2005

RIBEIRO, L. G. & SANTOS, D. F. & GUEDES, L. G. R. & WEBER, M. Ferramenta Avaliativa Dinâmica a partir da Teoria de Resposta ao Item. In: I Encontro Regional em Modelagem e Análise Computacional de Sistemas. Goiânia. 2004

SANTANA, G. C. Fatores Críticos de Sucesso na Implantação de Soluções de Business Intelligence. Faculdade Ruy Barbosa, Salvador, 2000.

SANTANDER, V.F.A. & VASCONCELOS A.M.L. Mapeando o Processo Unificado em Realção ao CMM – Nível 2, XI CITS – Qualidade de *Software*, Curitiba, 2000.

SEI Software Engineering Institute, informações institucionais disponíveis em <http://www.sei.cmu.edu/cmm>

SEN, Arun & ATISH, P. Sinha & RAMAMURTHY K. Data Warehousing Process Maturity: An Exploratory Study of Factor Influencing User Perceptions. IEEE Transaction on Engineering Management, Vol. 53, Nº 3, 2006

SWAMINATHAN, H & HAMBLETON, R. K. Response Theory: Principles an Applications. Kluwer Nijhoff Publishing, 1985.

UNICAMP. Seminário sobre data warehouse. Disponível em:
[http://www.dei.unicap.br/~almir/seminarios/2000.1/datawarehouse/
index.htm](http://www.dei.unicap.br/~almir/seminarios/2000.1/datawarehouse/index.htm), 2000

Apêndice

Formulário de Pesquisa

Sobre a empresa na qual trabalha, é possível dizer que:	
1. Possui uma infra-estrutura de tecnologia robusta - máquinas, redes, sistemas operacionais e dados consistentes	1
2. Possui um grande número de aplicações de sistemas de informações dentro da empresa que podem ser utilizados e beneficiados pelo <i>data warehouse</i>	1
3. Possui alta complexidade dos dados que estão em uso atualmente dentro da empresa	2
4. Os usuários interessados em <i>data warehouse</i> possuem entendimento e visão do que o <i>data warehouse</i> pode fazer pela empresa	2
5. Os usuários possuem familiaridade com <i>data warehouse</i>	3
6. Faz-se necessário ter extensivo treinamento dos usuários chaves de dw para desenvolver as habilidades necessárias ao entendimento e apropriado uso do <i>data warehouse</i>	1
7. Existe alguma barreira de conhecimento significativa ao <i>data warehouse</i> .	6
8. Os principais usuários do DW são tecnicamente preparados para explorar plenamente suas potencialidades.	1
9. Possui confiança nos fatos, gráficos e diagramas para tomada de decisão	1
10. É aberta ao compartilhamento das informações pela organização	3
11. Incentiva e recompensa as tomadas de decisões baseadas em fatos	6
12. Os requisitos de dados estão em formatos customizados.	1
13. Os requisitos de dados são detalhados em diferentes níveis pelos usuários de negócios.	1
14. Utiliza relatórios atualizados para análises de tendência e para análises históricas.	1
15. Existem recursos financeiros comprometidos para o <i>data warehouse</i> .	6
16. Existe deficiência de pessoal de sistemas de informação disponíveis ao DW.	1
17. Todos os recursos técnicos necessários ao DW, tais como servidores de banco de dados e ferramentas de OLAP estão comprometidos com o mesmo.	1
18. Existe uma alta qualidade dos dados atualmente disponíveis na empresa.	1
19. Existe uma alta confiabilidade nos dados atualmente disponíveis na empresa.	1
20. Existem acordos claros sobre a definição dos dados na empresa.	1
21. Existem acordos claros sobre as regras, os processos que guiam os negócios da empresa.	5
22. As resoluções de possíveis problemas são, na maioria, satisfatórias	2
23. Os vários desafios técnicos e inter-pessoais que podem ocorrer durante a implementação do <i>Data warehouse</i> é gerenciado efetivamente.	2
24. Os vários desafios técnicos e inter-pessoais que podem ocorrer na transição para um novo ambiente de dados é gerenciado efetivamente.	3
Quanto ao gerente do projeto (ou gerente do <i>data warehouse</i>, quando houver), é possível dizer que:	
25. Sabe lidar com resistências políticas que possam ocorrer durante a implementação do <i>Data warehouse</i> .	2
26. Sabe gerenciar as mudanças, e suas conseqüências, que possam ocorrer durante a implementação do <i>Data warehouse</i> .	2
27. Consegue dar todo o suporte necessário ao processo de implementação do <i>Data warehouse</i>	2

Orientação: Por favor, responder cada pergunta com um dos números da escala abaixo

1–Discordo Totalmente 7–Concordo Totalmente

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)