

José Marcelo Pereira da Silva

"Linguagem Documentária para Mediação de Recuperação de
Informações Empresariais em Sistemas de Data Warehouse"

Marília
2002



JOSÉ MARCELO PEREIRA DA SILVA

**“LINGUAGEM DOCUMENTÁRIA PARA MEDIAÇÃO DE RECUPERAÇÃO DE
INFORMAÇÕES EMPRESARIAIS EM SISTEMAS DE *DATA WAREHOUSE*”**

*Dissertação apresentada ao Programa de Pós-Graduação em
Ciência da Informação do Departamento de Ciência da Informação
da Faculdade de Filosofia e Ciências - UNESP - Campus de Marília,
para obtenção do título de Mestre.*

Orientador: Prof. Dr. Alexandre Soares Lopes Filho

MARÍLIA-SP

2002



MARÍLIA-SP

2002



JOSÉ MARCELO PEREIRA DA SILVA

**“LINGUAGEM DOCUMENTÁRIA PARA MEDIAÇÃO DE RECUPERAÇÃO DE
INFORMAÇÕES EMPRESARIAIS EM SISTEMAS DE *DATA WAREHOUSE*”**

Dissertação apresentada ao Programa de *Pós-Graduação em
Ciência da Informação* do Departamento de Ciência da Informação
da Faculdade de Filosofia e Ciências – UNESP – Campus de Marília,
para obtenção do título de Mestre.

Orientadora: Doutora Mariângela Spotti Lopes Fujita

1410001352



MARÍLIA-SP

2002



UNIVERSIDADE DO ESTADO DE SÃO PAULO - UNESP - CAMPUS DE MARÍLIA
FACULDADE DE FILOSOFIA E CIÊNCIAS

ORIENTADORA : PROFA. DRA MARIÂNGELA SPOTTI LOPES FUJITA

BANCA EXAMINADORA:

PROF. DR. JOSÉ AUGUSTO CHAVES GUIMARÃES

PROFA. DRA MARIÂNGELA SPOTTI LOPES FUJITA

PROFA. DRA. MARTA POMPIM VALENTIM

MARÍLIA

2002



AGRADECIMENTOS

À Professora Doutora **Mariângela Spotti Lopes Fujita**, professora e orientadora sempre presente, um agradecimento muito especial. Que Deus a ilumine para orientando e conduzindo os caminhos sinuosos de seus alunos.

A **Marina**, esposa e companheira nos momentos bons e nos momentos difíceis.

As minhas amadas filhas **Gabriela** e **Mônica**, que são a essência de minha existência.

Aos meus pais **João** (*In memorian*) e **Cassiana**.

Aos meus irmãos **João Paulo** e **Júlio César**.

A **todos meus familiares** que me deram força e carinho na conquista dos meus objetivos.

Agradeço a **todos os meus amigos** de vida acadêmica e profissional que, de alguma forma, contribuíram nesta trajetória de formação.

RESUMO

Os sistemas de informações tornam-se, a cada dia, os mais importantes recursos para a produção do conhecimento, contudo, recuperar informações de forma eficiente está sendo o grande desafio de cientistas e profissionais da informação de todo o mundo. Como produto dos sistemas de informações, empresas especializadas em tecnologia criam sistemas de armazenamento de informações para tornar mais eficiente a sua recuperação. Os *Data Warehouse* são bancos de dados para armazenamento de informações empresariais centralizadas que contribuem de forma significativa para esta recuperação, todavia, isso acontece somente dentro de uma estrutura padronizada para uma interface de busca e recuperação previamente definidas. As ferramentas de recuperação de informações não estruturadas existentes no mercado não oferecem recursos para tratamento da pluralidade lingüística que existe entre a linguagem dos usuários de uma área de especialidade e a linguagem das bases de dados empresariais. A proposição, portanto, é o estudo teórico e metodológico de Representação Documentária para elaboração de um modelo de ferramenta de mediação de recuperação de informações em sistemas de *Data Warehouse* de empresas que sirva como linguagem de mediação entre os usuário e as bases de dados empresariais, cuja elaboração pelo profissional de tratamento da informação deverá ser permanentemente subsidiada por uma base terminológica das áreas de especialidades. O desenvolvimento dessa proposta teve como objetivos: investigar a interdisciplinaridade entre as áreas de Representação Documentária e Terminologia mediante referencial teórico publicado; comprovar a pluralidade lingüística no ambiente especializado através da coleta de linguagem dos usuários; elaborar modelo da ferramenta de mediação para recuperação de informações armazenadas em bancos de dados empresariais. A pesquisa revelou que a Representação Documentária contém conceitos importantes indicadores da necessidade da linguagem de especialidade do sistema de informação com a linguagem do usuário no processo de recuperação, e a Terminologia demonstrou a importância do termo e a necessidade de um banco de dados terminológico para garantir a memória e o contexto da linguagem do sistema de informação. A pluralidade lingüística revelou, também, a necessidade de se criar mecanismos de recuperação das informações, haja visto a constatação de que uma mesma informação pode ser solicitada por

termos diferentes. O processo de representação da informação, mediante linguagem de mediação, apresentou a solução desse problema ao proporcionar um relacionamento entre os termos representativos das linguagens do sistema de informação e do usuário. Concluiu-se que as duas áreas proporcionaram a interdisciplinaridade necessária à elaboração da proposta de modelo de ferramenta de mediação de recuperação de informações em sistemas de *Data Warehouse* de empresas que contempla os conceitos teóricos de organização do conhecimento com características adequadas para tratar de recuperação de informações de uma área de especialidade, oferecendo aos usuários de informações em ambientes empresariais um novo recurso para recuperar informações e ampará-los no processo de tomada de decisões.

ABSTRACT

The information systems are becoming the most important resources for the production of knowledge, but to recover the information in an efficient way is the great challenge of scientists and information professionals around the world. As a product of information systems, technology specialized enterprises create systems of information storage to make its recovering more efficient. *Data Warehouse* are data base for storing the centralized enterprise information which contribute for this recovering, however, this happens just in a patterned structure for a search and recovering interface previously determined. The not-structured tolls of information recovering available don't provide resources for the linguistic plurality treatment between the users' language from a specialized area and the language of enterprises' data bases. The proposal, though, is the theoretic and methodological study of Documentary Representation for the elaboration of a mediation toll model for information recovering in enterprises' *Data Warehouse* systems which is used as mediation language between the users and the enterprises' data bases, whose elaboration by the information professional must be permanent sponsored by a terminological base from the specialized areas. The development of this proposal had these aims: to investigate the interdisciplinarity between Documentary Representation and terminology with a published theoretical referential; to prove the linguistic plurality on specialized environment through the language analyses which the users produce; to elaborate a mediation toll model for information recovering in stored in enterprises' data bases. The research showed that the Documentary Representation has important conceits which indicate the necessity of specialized language of the information system with the user's language in the recovering process, and the terminology showed the term importance and the necessity of a terminological data base to grant the memory and the language context of the information system. The linguistic plurality also revealed the necessity to create mechanisms of information recovering, as shown by the confirmation which the same information may be asked by different terms. The process of information representation, with mediation language, presented the solution of this problem when it promoted a relationship between language representative terms of the information system and the user's language. We can conclude that both areas provided the interdisciplinarity needed to

elaborate the proposal of mediation toll model of information recovering on enterprises' *Data Warehouse* system which contemplate the theoretical conceits of knowledge organization with suitable features to treat information recovering of a specialized area, providing to the information users on business environment a new resource to recover information and help them in the decision taking process.

Introdução	35
Capítulo 1. Fundamentos de Informação e Organização do Conhecimento	39
1.1. Conceitos	39
1.2. Fundamentos da Organização do Conhecimento	41
1.3. Fundamentos da Organização da Informação	43
1.4. Fundamentos da Organização da Medição	45
1.5. Fundamentos da Organização da Medição	47
1.6. Fundamentos da Organização da Medição	49
1.7. Fundamentos da Organização da Medição	51
1.8. Fundamentos da Organização da Medição	53
1.9. Fundamentos da Organização da Medição	55
1.10. Fundamentos da Organização da Medição	57
1.11. Fundamentos da Organização da Medição	59
1.12. Fundamentos da Organização da Medição	61
1.13. Fundamentos da Organização da Medição	63
1.14. Fundamentos da Organização da Medição	65
1.15. Fundamentos da Organização da Medição	67
1.16. Fundamentos da Organização da Medição	69
1.17. Fundamentos da Organização da Medição	71
1.18. Fundamentos da Organização da Medição	73
1.19. Fundamentos da Organização da Medição	75
1.20. Fundamentos da Organização da Medição	77
1.21. Fundamentos da Organização da Medição	79
1.22. Fundamentos da Organização da Medição	81
1.23. Fundamentos da Organização da Medição	83
1.24. Fundamentos da Organização da Medição	85
1.25. Fundamentos da Organização da Medição	87
1.26. Fundamentos da Organização da Medição	89
1.27. Fundamentos da Organização da Medição	91
1.28. Fundamentos da Organização da Medição	93
1.29. Fundamentos da Organização da Medição	95
1.30. Fundamentos da Organização da Medição	97
1.31. Fundamentos da Organização da Medição	99
1.32. Fundamentos da Organização da Medição	101
1.33. Fundamentos da Organização da Medição	103
1.34. Fundamentos da Organização da Medição	105
1.35. Fundamentos da Organização da Medição	107
1.36. Fundamentos da Organização da Medição	109
1.37. Fundamentos da Organização da Medição	111
1.38. Fundamentos da Organização da Medição	113
1.39. Fundamentos da Organização da Medição	115
1.40. Fundamentos da Organização da Medição	117
1.41. Fundamentos da Organização da Medição	119
1.42. Fundamentos da Organização da Medição	121
1.43. Fundamentos da Organização da Medição	123
1.44. Fundamentos da Organização da Medição	125
1.45. Fundamentos da Organização da Medição	127
1.46. Fundamentos da Organização da Medição	129
1.47. Fundamentos da Organização da Medição	131
1.48. Fundamentos da Organização da Medição	133
1.49. Fundamentos da Organização da Medição	135
1.50. Fundamentos da Organização da Medição	137
1.51. Fundamentos da Organização da Medição	139
1.52. Fundamentos da Organização da Medição	141
1.53. Fundamentos da Organização da Medição	143
1.54. Fundamentos da Organização da Medição	145
1.55. Fundamentos da Organização da Medição	147
1.56. Fundamentos da Organização da Medição	149
1.57. Fundamentos da Organização da Medição	151
1.58. Fundamentos da Organização da Medição	153
1.59. Fundamentos da Organização da Medição	155
1.60. Fundamentos da Organização da Medição	157
1.61. Fundamentos da Organização da Medição	159
1.62. Fundamentos da Organização da Medição	161
1.63. Fundamentos da Organização da Medição	163
1.64. Fundamentos da Organização da Medição	165
1.65. Fundamentos da Organização da Medição	167
1.66. Fundamentos da Organização da Medição	169
1.67. Fundamentos da Organização da Medição	171
1.68. Fundamentos da Organização da Medição	173
1.69. Fundamentos da Organização da Medição	175
1.70. Fundamentos da Organização da Medição	177
1.71. Fundamentos da Organização da Medição	179
1.72. Fundamentos da Organização da Medição	181
1.73. Fundamentos da Organização da Medição	183
1.74. Fundamentos da Organização da Medição	185
1.75. Fundamentos da Organização da Medição	187
1.76. Fundamentos da Organização da Medição	189
1.77. Fundamentos da Organização da Medição	191
1.78. Fundamentos da Organização da Medição	193
1.79. Fundamentos da Organização da Medição	195
1.80. Fundamentos da Organização da Medição	197
1.81. Fundamentos da Organização da Medição	199
1.82. Fundamentos da Organização da Medição	201
1.83. Fundamentos da Organização da Medição	203
1.84. Fundamentos da Organização da Medição	205
1.85. Fundamentos da Organização da Medição	207
1.86. Fundamentos da Organização da Medição	209
1.87. Fundamentos da Organização da Medição	211
1.88. Fundamentos da Organização da Medição	213
1.89. Fundamentos da Organização da Medição	215
1.90. Fundamentos da Organização da Medição	217
1.91. Fundamentos da Organização da Medição	219
1.92. Fundamentos da Organização da Medição	221
1.93. Fundamentos da Organização da Medição	223
1.94. Fundamentos da Organização da Medição	225
1.95. Fundamentos da Organização da Medição	227
1.96. Fundamentos da Organização da Medição	229
1.97. Fundamentos da Organização da Medição	231
1.98. Fundamentos da Organização da Medição	233
1.99. Fundamentos da Organização da Medição	235
1.100. Fundamentos da Organização da Medição	237

SUMÁRIO

Apresentação	12
Capítulo 1: Introdução	15
Capítulo 2: Representação Documentária	24
Capítulo 3: Terminologia: e sua interface com linguagens documentárias no processo de representação documentária	38
Capítulo 4: Recuperação de informações empresariais em sistemas de <i>Data Warehouse</i>	59
Capítulo 5: A Terminologia dos usuários de bancos de dados <i>Data Warehouse</i> : relato de experiência	77
Capítulo 6: Modelo da ferramenta de mediação	93
Conclusão	112
Referências Bibliográficas	119
APÊNDICES:	
Apêndice A: Instruções aos participantes da pesquisa sobre a técnica de coleta de dados do protocolo verbal ou "pensar alto"	124
Apêndice B: Transcrição do Protocolo Verbal com as anotações do usuário 1	125
Apêndice C: Transcrição do Protocolo Verbal com as anotações do usuário 2	128
Apêndice D: Padrão textual para lay-out do documento do usuário informante 1	129
Apêndice E: Padrão textual para lay-out do documento do usuário informante 2	130
ANEXOS:	
Anexo A: Estrutura das tabelas de dados do <i>Data Warehouse</i>	131
Anexo B: Lista das tabelas com os atributos das dimensões	132
Anexo C: Notações de transcrição	133
Glossário	134
Lista de Ilustrações	9
Lista de Abreviaturas	10
Quadro A - Análise comparativa dos termos diferentes utilizados entre o padrão textual elaborado pelos usuários informantes	90
Quadro B - Análise comparativa dos termos diferentes constantes nos sistemas de informação e nos documentos gerados pelos participantes	91

LISTA DE ILUSTRAÇÕES

Figura 1 – Modelo dimensional típico	66
Figura 2 – Modelo dimensional com entidades eventuais	68
Figura 3 : Fluxograma lógico da ferramenta de mediação	101
Figura 4 : Fluxograma físico da ferramenta de mediação	103
Figura 5: Gráfico de relacionamento de termos	104
Figura 6: Interface de entrada de dados para pesquisa	105
Figura 7: Interface com a listagem dos termos relacionados ao termo pesquisado	106
Figura 8: Interface de solicitação de registro de termos	107
Figura 9: Interface do resultado da consulta de termos	107



LISTA DE ABREVIATURAS

DBA - *Data Base Administrator* - Administrador de Banco de Dados

DOI - Departamento de Organização e Infomática

OLTP - *On Line Transaction Processing* - Processamento de transações instantâneas.

SAD - Sistema de Apoio à Decisão

SGBD - Sistema Gerenciador de Banco de Dados

SQL - *Struct Query Language* - Linguagem de acesso estruturada



APRESENTAÇÃO

O autor desta pesquisa fundamentou-se na experiência profissional adquirida como profissional da área de tecnologia da informação, especializando-se em recuperar informações em bases de dados de sistemas de informações empresariais, solicitadas por usuários na empresa que trabalha. Sua principal missão como profissional dentro da empresa é analisar as necessidades dos usuários e construir ferramentas através de recursos tecnológicos, para satisfazer os desejos dos usuários no que diz respeito aos propósitos da empresa.

Nesse contexto, durante aproximadamente dez anos de serviços prestados, o autor teve oportunidade de participar e coordenar vários projetos de sistemas de informações e processos de recuperação previamente formulados para permitir aos usuários perspectivas de análise dos negócios da empresas.

Dessa necessidade o autor sempre buscou a melhor forma em atender às necessidades, encontrando soluções que satisfizessem os usuários com o menor nível de dependência do profissional de tecnologia, acreditando ser essa uma solução eficiente.

Com o objetivo de ampliar seus conhecimentos no tratamento da informação, o autor interessou-se em estudar na área da Ciência da Informação o processo de recuperação utilizado no meio científico, visando, assim, poder abstrair técnicas e conceitos da área para elaborar uma solução técnica para os problemas de recuperação presenciados pelo autor no ambiente de trabalho.

Como resultado desse estudo, o autor propôs-se modelar uma ferramenta como solução significativa para diminuir as dificuldades dos usuários dos sistemas de informações da empresas em recuperar informações das bases de dados.



As constantes necessidades de recuperação das informações armazenadas nas bases de dados dos sistemas de informações, apresentadas pelos usuários aos profissionais de tecnologia para executar suas tarefas, despertaram a busca de uma solução que auxilie os usuários na recuperação das informações. Diagnosticou-se que uma das principais dificuldades encontradas pelos usuários, quando recorrem aos profissionais de tecnologia, está em localizar as informações na base de dados e identificar o nome do campo determinado pelo profissional de tecnologia, onde está armazenada a informação.

A partir de diagnóstico, analisou-se o desempenho das ferramentas de recuperação das informações disponíveis aos usuários da empresa e constatou-se que a linguagem sofisticada apresentada nessas ferramentas dificulta sua utilização, embora tais ferramentas ofereçam diversos recursos. Nesse sentido, estudou-se uma solução que facilite o acesso às informações baseado na linguagem dos usuários dos sistemas de informações da empresa.

Um dos principais objetivos deste trabalho é a realização de uma pesquisa exploratória mediante a revisão de literatura que demonstre os conceitos para organização do conhecimento, permitindo conhecer cientificamente formas de tratamento da informação. Nesse enfoque, fez-se o aprofundamento para identificar critérios de armazenamento da informação vinculados aos processos de recuperação a fim de investigar a possibilidade de elaboração da interface de um sistema de recuperação.

Com essa perspectiva, a pesquisa busca encontrar na área de tratamento da informação a fundamentação teórica necessária que viabilize a modelagem de um sistema de mediação para recuperação de informações em ambiente empresarial.

Para isso, estudam-se os critérios e definições de representação da informação na área da Análise Documentária para conhecer conceitos relevantes que devam ser incorporados ao modelo a ser sugerido. Tais conceitos são especificados de forma mais detalhada pela Linguagem Documentária, na qual intensificou-se a investigação por ser um instrumento cuja função é garantir que o conhecimento não se perca.

Na investigação da evolução histórica do tratamento da informação surge a necessidade de estudo de outras áreas como a terminologia, pois, ao se falar de armazenamento de palavras, é importante manter o seu significado, assim nesta pesquisa exploram-se os critérios da área de terminologia com a finalidade de preservar a identidade do termo quando isolado do texto.

Na análise das necessidades de recuperação de informações empresariais, fez-se o levantamento do problema existente em ambiente empresarial e entrevista com os usuários do ambiente empresarial para comprovação do problema.



INTRODUÇÃO

A informação torna-se, a cada dia, um dos mais importantes ingredientes para o desenvolvimento da humanidade, principalmente para a evolução da ciência no campo tecnológico. A evolução tecnológica está modificando significativamente o perfil das empresas, num processo contínuo e sem volta; isto abrange desde a sua base, que é a operação e a execução das tarefas para cumprimento das metas estabelecidas, passando pela estrutura de gerenciamento das operações e definição de metas até a execução das atividades definidas pelo setor estratégico da empresa.

Os avanços tecnológicos estão contribuindo para a descoberta de técnicas cada vez mais aprimoradas que permitem projetar e desenvolver sistemas de armazenamento e recuperação de informações de forma mais rápida e eficiente. Considerando os recursos tecnológicos disponíveis e acessíveis na atualidade, muitas soluções são implementadas nas mais diversas áreas de atuação para facilitar, melhorar e agilizar a vida do homem em suas tarefas. Isso lhe possibilita a utilização de tempo e de energia, economizados nas tarefas realizadas com a ajuda de novos recursos tecnológicos.

A principal finalidade de um sistema de informação em uma empresa, além de colher e armazenar informações, é recuperar informações como matéria-prima na realização de análise para obter e gerar conhecimento sobre o andamento das atividades. As informações consolidadas fornecem perspectivas de tendências dos acontecimentos internos e externos ao resgatar os dados históricos, permitindo tomadas de novas estratégias ou correções das estratégias aplicadas para atingir os objetivos sociais e financeiros desejados pelas empresas.

A independência dos usuários dos sistemas de informações das empresas na localização e recuperação das informações desejadas, além de diminuir as solicitações de suporte dos profissionais da área de tecnologia da informação, faz com que todos os usuários possam buscar

por si e de forma direta as informações necessárias a serem incorporadas às suas tarefas a fim de aumentar o desempenho das atividades nas empresas.

A recuperação das informações armazenadas nos bancos de dados dos sistemas de informações empresariais é dificultada pelas limitações técnicas ou funcionais de acesso às informações depositadas nos computadores. Torna-se cada vez mais necessário criar novos recursos tecnológicos, para fornecer aos usuários que trabalham com as informações empresariais meios de facilitar a recuperação das informações.

A necessidade de maior interação entre os setores e os diversos departamentos faz da tecnologia uma das mais importantes ferramentas para a sobrevivência das empresas em um mundo cada vez mais competitivo, através do desenvolvimento e utilização de sistemas de informação.

Na análise de alguns requisitos para a criação de novos recursos tecnológicos, pode-se constatar que um dos principais obstáculos ao processo de recuperação das informações é a dificuldade de representação de um vocabulário específico para identificar as informações empresariais nos bancos de dados, pois a indicação e definição de termos técnicos ou abreviaturas não convencionais por profissionais da área tecnológica dificultam a interpretação dos dados.

Embora existam muitas ferramentas para fazer a recuperação de informações armazenadas em banco de dados, elas não possuem recursos de tratamento da linguagem para auxiliar os usuários na busca e recuperação das informações desejadas para atender às suas necessidades.

Vários trabalhos são desenvolvidos e implantados nos sistemas de informação das empresas para a otimização de tarefas e controle dos dados em diversos setores da empresa. As empresas construíram recursos para desenvolver e dispor aos usuários gestores as ferramentas de *Data Mining*. Em virtude de estas aplicações ocasionarem um grande impacto nos setores que fazem a garimpagem de dados para análise de resultados e de tendências, surgiu a necessidade de uma metodologia mais eficiente para a busca e recuperação da informação.

Data Mining é uma técnica que permite buscar, em grandes bases de dados, informações que, aparentemente, estão camufladas ou escondidas, permitindo maior agilidade nas tomadas de decisões.

As ferramentas de *Data Mining* podem determinar padrões de comportamento, como associação de produtos durante um processo de compras; com isto, as ferramentas *Data Mining* indicam a presença de oportunidades e "*insights*" em relação àquele público consumidor. O diferencial do *Data Mining* está no fato de que as descobertas de padrões de comportamento se dão por uma lógica de algoritmos¹. São ferramentas de descobertas matemáticas feitas sobre os registros corporativos já processados, contra descobertas empíricas.

Com o objetivo de atender a essas necessidades e aprimorar as técnicas de desenvolvimento de ferramentas de recuperação de informações, a partir de estudos científicos, esta pesquisa baseou-se no trabalho de desenvolvimento do projeto "Alternativa Metodológica para Design e Desenvolvimento de *Data Warehouse* com base em *Data Mart*".

Para o desenvolvimento do projeto foi necessário criar um banco de dados com as características de um dicionário, contendo os principais arquivos dos sistemas transacionais, porém a implementação desse projeto não foi a solução definitiva para os problemas de recuperação por parte dos usuários, existe a necessidade de explorar informações analíticas nessa base de dados para atender as suas necessidades.

Para dar maior flexibilidade de acesso às informações, disponibilizaram-se aos usuários ferramentas de "garimpagem" de dados e dicionários. Porém permaneceu a dificuldade de recuperação, pois ainda há o fator lingüístico a ser superado. Sem uma solução continuam os problemas de não se encontrarem no dicionário as informações desejadas, por estarem identificadas de forma diferente da palavra pesquisada. Nesses casos os usuários recorreram constantemente aos analistas de sistema para esclarecê-los.

¹ Lógica de algoritmos são instruções ordenadas em uma linguagem técnica para realização de tarefas.



A linguagem utilizada na alimentação dos dicionários de dados é geralmente elaborada por técnicos em informática a qual é diferente dos termos utilizados pelos usuários, isso dificulta a recuperação da informação armazenada nas bases de dados, por não se saber em qual estrutura está armazenada a informação ou pela falta de recursos para descobrir com qual termo foi definida.

Dessa forma existe a necessidade de se encontrar uma solução para que os usuários possam obter respostas para as suas dúvidas a respeito da informação que desejam recuperar, sem ter que recorrer ao profissional de suporte ou aos analistas de sistemas, os quais, na maioria dos casos, estão sobrecarregados com outras tarefas, ficando os usuários limitados em suas atividades e na dependência da informação.

Como alternativa sugerida para suprir as necessidades dos usuários na busca de informações da base de dados dos sistemas das empresas, poder-se-iam considerar as seguintes soluções: a) Programa de Treinamento de Usuário - apresentando a todos os usuários, em forma de seminários periódicos, as dúvidas mais comuns sobre como está identificada ou denominada a informação desejada para pesquisa, objetivando unificar a linguagem interna; b) manter a sistemática de suporte de atendimento aos usuários, permitindo que esse departamento ministre minicursos para os usuários e áreas da empresa que tenham mais necessidade de informação, com base nas estatísticas de atendimento; c) implementação de uma ferramenta de consulta para uso dos usuários, permitindo-lhes identificar o nome técnico ou representativo dado pelos profissionais da informação ou por outros departamentos da empresa.

No estudo das soluções acima destacadas, possivelmente a que trará resultados mais dinâmicos e objetivos e beneficiará os usuários dos sistemas de informações existentes e os novos que poderão surgir, será a implementação de uma ferramenta de consulta acessível aos usuários, permitindo-lhes identificar o nome técnico ou representativo e a localização da informação desejada com base no próprio vocabulário.

A constatação da realidade observada sugere a necessidade de elaboração de uma ferramenta de mediação que possibilite ao usuário final fazer pesquisas e obter a palavra que

identifique a informação que ele procura no dicionário de dados, permitindo a utilização de sua própria linguagem durante a pesquisa. Essas informações poderão ser acessadas futuramente por usuários de diversos departamentos e de diferentes regiões do país.

Essa necessidade, contudo, identificou de forma preliminar a realização de uma pesquisa exploratória, mediante a revisão de literatura para investigar a possibilidade de elaboração da interface do sistema de recuperação.

A idealização da proposta desta pesquisa iniciou-se a partir do estudo de procedimentos para a construção de Linguagem Documentária, existente na tese de doutorado "Linguagem Documentária em Odontologia", da Fujita (1992), possibilitando verificar que a recuperação das informações pelos usuários dos sistemas da empresa precisa utilizar formas de tratamento de Representação Documentária dotada de uma estrutura de termos relacionados.

Analizando as relações dos significados das informações para os diversos setores das empresas e os diferentes termos utilizados para referir-se a elas, verificam-se algumas propriedades existentes na relação entre termos e pode-se caracterizá-las de forma a manter um elo entre eles.

Como se trata de uma solução de recuperação de informações armazenadas em bancos de dados empresariais, é natural que o sistema de mediação, para cumprir seu papel de mediador na recuperação de informações, precise de mecanismos internos de buscas e verificação de termos que estejam registrados em algum suporte. Em função dessa necessidade, é essencial para o sistema de mediação a construção de uma base para armazenamento dos termos, a qual comporte o registro dos termos da linguagem natural dos usuários e das denominações das informações nas bases de dados das empresas.

Considerando que a área da terminologia trata da fixação de conceitos, visando elaborar definições orgânicas e estabelecer princípios para a criação de novos termos que possibilitem uma comunicação mais precisa entre especialistas de diversas áreas do conhecimento, faz-se necessário pesquisar também os fundamentos da terminologia como campo de estudo, relatando a

reconstituição histórica da ciência terminológica para demonstrar sua importância na organização do conhecimento humano.

Dessa forma as constatações teóricas preliminares revelam a possibilidade de investigação interdisciplinar entre a Análise Documentária e a Terminologia. Esta afirmação está bem definida por Tálamo (1992, p.199) quando se trata de construção de tesouros, na aplicação dos instrumentos de controle terminológico e de representação da informação.

Assim, a proposição dessa pesquisa é o estudo teórico e metodológico da Representação Documentária para elaboração do modelo de mediação de recuperação de sistemas de *Data Warehouse* de empresas, o qual sirva como linguagem de mediação entre o usuário e os bancos de dados empresariais, cuja elaboração pelo profissional da informação deverá ser permanentemente subsidiada por uma base terminológica das áreas de especialidades.

Data Warehouse é considerado um sistema de informações construído para atender setores estratégicos das empresas como ferramenta de auxílio na tomada de decisões.

Para tanto a pesquisa tem como objetivos instrumentais:

- estudo teórico e metodológico em Representação Documentária e Terminologia para demonstrar a interdisciplinaridade entre Análise Documentária e Terminologia;
- coleta e estudo da linguagem utilizada pelos usuários como prova da diferença lingüística entre o sistema de informação e os usuários.
- elaboração de um modelo da ferramenta de interface para auxiliar na recuperação das informações armazenadas nas bases de dados dos computadores.

Para a elaboração desta pesquisa, considerando as atividades necessárias para atingir o objetivo para a proposta da ferramenta-modelo de mediação, planejou-se o desenvolvimento em etapas, com os trabalhos executados conforme descrito nos parágrafos seguintes.



Primeiramente, elaborou-se o referencial teórico de estudos em Representação Documentária para o resgate de técnicas e conceitos da área na representação e recuperação das informações, para especializar ainda mais o processo de representação da linguagem do usuário e a relação com a linguagem técnica, efetuou-se o estudo da terminologia para conceituação e aplicação complementar do processo de busca e recuperação da informação.

Na segunda fase, trabalhou-se no estudo dos sistemas de recuperação da informação em empresas com *data warehouse* com a finalidade de conhecer sua funcionalidade, suas vantagens e prováveis deficiências, para conhecimento e análise dos problemas existentes.

Na terceira fase, adotaram-se métodos de recuperação de informações com base na experiência adquirida pelo autor da pesquisa em um ambiente com os requisitos necessários descritos na fase acima, apropriados para o desenvolvimento desta pesquisa.

Na quarta fase, desenvolveu-se a atividade de identificação das diferenças terminológicas no ambiente empresarial, por meio de coleta de palavras selecionadas da linguagem dos usuários do setor adotado como estudo de caso.

Na quinta fase, como finalidade principal de toda a pesquisa, após estudos de cada etapa acima, elaborou-se um modelo de ferramenta auxiliar para solucionar os problemas de recuperação das informações das bases de dados empresariais.

Para apresentação do resultado dos estudos realizados para o desenvolvimento da pesquisa, organizou-se o trabalho conforme descrito na sequência.

No capítulo dois, realizou-se levantamento bibliográfico sobre a representação documentária e controle de vocabulário, buscou-se recuperar a aplicação dos conceitos e técnicas da Análise Documentária, no tratamento terminológico da linguagem natural relacionada à linguagem dos sistemas de informações, como caminho científico para a elaboração do modelo

alternativo da ferramenta intermediária de recuperação das informações existentes nos sistemas empresariais.

No capítulo três, desenvolveu-se a pesquisa em uma área mais específica do conhecimento, agregando conceitos de controle específico para o tratamento das palavras dos usuários, e as palavras-chave de recuperação, fez-se o estudo histórico das estruturas de tratamento terminológico cuja Terminologia importa conceitos da Representação Documentária.

No capítulo quatro, destacando-se as funcionalidades de sistemas de *Data Warehouse*, considerado um dos principais sistemas de informações de recuperação de informações empresariais, o histórico do processo de desenvolvimento e aprimoramento dos sistemas de informações, as dificuldades de acesso às informações pelos usuários em suas atividades diárias, o perfil dos usuários que operam os sistemas transacionais e dos que fazem uso das informações armazenadas nos bancos de dados.

No capítulo cinco, realizou-se o estudo da terminologia dos usuários que utiliza os sistemas de informações da empresa, para obter informações dos usuários aplicaram-se técnicas de identificação dos termos, relatam-se a aplicação da técnica de coleta de termos chamada de "Protocolo Verbal", os procedimentos do processo de coleta, a diferença existente nos termos coletados dos usuários, como fatos sinalizadores das diferenças como elemento importante para a elaboração desta pesquisa.

Ressalta-se a importância das relações semântica e sintática nos sistemas e esquemas para recuperação da informação, com base nos trabalhos de pesquisa e referências existentes em Representação Documentária. Estar-se-á enriquecendo esta pesquisa com os conhecimentos aplicados pela Representação Documentária, nos critérios de armazenagem e recuperação de informação, enfocando a utilização da linguagem natural dos usuários que realizam as pesquisas de recuperação.

Continuamente as empresas e os profissionais da área de tecnologia da informação vêm desenvolvendo ferramentas tecnológicas para otimizar o processo de recuperação das



informações armazenadas nos sistemas empresariais, devido à importância adquirida pela “informação” nas últimas duas décadas; a cada dia novas necessidades de tratamento e organização são exigidas, tornando constante o processo de busca na eficiência de armazenamento e recuperação de informações.

Na análise do uso das ferramentas disponibilizadas para a recuperação de informações, observa-se que as ferramentas, embora tenham grandes inovações tecnológicas, são ineficazes, em virtude das diferenças existentes entre a linguagem utilizada pelo usuário e a linguagem utilizada pelos técnicos na definição da estrutura dos bancos de dados empresariais. A busca de uma solução que auxilie os usuários no uso das ferramentas para recuperação é exercício constante dentro das empresas.

A necessidade de desenvolvimento de técnicas que ofereçam maior flexibilidade de acesso às informações é o desafio de cientistas e estudiosos do mundo inteiro. O propósito da presente pesquisa é encontrar recursos técnicos e científicos que possam fornecer a flexibilidade necessária para o acesso às informações.

As investigações realizadas para o construir a solução dos problemas destacados nesta pesquisa forneceu elementos importantes para a elaboração do modelo da ferramenta de mediação, com fruto desse trabalho o modelo da ferramenta de mediação tem por objetivo auxiliar na recuperação de informações armazenadas nos bancos de dados empresariais.



CAPÍTULO 2

REPRESENTAÇÃO DOCUMENTÁRIA

Esta pesquisa, apoiada em estudos da área da Ciência da Informação e com base nas perspectivas de soluções dos problemas de recuperação de informações, vincula-se à área de Análise Documentária para garantir de forma adequada a representação, armazenagem e recuperação de informações, estudando-se os padrões e variações das formas e dos códigos de representação, visando melhorar e ampliar as fronteiras de domínio do conhecimento humano.

Neste capítulo expõe-se a fundamentação teórica da Representação Documentária para evidenciar a função das Linguagens Documentárias, como instrumentos importantes na elaboração do modelo, cuja finalidade básica se funde ao funcionamento e objetivo do sistema de mediação.

Segundo Ginez de Lara (1993, p. 4) a Análise Documentária atua especificamente no interior da Documentação, através da análise e tratamento da informação, com finalidade de garantir sua recuperação. Pode-se afirmar que, sendo essa a atividade da Análise Documentária, está garantindo também com a recuperação da informação armazenada que o conhecimento acumulado não se perca.

Segundo Barreto (1994, p. 3 citado por GINEZ DE LARA, 1999, p. 136) “informação é um agente mediador na produção de conhecimento e qualifica-se, em forma e substância, como estruturas significantes com a competência de gerar conhecimento”. A mediação ocorre na aplicação de técnicas que facilita a comunicação entre entidades elementares com estruturas organizacionais diferenciadas, portanto o uso das técnicas de acesso as informações podem ser consideradas o processo de mediação em execução.

Segundo Ginez de Lara (1999, p. 138) o enfoque da mediação em Documentação acontece mais exatamente entre a linguagem do sistema e a linguagem do usuário, pois esta diretamente relacionado a um problema lingüístico-comunicacional.

Em Análise Documentária a informação é um elemento mediador com a finalidade de disseminar e produzir conhecimento, dessa forma, qualquer ferramenta que faça parte do processo de armazenamento e recuperação da informação para a geração de conhecimento pode ser considerada uma ferramenta de mediação.

Como o resultado final dessa pesquisa é propor um modelo de interface para recuperação de informações em sistemas computadorizados a partir da linguagem do usuário. Tal modelo está caracterizado pela necessidade de uma estrutura de representação e ligação entre as linguagens dos sistemas e dos usuários. Torna-se indispensável na elaboração do modelo a aplicação dos estudos da representação documentária, considerando-se que sua função é “execer a *mediação*, agenciando a produção do conhecimento para modificar os estoques de informação dos indivíduos e das sociedades” (GINEZ DE LARA, 1999, p.136).

Do ponto de vista de Alonso (1997, p. 28), ressalta-se a informação como elemento significativo que marcou o início da era pós-moderna, constituindo-se o centro de toda estrutura econômica e social no desenvolvimento da humanidade.

Existem diferentes definições sobre informação, e sua abrangência permite expor pontos de vistas diferentes sobre seu significado. A definição que melhor incorpora o valor da informação em um contexto lingüístico é a de Garvin (1993, p.3-22 citado por ALONSO, 1997, p. 30) que define a informação “um sistema de signos cuja estrutura se especifica em termos de níveis de construção, de níveis de organização e mais de um nível de integração”.

Confirma-se o descrito no parágrafo nos estudos de Kobashi (1994, p.80) que define e caracteriza as três operações da Análise Documentária da seguinte forma: a) Análise: leitura e segmentação do texto para identificação e seleção de conceitos; b) Síntese: construção do texto documentário com os conceitos selecionados; c) Representação; a representação possui duas

naturezas distintas: na primeira, ela é construída através de um processo de condensação intensiva do texto original. Essa representação gera diferentes tipos de resumo como produtos documentários. Na segunda, a representação é realizada através do uso de um código comutador, ou seja, uma linguagem documentária.

Do ponto de vista de Ginez de Lara (1999, p. 136) Representação Documentária pode ser definida como “o resultado das operações de análise e síntese de conteúdo de textos, cujo objetivo é a transferência da informação”.

No âmbito da Representação Documentária, com a perspectiva de incorporar os elementos necessários, Ginez de Lara (1999, p.136) divide a definição da informação a partir de três elementos participantes do processo informativo, definindo a informação de cada elemento da seguinte forma: a) o gerador como “resultado da relação estática entre objetos materiais, independente da ação de seres humanos”; b) enquanto mensagem como “símbolos produzidos por um gerador para efetivar um processo de transferência” (BARRETO, 1994, p.3, citado por GINEZ DE LARA, 1999, p. 136); c) o receptor que “reforça a intenção semântica da transferência, adjetivando o conceito com o significado da mensagem, seu uso efetivo e ação resultante do uso”.

A partir dessas definições, concorda-se com Cintra (1994, p.10) que reforça a idéia de que a informação se constrói no encontro de duas dinâmicas: a de quem “emite” e a de quem “recebe”. Viabilizando a transmissão de conhecimentos entre os indivíduos num processo contínuo.

A informação pode ser considerada como a representação gráfica ou sonora do conhecimento exteriorizado pelo emissor, disponibilizado em um veículo de comunicação acessível aos indivíduos; a compreensão dessa representação tem por finalidade a aquisição de conhecimento pelo receptor.

O armazenamento da informação é recurso que faz com que o conhecimento humano não se perca. Pode-se registrar o conhecimento transformando-o em informação; para isso, o homem

inventou sofisticadas técnicas de representação e de transferência do conhecimento do indivíduo; uma vez registradas em uma estrutura lógica outros indivíduos podem ter acesso a elas.

Somente o registro do conhecimento não significa que este será recuperado; a forma como a informação está armazenada e organizada é fundamental para o processo de recuperação. Para que a informação possa gerar conhecimento, a divulgação do conhecimento ocorre em duas etapas: a primeira, quando o indivíduo-emissor registra o conhecimento em um suporte; a segunda, quando outro indivíduo-receptor faz o resgate do conhecimento registrado e transforma-o em informação, desde que o indivíduo compreenda a mensagem registrada no suporte.

Acredita-se que a disseminação do conhecimento ocorre na segunda fase do processo de comunicação, quando acontece a transferência do conhecimento para o meio de comunicação, e na recepção da informação, que é quando ocorre a aquisição do conhecimento, tornando o receptor um provável veículo e emissor de conhecimento.

A necessidade de novos conhecimentos para encontrar solução para os problemas do cotidiano da sociedade leva a uma busca constante e favorece o surgimento de novas idéias e novas descobertas, contribuindo para a evolução da sociedade; resgatam-se, desta forma, as descobertas do passado, com o intuito de conhecer e compreender os seus sucessos e fracassos.

Essa roda viva, que é o conhecimento gerando novas idéias e novos conhecimentos, tem como elemento básico de transferência a disseminação da informação; do ponto de vista de Cintra (et al, 1994, p.16), nos últimos anos houve um aumento acentuado de informações e uma grande necessidade de armazená-las e recuperá-las nas mais diferentes áreas de interesse.

Atualmente não se consegue calcular quantas revistas científicas são publicadas por ano e muito menos o volume de artigos, dada a dificuldade de recuperar a informação desejada. Tornou-se "mais fácil redescobrir algo que saber se alguém já o descobriu antes" (CINTRA, 1994,p 16). A necessidade de uma solução para esse tipo de problema é incontestável, pois sem isso o desenvolvimento científico ocorre de forma lenta e limitada.

Naturalmente, para cada obstáculo que dificulte o desenvolvimento da sociedade, estudiosos e cientistas concentram seus esforços para desenvolver meios de superá-lo; depois de superados, trabalham para que esses meios evoluam num processo de melhoria contínua.

A necessidade de aprimorar os conhecimentos está aliada à capacidade humana de utilizar o conhecimento adquirido com sabedoria. Quando se trata de conhecimento individual em favor do próprio indivíduo, mesmo que esse conhecimento seja registrado em documentos para atender suas necessidades, a recuperação destas informações é instantânea, pois os sistemas de recuperação são dinâmicos e são processados dentro da mente do indivíduo.

Para o conhecimento coletivo torna-se algo bem mais complexo. O conhecimento do indivíduo que contribui para o bem coletivo deve ser registrado de forma acessível para os outros indivíduos. A grande dificuldade é o registro da informação de forma tal que possa ser de fácil recuperação pelos usuários; mais importante que recuperar a informação, é que esta possa gerar conhecimento para os indivíduos que a pesquisam.

Para a recuperação de informações foram criados sistemas e esquemas para a guarda e organização, os quais são constantemente estudados, objetivando à eficácia do processo de recuperação da informação quando pesquisada. Os descritores de assuntos encontrados em fontes de informações referenciais ou de texto completo são os principais canais de recuperação da informação. Para a recuperação de qualquer informação por assunto é necessário o uso de linguagens documentárias capazes de proporcionar, de forma flexível e consistente, a compatibilidade entre a descrição do assunto de um documento e a linguagem de busca de um usuário que faz associações semânticas e sintáticas simultâneas entre conceitos.

O papel da Representação Documentária é garantir que o conhecimento não se perca. O estudo de formas adequadas de depósito do conhecimento deve ter a preocupação permanente de obter eficiência na recuperação quando o que está registrado for solicitado.

Por tanto, a Representação Documentária destina-se em aplicar a forma mais apropriada de registro do conhecimento em um suporte, visando maior acessibilidade e facilidade de



recuperação do conhecimento registrado nos depósitos; pode-se afirmar que a Representação Documentária é umas das operações da Análise Documentária.

Por isso, a Linguagem Documentária atua no espaço existente entre o emissor e o receptor da informação, concentrando suas atividades nos suportes de armazenamento, como documentos de representação. O objetivo é fazer com que as informações depositadas nesses documentos cheguem até o receptor, e da maneira mais rápida, quando solicitadas; isso faz parte das funções elementares desse instrumento.

O uso da linguagem natural é um grande problema no processo de recuperação de informações, o problema está relacionado ao significado das palavras para os sistemas de recuperação, em função da sua área de especialidade, pois uma mesma palavra pode ganhar sentidos diferentes dependendo da área de especialidade em que foi empregada.

Cabré (1993, p.111) diz que a Análise Documentária é uma disciplina relativamente recente que se ocupa dos documentos que contêm informações. É também uma atividade prática que tem por objetivo facilitar a informação dos distintos especialistas e dos usuários que têm necessidade dela.

Do ponto de vista da Cabré (1993, p.111), Análise Documentária é a recompilação, análise, representação, ordenação e armazenamento de documentos, a fim de que a informação seja recuperada para usos distintos, buscando atender os usuários.

A Linguagem Documentária atua como meio de acesso aos documentos armazenados em determinado suporte. Os usuários que desejam recuperá-los procuram obter o menor número de ruídos ou de documentos não recuperados, conforme a idéia de Cintra (1994, p.24) define que "Linguagens Documentárias são instrumentos intermediários, ou instrumentos de comutação, que realizam a tradução da síntese dos textos".

Segundo Gardin (1968, citado por Cintra, 1994, p.25), uma Linguagem Documentária é um conjunto de termos, providos ou não de regras sintáticas, utilizado na representação de



conteúdo de documentos técnico-científicos, com fins de classificação ou busca retrospectiva de informações.

A finalidade da representação de conteúdos é identificar o conteúdo através de um processo que se inicia pela análise do texto; portanto, a Linguagem Documentária é utilizada na entrada do sistema, quando o documento é analisado para registro.

É importante ressaltar que palavras isoladas podem perder seu significado ou às vezes apresentam tantos significados que acabam perdendo sua identidade. Esse problema é resolvido pelas Linguagens Documentárias através do processo de controle de vocabulário para que cada unidade corresponda a um conceito ou noção.

Em Análise Documentária, portanto, a Representação faz uso das Linguagens Documentárias para o processo de controle do vocabulário com um conjunto limitado de termos, os quais podem ser utilizados para representação do conteúdo do documento. A Linguagem Documentária utilizada pode ser uma lista de cabeçalho de assunto, um esquema de classificação, um tesouro ou simplesmente uma lista autorizada de frases e palavras-chave.

As aplicações práticas dos conceitos e técnicas estão presentes nos tesouros, listas de cabeçalhos de assunto e nos sistemas de classificação bibliográfica.

Esses sistemas visam recuperar as informações que representam o conhecimento da sociedade, com o propósito de assegurar que o conhecimento cresça permanentemente em favor do bem-estar sociedade. A eficácia de um sistema de recuperação deverá estar fundamentada no uso das definições de tratamento das Linguagens Documentárias e se faz necessário que "para existir recuperação da informação desejada é preciso que tanto a linguagem de busca quanto a fonte de informação estejam adequadamente representadas através de uma linguagem documentária". (FUJITA, 1992, p.18).

Segundo Campos (2001, p. 25), atualmente há muitos estudos sobre a problemática relativa à representação da informação e do conhecimento, extrapolando o domínio da

Documentação. Segundo a autora isso pode ser um erro, não se deve desconsiderar as teorias da representação que têm como objetivo a organização do conhecimento.

Os serviços de Análise Documentária objetivam à eficiência na recuperação da informação através do uso de linguagens documentárias para controle do vocabulário e da sintaxe, orientando o indexador sobre quais termos devem ser utilizados para representação do conteúdo do documento. A linguagem documentária de um serviço de análise controla e faz a ligação das relações semânticas existentes entre os termos descritores, além de coordenar os termos que relacionam-se ao tema fazendo uso da sintaxe.

O sucesso da recuperação da informação depende, essencialmente, da lógica adotada para a formação da sistemática no tratamento da informação e a identificação das propriedades existentes nas relações semânticas e sintáticas.

Para Gardin (1973, p. 137-168), uma Linguagem Documentária deve integrar três elementos básicos:

- um léxico, identificado como uma lista de elementos descritores, devidamente filtrados e depurados;
- uma rede paradigmática para traduzir certas relações essenciais e geralmente estáveis, entre os descritores;
- uma rede sintagmática destinada a expressar as relações contingentes entre os descritores.

Entende-se o léxico como o conjunto de palavras de um idioma ou área de especialidade, ou seja, o conjunto de vocábulos da língua de um dado contexto, da qual se reiteram palavras-chave para serem identificadas na lista de elementos dos descritores.

A rede paradigmática refere-se a uma padronização da linguagem utilizada, como forma de manter uma relação entre as palavras cujo significado é de senso comum entre os especialistas da área.

De acordo com Natali (1978, p.37), uma rede sintagmática é lista organizada ou não de relações que podem ser determinadas entre os termos retidos no momento da indexação, a fim de reproduzir a informação sintagmática transmitida pelos textos analisados.

Da mesma forma que numa língua, o vocabulário é a relação das palavras utilizadas, numa Representação Documentária o vocabulário corresponde à relação de termos e/ou símbolos empregados para a identificação temática de documentos. Sendo a sintaxe a estrutura gramatical de uma língua, abrange o conjunto de regras relativas a combinações de elementos que, juntos, possuem significado diferente daquele que apresentam originalmente. Numa Linguagem Documentária a sintaxe refere-se ao conjunto de suas regras de uso para a combinação de termos e/ou símbolos, a fim de expressarem, com a menor ambigüidade possível, o conteúdo temático dos documentos. Estabelece regras sobre o uso dos termos, além de revelar as relações existentes entre os conceitos e os termos de linguagens documentárias.

Considerando as afirmações de Cintra (1994, p.31-32), a configuração interna das linguagens documentárias está estruturada de maneira lógico-semântica; tal estrutura é sustentada pela representação hierárquica, que liga as relações em função da representação do significado no sentido vertical, amparado por unidades informacionais que se relacionam no sentido horizontal e representam as relações não-hierárquicas, denominadas relações associativas.

As relações hierárquicas podem ser genéricas, específicas ou partitivas; as relações desse critério estão baseadas na lógica de superordenação e subordinação, sendo que na relação genérica a superordenação caminha da espécie para o gênero, enquanto que, na relação específica caminha no sentido inverso, ou seja, do gênero para a espécie. A relação partitiva de noção superordenada refere-se a um objeto como um todo e as noções subordinadas a objetos como suas partes.

A estrutura das relações hierárquicas permite subdivisões sucessivas para melhor representar as noções de superordenação e subordinação constitui as espécies e as partes do relacionamento, possibilitando nova subdivisão.

As relações não-hierárquicas referem-se à ligação entre termos que estão em campos semânticos distintos, mas que fazem parte do mesmo contexto, tornando-se próximos e constituindo um conjunto de termos com significados semelhantes. Na linguagem Documentária as relações não-hierárquicas são direcionadas para as relações associativas.

A relação associativa é considerada um dos relacionamentos mais difíceis de serem estabelecidos, refere-se ligação entre "termos que não possuem relacionamento hierárquico e nem de equivalência" (CAMPOS, 1994, p.100-101), esta relação é aplicada como recurso na ligação de termos em uma categoria conceitual.

As relações de equivalência por sua vez operam no nível de sinonímia e da polissemia; estabelecem as remissivas, com a finalidade de encaminhar o usuário para os termos preferidos pelo sistema, caracterizando uma chave de acesso do sistema, em relação aos termos equivalentes informados pelo usuário.

Do ponto de vista de Cintra (1994, p.35), o sistema nocional de uma área constitui o arcabouço de sustentação das Linguagens Documentárias. Pode-se verificar que na construção de uma linguagem para uma área está-se enfocando o contexto para o desenvolvimento de sistema de recuperação, por intermédio do controle do vocabulário da linguagem dos usuários da área de especialidade.

Muito importante ressaltar que a construção de uma linguagem com controle de vocabulário deve estar baseada em uma área de especialidade, pois, ao proceder a elaboração de uma linguagem que abrange mais de uma área, poderá ocorrer a ausência de um sistema nocional, comprometendo o significado e a compreensão dos termos.



De acordo com a norma ISO 1087, um sistema nocional caracteriza-se como um “conjunto estruturado de noções que refletem as relações estabelecidas entre as noções que o compõem no qual cada noção é determinada pela sua posição no sistema”.

Os termos e símbolos são as representações do sentido de um sistema nocional, as quais são influenciadas pelo contexto sócio-cultural. A ISO 1087 define conceito ou noção como “unidade de pensamento constituído por propriedades comuns a uma classe de objetos”.

Uma linguagem enfocada em um contexto, ou seja, uma área de especialidade, naturalmente elimina as possíveis diferenças de significado de uma palavra. Segundo Cintra (1994, p.53), “não podem existir, por exemplo, duas ou mais palavras que se refiram a um mesmo conceito ou uma palavra para designar vários conceitos”.

Ressaltam-se alguns fenômenos da lingüística, reconhecidos pela Linguagem Documentária no controle de vocabulário para tratamento da informação: a polissemia que decorre da organização sintático-semântica, com o intuito de distinguir a plurissignificação como fenômeno geral; a ambigüidade que ocorre quando há mais de uma interpretação em função da plurissignificação, que como polissemia, poderá também levar à ambigüidade.

A plurissignificação não é permitida em sistemas de recuperação, pois a presença de termos com vários significados ou várias interpretações contribui para o aumento do ruído na recuperação das informações, diminuindo o nível de eficiência do sistema. Nesse sentido, Cintra (1994, p.56) afirma que “numa linguagem documentária, tanto a polissemia quanto a ambigüidade devem ser neutralizadas, para que seja garantida a monossemia”.

A sinonímia é uma relação de palavras equivalentes em que a identidade de relacionamento entre os elementos dependerá das circunstâncias. Desta forma, a equivalência é considerada um recurso normalizador importante para a compreensão de uma linguagem documentária.

A sinonímia contempla o conceito de termos quase-sinônimos, considerados formas que não são intercambiáveis em todos os enunciados relativos a um mesmo domínio.

A hiponímia é uma relação parte/todo de termos constitutivos de uma classe e, segundo Lyons (1977, p.235, citado por CINTRA, 1994, p.62) a hiponímia é “a relação existente entre um lexema mais específico, ou subordinado e um lexema mais geral ou superordenado”.

A linguagem documentária que incorpora o controle de vocabulário e, conseqüentemente, o controle dos termos e as relações entre eles, em sua maioria, estão diretamente relacionada ao tratamento terminológico, como recurso importante na construção de uma linguagem.

Através da terminologia, os especialistas em documentação descrevem o conteúdo dos documentos que devem formar parte de uma biblioteca especializada (estando informatizada ou não).

Existem dois níveis de representação de conteúdo do documento, a representação textual e a representação simbólica; a textual é representada pelo resumo e a simbólica é representada pelo índice, o que torna a terminologia importante para ambos.

O conteúdo de um documento é objeto de resumo e de indexação. A indexação é uma operação que consiste em sintetizar o conteúdo informativo de um documento através de palavras-chave ou descritores, que são termos e que podem estar controlados por um tesauro.

A terminologia, através de suas regras e conceitos, contribui para que esse processo ocorra de forma transparente, pois os trabalhos são elaborados em áreas específicas para atender usuários com interesse no assunto da área; sendo assim, os termos são de domínio comum para os usuários que a utilizam.

Isso faz com que o conceito para o tratamento terminológico tenha sua característica dinâmica cada vez mais acentuada. Com o amadurecimento desses conceitos e com a contribuição da informática, esses sistemas terminológicos já existentes poderão ser

representados no computador. Já existem muitos modelos sendo construídos e utilizados para esse fim, com o objetivo de tornar as informações muito mais acessíveis e sua recuperação mais rápida e eficaz.

Diversas pesquisas estão sendo realizadas para a aplicação dos conceitos terminológicos em sistemas de armazenamento, organização e recuperação de informações. Seguindo as definições teóricas da ciência terminológica, os sistemas sempre buscam trabalhar com uma área de conhecimento específica para o tratamento terminológico, com o objetivo de atender às necessidades dos usuários do sistema.

Nas definições de Terminologia e de Análise Documentária encontram-se muitos conceitos com as mesmas diretrizes; suas definições se assemelham e se complementam nos aspectos metodológicos para organização do conhecimento. A terminologia se preocupa mais em estudar a conceituação e a identificação dos signos e os termos adequados para o contexto, enquanto instrumento de análise, a Representação Documentária estuda se a terminologia adotada para o assunto está apropriada ao contexto, para garantir a eficácia da recuperação da informação para determinada comunidade de usuários.

A identificação da terminologia utilizada nos documentos da área de especialidade para a definição de conceitos faz parte da compilação de termos para a elaboração de tesauros. Na elaboração de tesauros, a compilação de termos é a primeira etapa seguida do tratamento de relações entre termos, a qual somente é possível após a identificação e conceituação dos mesmos.

A organização do conhecimento é tratada pela terminologia por áreas de especialidades, formada por termos que representem os assuntos da área usados nas linguagens especializadas para transferência de conhecimento. Segundo Cabré (1993, p.111), isso também “constitui da base para a formulação de textos técnicos, para a tradução de textos especializados e para a descrição, armazenamento e recuperação da informação especializada”.

Por isso, a terminologia padronizada é também um elemento de controle para os documentalistas; e nesse sentido, Cabré (1993, p.112) afirma em sua obra que a terminologia está a serviço da Análise Documentária.

Neste capítulo elaborou-se a fundamentação da área de representação documentária para identificar as características que deverão ser contempladas pelo modelo do sistema de mediação como ferramentas auxiliares, detectando a metodologia aplicada e a forma adequada para armazenagem e recuperação das informações. Observou-se a importância da elaboração dos sistemas em função do volume de documentos publicados, dificultando a disseminação do conhecimento.

Estudaram-se os conceitos para Representação Documentária, constatou-se materiais relevantes que devem ser considerados e incorporados ao modelo final da ferramenta auxiliar de recuperação, com adoção de elementos encontrados na literatura a respeito.

O processo investigatório enfocou as possibilidades e determinações de se adotarem os sistemas e esquemas de armazenamento e recuperação de informações, analisando os obstáculos a serem superados no processo de passagem da Linguagem Natural para a Linguagem Documentária e na recuperação das informações. Relataram-se algumas aplicações práticas dos conceitos e técnicas existentes na Representação Documentária as quais estão em uso e continuam sendo implementadas como os tesauros e os sistemas de classificação documentária.

A necessidade de encontrar subsídios para garantir a eficiência da recuperação das informações pela ferramenta auxiliar de recuperação fez com que se explorassem novas áreas do conhecimento, e com esse objetivo no próximo capítulo apresentam-se os estudos sobre terminologia, para destacar e evidenciar a importância do tratamento do termo, bem como manter o seu significado, para não haver perda do sentido no processo de recuperação.

CAPÍTULO 3

TERMINOLOGIA: E SUA INTERFACE COM LINGUAGENS DOCUMENTÁRIAS NO PROCESSO DE REPRESENTAÇÃO DOCUMENTÁRIA

Este capítulo se propõe evidenciar os fundamentos da terminologia como campo de estudo, relatando através do levantamento bibliográfico a sua reconstituição histórica para demonstrar sua importância na organização do conhecimento humano. Ressalta-se, através da evolução histórica da terminologia, a base conceitual das técnicas para tratamento do termo, considerando que a ciência terminológica estuda também o significado do termo; dessa forma, incluímos o estudo da terminologia nesta pesquisa para destaque de suas características que permitam identificar subconjuntos lingüísticos, para auxiliá-la no estudo da estrutura apropriada para armazenamento das informações e dar subsídios para a modelagem da solução desejada.

Para a formação de uma base sólida de fundamentação científica deste projeto e a obtenção de conhecimentos das estruturas de tratamento de termos, que levem à elaboração de sistemas eficazes de recuperação das informações empresariais, se faz necessário o estudo da evolução da terminologia, expondo da ciência de estudo de termos os conceitos e fundamentos no tratamento dos termos, no que diz respeito ao significado do termo, os níveis de relacionamento entre termos, e termos diferentes com o mesmo significado, introduzindo à pesquisa um caráter interdisciplinar, pois são considerados os conceitos da Linguagem Documentária na análise do contexto dos termos utilizados e a Terminologia com os conceitos de identificação do significado dos termos.

A análise dos instrumentos para contextualização dos termos e seus significados nos setores da empresa está diretamente relacionada com as linguagens documentárias, pois a necessidade de uma política de controle de vocabulário é imprescindível; podem ser encontrados no estudo da terminologia subsídios importantes que contribuem para a identificação de alguns elementos, para constatar necessidades de tratamento diferenciado para diferentes situações

relacionadas aos termos, por exemplo, termos com mais de um significado entre setores da empresa.

Portanto, faz-se necessário procurar os recursos ou mecanismos estudados na área terminológica, para se obter a qualidade da informação recuperada, pesquisando as técnicas de filtragem da informação, com o menor índice de ruído, partindo do ponto de vista de que "o recurso à terminologia dos domínios é, para as linguagens documentárias, fundamental, uma vez que permite identificar o subconjunto lingüístico a partir do qual as palavras passam a significar". (GINEZ DE LARA, 1999, p.108).

Cabe aqui ressaltar que a finalidade dessa pesquisa está relacionada com as dificuldades enfrentadas pelos usuários das empresas, na tarefa de recuperação das informações armazenadas nos bancos de dados empresariais, para suas atividades diárias, mesmo existindo ferramentas de pesquisa, compilação e recuperação dos dados armazenados, esse estudo busca encontrar proposta de soluções relacionadas ao problema das diferenças lingüísticas entre os usuários dos sistemas de informações e as definições técnicas dos dados armazenados.

Uma grande parte das empresas de médio e grande porte possuem unidades de fabricação ou de distribuição espalhadas por diversas regiões do país, com a finalidade de obter mão-de-obra com remuneração menor e aproximá-la dos locais fornecedores de matéria prima, diminuindo, assim, o custo com transporte e tempo desperdiçados na fabricação ou industrialização de produtos.

As empresas que apresentam o perfil acima, normalmente, realizam contratações de profissionais que residem próximo das instalações dessas unidades, considerando que algumas dessas regiões têm renda per capita baixa, e, conseqüentemente, nível baixo de formação cultural, com escassez de profissionais especializados para ocuparem cargos que exigem conhecimentos prévios de conceitos e métodos na área de atuação.

A seleção e contratação da maioria dos profissionais que atuam nas empresas, normalmente, são realizadas com as ofertas de mão-de-obra que o mercado local tem à disposição; conseqüentemente, o preenchimento de algumas das vagas nessas empresas é com profissionais sem experiências, devido ao custo-benefício de seleção dos profissionais com o perfil desejado em outras regiões, o que não é viável financeiramente para as empresas.

Considerando que a construção do vocabulário da linguagem natural do indivíduo está diretamente ligada à sua formação cultural, ao meio social em que vive e à sua capacidade intelectual em adquirir novos conhecimentos. Tais elementos contribuem para o enriquecimento do próprio vocabulário, passando a usar de forma natural no dia-a-dia do indivíduo.

A terminologia assume um importante papel no âmbito social, pois através da aplicação de seus conceitos poderemos trabalhar na recuperação de informações utilizando a linguagem natural de indivíduos das mais diferentes formações intelectuais, pois a terminologia está diretamente relacionada a áreas de conhecimentos específicos para estudo. O desenvolvimento sistemático das bases e métodos de tratamento terminológico atingiu, “o reconhecimento generalizado de sua importância social e política, tanto local como mundial” (CABRÉ, 1993, p.21).

A terminologia como ciência é uma atividade jovem e em pleno desenvolvimento, embora suas origens remontem a tempos antiquíssimos, quando o homem primitivo deu nome a seus primeiros instrumentos de trabalho em suas atividades elaboradas. Como fenômeno lingüístico, portanto, é tão antiga como o próprio trabalho do homem.

No entanto, no início da idade moderna é que aparecem as primeiras intenções de ordenação sistemática de terminologias e nomenclaturas, como conseqüência da crescente necessidade de comunicação sobre o conhecimento cada vez mais especializado. Cabe assinalar, entre estas obras pioneiras, as seguintes:

1 - Versalius (1514-1564), o fundador da nova anatomia, elabora a terminologia de sua disciplina com intenções normalizadoras e racionalizadoras; em química, especificamente na designação de

vínculos, os franceses Lavoisier (1743-1794) e Bertholet (1748-1822) realizam um trabalho significativo para o estabelecimento de nomenclaturas; o sueco Carl B. Linné (1707-1778), definiu as nomenclaturas da botânica e da zoologia, as quais são iguais ao trabalho de Versalius e Lavoisier em anatomia e em química, o qual permanece vigente nestes ramos científicos.

2 - Lavoisier, sem ser lingüista, reconheceu a necessidade de um equilíbrio entre o nível epistemológico de uma ciência e seu nível de expressão lingüística, e disse: “Toda ciência consta, necessariamente, de muitos fatores que a comportam e de idéias que a fazem agir. A palavra deve fazer surgir uma idéia; a idéia deve expressar a realidade: estas são impressões da mesma categoria. Deduz que o aperfeiçoamento da linguagem depende diretamente do aperfeiçoamento da ciência, assim como, a ciência da linguagem, a necessidade da sintonia dessas áreas contribui para o surgimento de expressões precisas para o desenvolvimento de qualquer campo da ciência.

De acordo com Felber (1979, p.64 citado por CAMPOS, 1994, p.129), o precursor dos estudos que propiciaram à terminologia o *status* de área do conhecimento foi o engenheiro austríaco E. Wüster, que nos anos 30, ao organizar a Terminologia de Eletrotécnica com o objetivo de garantir comunicação precisa nesse campo, terminou por desenvolver a Teoria Geral da Terminologia em sua tese intitulada “*Internationale Sprachnormung in der Technik*” - Normalização Lingüística Internacional no Campo da Técnica.

Com a publicação do trabalho de Wüster e com a disseminação de suas idéias sobre a Teoria Geral da Terminologia, que mais tarde foi considerada como parte constituinte da Ciência de Terminologia, várias linhas de pesquisa se formaram e se apresentam refletidas em três escolas clássicas - a Escola de Viena, a Escola de Praga e a Escola Soviética de Terminologia. “Estas Escolas possibilitaram a formação de instituições por todo o mundo” (FELBER, 1979, p.47, citado por CAMPOS, 1994, p.130).

Cada uma das escolas de estudo e pesquisa seguiu uma determinada tendência; a escola de Praga com I.K. Weisenberg como colaborador e continuador da obra de F.Saussure é

especializada na lingüística funcional; realiza pesquisas sobre linguagem normalizada com fins comunicativos. A escola russa dedica-se, principalmente, à normalização e ao tratamento da terminologia do ponto de vista da lingüística, coordenando a teoria e a prática, elaborando vocabulários especializados e dicionários bi e plurilíngües. "Na França, as pesquisas concentraram no estudo do seu idioma, os estudos são tomados em sentido genérico, os franceses ocupam-se do estudo dos neologismos e sua normalização" (CURRÁS, 1995. p. 43).

O interesse sistemático pela terminologia surge simultaneamente, como se disse, em vários países europeus: Áustria, extintas União Soviética e Checoslováquia. A partir desses três núcleos de origem, o estudo e a prática da terminologia se estende numa primeira fase em direção ao oeste (França, Canadá, Quebec) e em direção ao norte (Bélgica, Dinamarca e países nórdicos); em uma segunda etapa, mais recentemente estende-se em direção ao sul (países do norte da África, e África sub-sahariana, América central e meridional, Portugal e Espanha); mais recentemente, em direção ao leste (China e Japão). Cada um desses eixos está marcado pelas características do contexto em que se desenvolve a terminologia e pelas finalidades que deve cumprir nesse entorno.

Auger (1988 citado por CABRÉ, 1993, p.39), estabelece duas grandes tendências em terminologia, segundo as grandes finalidades a que se propõem seus diversos enfoques: a terminologia orientada para a tradução (comum traducional) e a direcionada para a planificação (normal "aménagiste", segundo a denominação quebequesa).

Cabré (1993, p.39), ao considerar três tendências terminológicas baseadas nos acontecimentos dos anos seguintes às definições de Auger, ressalta três tendências terminológicas que se destacam em seus estudos: a terminologia orientada ao sistema lingüístico (corrente lingüística-terminológica); a terminologia orientada à tradução (corrente traducional) e a terminologia orientada à planificação lingüística (corrente normalizadora).

As ciências terminológicas adquiriram importância nas últimas duas décadas, transformando-se em uma área de estudo interdisciplinar, pois esta ciência passou a inter-

relacionar-se com outras áreas de estudo, tornando necessário encontrar uma definição para a palavra terminologia, que represente de forma adequada a abrangência dessa ciência.

A palavra terminologia apresenta-se na literatura, de uma forma geral, com três significados distintos: a) uma lista de termos e seus significados, os termos de uma área de especialidade é um conjunto de princípios teóricos; b) como campo que abrange o estudo científico dos termos de uma área particular de conhecimento em uma certa língua, semelhante à Lexicologia especializada; c) estudo dos princípios teóricos básicos para o trabalho terminológico.

O primeiro significado do termo nos leva ao campo dos dicionários técnicos, dos vocabulários e léxicos. Assim, terminologia é entendida como a apresentação ordenada de um certo grupo de conceitos e termos de uma área de assunto específico. O segundo significado considera o estudo científico do conjunto de termos de uma dada língua, em área especializada. O terceiro significado enfoca o sentido etimológico.

As definições do termo “terminologia” denominam, neste momento, três conceitos:

Terminologia 1 :Ciência Terminológica; Área de conhecimento interdisciplinar e transdisciplinar que trata de conceitos e suas representações (termos, símbolos, etc.);

Terminologia 2 : Conjunto de termos que representa o sistema de conceitos de um campo especializado individual;

Terminologia 3 :Publicações nos quais o sistema de conceitos de um campo especializado está representado por termos.

Sem dúvida alguma, T2 e T3 não representam maiores problemas, já que são precisamente estas as acepções mais conhecidas e divulgadas; falamos da terminologia médica, computacional e criminológica, como um vocabulário especializado pertencente a algum ramo do saber da humanidade e denominamos como “terminologia” a publicação na qual se listam os termos especializados, quer dizer, dos dicionários, glossários etc.

Mas o conceito de T1 não acabou no acervo conceitual de especialista dedicado ao estudo de fenômenos lingüísticos com suficiente profundidade, isso, se observa com maior vigor nos países em desenvolvimento do que nos países industrializados.

Consideramos, contudo, a equivalência “terminologia = ciência terminológica”, muito distante de ser uma definição, já que não oferece nenhuma informação a respeito das características do termo. Na denominação “ciência terminológica” ou “ciência da terminologia”, para concluir, não é mais que uma tautologia, já que “ciência” e “-logia” tem o mesmo conteúdo semântico.

Para a norma ISO 1087, “Terminologia” é o “estudo dos conceitos e dos seus respectivos termos considerados em seu funcionamento social e pertencente a áreas da experiência humana”. Enquanto objeto concreto, as terminologias são um “conjunto de termos que representam um sistema de conceitos de um domínio particular” (ISO 1087).

O ser humano manifesta a realidade por intermédio de *objetos* que são reconhecidos pelas suas características principais, consideradas como as propriedades do *objeto*; as propriedades comuns de um objeto a outros objetos podem ser definidas como esquemas cognitivos. Para a norma ISO 1087, o “objeto” constitui “o elemento passível de ser percebido e que pode ser representado por um termo ou símbolo”. Os termos “*reconhecer*” e “*perceber*” encaminham a análise à ‘descoberta’, supondo a realidade como um dado passível de ser descrita como tal.

Segundo a norma ISO 1087, a característica é uma “representação mental de uma propriedade do objeto e que serve para delimitar seu conceito”. Podemos concluir que as normas tendem a conferir às representações um caráter eminentemente extensional, o que poderia pressupor uma intrínseca ligação entre as palavras e as coisas.

Para que a Terminologia exerça a função comunicativa, pesquisadores estudam procedimentos para atender uma dimensão cognitiva do mundo real, utilizando a Teoria da Referência para tratar os conceitos como elementos do conhecimento. Do ponto de vista de Ginez

de Lara (1999, p.111) “os pesquisadores buscam na Teoria da Referência os procedimentos para relacionar as estruturas de conhecimento à linguagem e definir seus elementos constituintes”.

A definição de Dubuc (1977, p.6, citado por GINEZ DE LARA, 1999, p.112) enfatiza que a Terminologia está voltada essencialmente às funções de expressão e comunicação, destacando as necessidades do usuário como “norteadoras” do desenvolvimento de terminologias específicas. Segundo Dubuc “a terminologia é a arte de identificar, analisar, com o objetivo de criar o vocabulário para uma técnica dada, uma situação concreta de funcionamento, de maneira a responder aos objetos de expressão do usuário”. A análise de Ginez de Lara (1999, p.112) complementa as idéias de Dubuc e descreve que a função da Terminologia mostra que a língua, em si, é insuficiente para garantir uma comunicação unívoca.

Considerando os aspectos teóricos e práticos da terminologia, Dubuc descreve uma definição global da terminologia com maior ênfase no aspecto prático, sem desconsiderar a dimensão teórica; do ponto de vista do autor, a terminologia “é uma disciplina que permite identificar o vocabulário de uma determinada especialidade em forma sistemática, analisar o vocabulário, e, se necessário, criá-lo e normalizá-lo em uma situação concreta de funcionamento com a finalidade de responder às necessidades de expressão dos usuários” (DUBUC, 1992, p.21).

A terminologia é considerada por Sager, Dubuc, Cabré e outros como uma ciência interdisciplinar, pois o seu estudo abrange matérias de outras disciplinas, para a definição de seu campo de estudo; a terminologia importa conceitos e elementos de outras disciplinas, como é o caso da Análise Documentária; alguns pesquisadores observam que estas se complementam na aplicação de conceitos de tratamento terminológico, sintático e semântico.

A aplicação do tratamento terminológico para recuperação de informações somente é possível nesse trabalho porque estamos enfocando uma comunidade de usuários que interagem com os sistemas de informações e entre si sobre uma mesma temática que está voltada para os assuntos ligados às atividades da empresa.

Devemos atentar que o uso da terminologia em um universo de muitas possibilidades de entendimento do significado do termo tornará qualquer sistema de recuperação de informações ineficaz operacionalmente, “A terminologia só tem sentido em relação às linguagens de especialidade e a comunicação” (CABRÉ, 1993, p.37).

Convém observar que no trabalho terminológico de diversos contextos, pode-se obter diferentes perspectivas de análise para um mesmo contexto, “provocando um amplo leque de aproximações de matéria e a sua realização aplicada”. (CABRÉ, 1993, p.37)

Do ponto de vista de Cabré (1993, p.37), seria difícil dar, hoje em dia, uma visão uniforme da terminologia, mas não é tão complicado encontrar uma ordem, como de fato já se faz, um conjunto de supostas bases compartilhadas pelas distintas aproximações a estas disciplinas. De início, devemos distinguir sobre seu estudo e sua prática, para:

- a) os lingüistas, a terminologia é uma parte do dicionário especializado por critérios temáticos e pragmáticos;
- b) os especialistas, a terminologia é um reflexo formal da organização conceitual de uma especialidade e um meio inevitável de expressão e de comunicação profissional;
- c) os usuários (diretos e intermediários), a terminologia é um conjunto de unidades de comunicação, úteis e práticos, cujo valor se mede em função de critérios de economia, de precisão e de adequação;
- d) os planificadores lingüísticos, a terminologia é um âmbito da linguagem onde se deve intervir para reafirmar a existência, a unidade e a permanência de uma língua, e para garantir, mediante sua modernização, sua continuidade como meio de expressão.

Com base nesses diferentes enfoques, “apresenta duas dimensões diferentes, mas relacionadas entre si: a dimensão lingüística e a dimensão comunicativa”.(CABRÉ, 1993, p.37)



A Teoria Geral da Terminologia, desenvolvida por Eugen Wüster, é uma disciplina científica que possibilita uma base para o trabalho terminológico. É a base de todas as Escolas de Terminologia.

O objetivo do trabalho terminológico é a fixação de conceitos, visando à elaboração de definições orgânicas, além de estabelecer princípios para a criação de novos termos e possibilitar, assim, uma comunicação mais precisa entre especialistas de diversas áreas do conhecimento, no âmbito da Ciência e da Tecnologia, em nível nacional e internacional.

O grande avanço dado pela Teoria Geral da Terminologia (TGT) foi sistematizar princípios terminológicos que deram a esta atividade um caráter científico próprio, diferenciando-a da atividade lexicográfica. Para muitos profissionais envolvidos com as questões de denominação, até aquele momento, não estava bem determinado o campo de atuação de cada uma das atividades.

Para a TGT, o trabalho terminológico inicia-se com o conceito, que possui uma unidade de denominação que é o termo. Um termo designa um conceito. Desta forma, é necessário garantir a unificação de conceitos e termos, o que caracteriza a Terminologia como sendo de natureza prescritiva. Porém a esfera do termo é diferente daquela do conceito. O conceito é o significado do termo. Para a Lexicologia, a unidade de trabalho é a palavra, que pode possuir conotações. Na TGT o conceito pertence, sempre, a uma língua especializada.

A identificação dos elementos necessários para garantir a unificação de conceitos e termos é fundamental no estudo das unidades terminológicas para a fixação de conceitos.

Embora, do ponto de vista de Cabré (1993, p.169), o conjunto das palavras especializadas constitua sua própria terminologia, acredita-se que a linguagem natural do usuário em uma comunidade forme um conjunto de palavras que melhora e facilita a comunicação na comunidade, como ocorre no mundo científico, constituindo a terminologia própria da comunidade. “Os

termos, que são as unidades de base da terminologia, designam os conceitos próprios de cada disciplina especializada” (CABRÉ, 1993, p.169).

Ao analisar-se o conjunto de palavras usadas com certa frequência, por diferentes indivíduos de uma comunidade, pode-se considerar que se trata de uma linguagem especializada no que diz respeito ao relacionamento entre os indivíduos da mesma comunidade, quando visa a um objetivo comum. Pode-se incluir nesse conceito a seguinte afirmação: “os termos são conhecidos fundamentalmente pelos especialistas de cada uma dessas matérias, e aparecem com uma frequência muito elevado nos documentos especializados de cada disciplina” (CABRÉ, 1993, p.169).

Ao adotar-se a terminologia como técnica para amparar o desenvolvimento da ferramenta de mediação para a recuperação de informações dos sistemas empresariais pode-se criar um repositório de termos da linguagem natural que, por se tratar de uma comunidade de indivíduos com objetivo comum, poderá tornar-se uma linguagem especializada para o ambiente em que será aplicado, pois tende a solucionar os problemas de comunicação e recuperação de informações na corporação.

Ressalta-se, porém, a necessidade de se levar em consideração o contexto em que um termo é utilizado pelos indivíduos da comunidade, pois, ao isolarmos o termo do seu contexto, poderá perder o seu significado relacionado ao assunto abordado na linguagem natural do indivíduo. Por serem unidade de forma e conteúdo, “os termos, como as palavras do léxico em geral, são unidades de signos distintos e significativos ao mesmo tempo, que se apresentam de forma natural no discurso especializado” (CABRÉ, 1993, p.169).

Os conceitos representam os objetos; portanto, cada conceito designa um termo. É estabelecida uma relação triádica, onde o termo é uma unidade lingüística que designa um conjunto de propriedades ou conceitos e representa um objeto dentro de um contexto determinado.

Na definição de estudiosos, os termos são considerados unidades sistemáticas com a finalidade de representação dos objetos, participando de um sistema lingüístico; do ponto de vista de Larét, o termo recebe outros atributos, que podem ser considerados como representação e a materialização do significado do termo, quando afirma que “os termos são unidades de forma e conteúdo de um sistema lingüístico determinado”, nesta mesma direção Cabré (1993, p.170) confirma que “os termos formam parte de um sistema estruturado”, o que poderá fornecer uma lógica para identificar as relações entre os termos.

Os termos são unidades sistemáticas pelo fato de que um conjunto de termos de um determinado segmento do conhecimento representa a sua organização conceitual; os indivíduos fazem uso das palavras do léxico geral em sua linguagem natural e com frequência de termos de conhecimento comum na comunidade, para obter exatidão, clareza e eficiência na sua comunicação.

O termo pode ser considerado bem aceito, quando representa seu significado com clareza e eficiência e passar a ser incorporado ao vocabulário natural da maioria dos indivíduos da comunidade, possibilitando tal termo estender seu uso numa abrangência regional, nacional ou internacional, tornando-se unânime seu uso no jargão da língua de uma área do conhecimento humano.

A idéia do parágrafo anterior poderá ser aplicada no mundo dos negócios e no mundo empresarial e, mais especificamente, no estudo de caso, objeto desta pesquisa; ao referir-se aos usuários que se comunicam nos departamentos da empresa, e entre os departamentos ou com usuários externos, considera que o ambiente dos usuários contém uma linguagem especializada, mesmo que as falhas de comunicação sejam frequentes por não existir uma normalização e padronização mais científica das expressões empregadas nos documentos e na sua linguagem.

O desenvolvimento tecnológico e o surgimento de novos meios de comunicação contribuirão para o aumento da importância da informação, com a facilidade e agilidade com que uma informação é comunicada; o constante aprimoramento das facilidades de acesso à

informação, por uma quantidade cada vez maior de indivíduos é um fator primordial no aumento do volume de informações nas últimas três décadas do século XX. O interesse pela busca de informações com valor agregado em todas as áreas pode estar ligado a diferentes objetivos, como sociais, religiosos, políticos e econômicos.

A diversidade das áreas do conhecimento e de interesse da sociedade e a concorrência na divulgação de informações de valor agregado estão transformando os meios de comunicação, levando cada canal a especializar-se na divulgação de informações por área de conhecimento a fim de abranger uma parte cada vez maior da sociedade, que também se subdivide em áreas de interesse. Cada canal de divulgação ou de fornecimento das informações busca obter informações de seu público para conhecer seus interesses e satisfazê-los.

A expressão “as indústrias da língua”, embora pareça um tremendo contra-senso, deve admitir que se pode realizar e é um fato real com o qual esbarramos diariamente. Currás (1991, p.51) descreve em sua obra que não somente se industrializa a língua, mas se fala também da engenharia da língua e da linguagem, da informática lingüística informatizada e da lexicografia computacional. Todas essas expressões refletem uma característica de nosso tempo, que faz referência à tecnologia aplicada a atividades cotidianas, sejam elas quais forem.

Segundo Cabré (1993, p.374), de uma forma geral, a indústria da língua abrange três grandes tipos de atividades: a investigação na informática lingüística, a produção e comercialização de programas informáticos e a formação de especialistas e usuários dos novos produtos.

No nosso tempo a indústria da língua está diretamente envolvida em todos os aspectos relacionados à produção, divulgação, armazenamento e recuperação de informações existentes na humanidade e tem interesse no desenvolvimento da comunicação para a circulação da informação com objetivos definidos.

Por outro lado, com o constante aperfeiçoamento de conceitos e técnicas surgem novos recursos no que se refere ao tratamento da informação, favorecendo ainda mais sua reutilização para geração de conhecimento através do processo de armazenamento e recuperação. Áreas de pesquisa concentram-se cada vez mais no objetivo de definir a melhor técnica de armazenamento das informações com a finalidade de obter na sua reutilização a maneira mais eficiente de recuperação, com a tendência de percentual zero de perda e de ruído no processo de recuperação.

A tecnologia oferece facilidade e praticidade para a aplicação das mais complexas técnicas e conceitos de armazenamento e recuperação de informações; na maioria dos casos, muitas técnicas tornar-se-iam inviáveis pelo processo manual, haja vista o tempo necessário para a sua implementação e o espaço físico para armazenamento do material a ser utilizado.

Alguns registros mostram que desde a metade da década de sessenta do século XX a informática e a terminologia estão ligadas ao armazenamento e divulgação de dados terminológicos e na elaboração de grandes bases de dados especializados, chamados de bancos de dados terminológicos; para elaboração e definição da base terminológica é necessário o estudo dos termos mais adequados na estruturação dos metadados, das bases de armazenamento de informações.

De acordo com Currás (1991, p.54), a informática possibilitou a elaboração de diversas aplicações na construção de bancos de dados terminológicos. Nesse processo, atualmente, pode-se contar com as técnicas de inteligência artificial e o desenvolvimento de sistemas especializados, que favorecem a automatização dos trabalhos terminológicos.

Fruto do desenvolvimento tecnológico, os bancos de dados são considerados um dos mais importantes recursos para armazenagem e recuperação de informações disponíveis. São esquemas e estruturas com dispositivos digitais que conservam informações formatadas e padronizadas e que ocupam pouco espaço, proporcionando ganho de velocidade na sua reutilização e recuperação, incluindo o processamento das informações, como filtragem, classificação, verificação e validação de bancos de dados, facilidade de transporte, reprodução e divulgação.

Segundo Currás (1991, p.55), bancos de dados terminológicos são registros de conceitos ou termos, realizados com a finalidade de serem utilizados por pesquisadores ou tradutores, onde os dados são acompanhados de referências, definições e outras informações auxiliares para sua melhor compreensão e utilização.

Do ponto de vista de Dubuc (1999, p.207), as bases terminológicas são bancos de dados especializados que têm por objetivo armazenar a informação terminológica existente e entregá-la aos tradutores para que tenham acesso direto mediante conexão aos discos ópticos compactos nas bases terminológicas.

A definição que se aproxima inteiramente do enfoque do trabalho de pesquisa, a do pensamento de Cabré (1993, p.396), define as Bases Terminológicas como uma recompilação estruturada e automatizada de informações sobre as unidades de significação e designação de uma área especializada. Tais bases terminológicas são destinadas a responder às necessidades de um grupo definido de usuários.

Analisando as definições encontradas nas obras dos autores acima, observa-se que todas apontam para a mesma direção, prevalecendo o uso da tecnologia para armazenamento e recuperação de informações previamente estruturadas.

O tipo de estrutura de uma base terminológica dependerá, em primeira instância, de qual será sua finalidade temática. Do ponto de vista de Currás (1991, p.56) deduz que o registro das informações que farão parte do banco de dados pode ser composto por "conceitos" e "termos", considerando-se que existem vocabulários e dicionários cuja utilização pode ser respectiva ou indistinta.

Segundo Currás (1991, p.56) vocabulários têm por objetivo o registro dos "conceitos" e dicionários, o registro dos "termos", relativos a assuntos tratados de abrangência universal e multidisciplinar.

O objetivo do projeto de pesquisa é incorporar "termos" e "conceitos", pois pretende-se oferecer a uma comunidade de usuários uma interface de interação com a linguagem natural do usuário, em que este poderá fazer a recuperação da informação pela palavra que desejar; neste caso, o mais importante é ter a palavra no banco de dados para poder localizar as informações que são armazenadas pelos sistemas de informação da empresa.

A maioria dos bancos de dados terminológicos do tipo vocabulário são monolíngüe, embora contenham também equivalência em outros idiomas. Por sua vez, os bancos de dados do tipo dicionário são bilíngües sedo a maioria multilíngüe, muito utilizados por tradutores.

Em todas as obras pesquisadas, com relação à construção de uma base terminológica, constatamos que a unidade informativa básica dos bancos terminológicos é a ficha terminológica. A redação de uma ficha deve responder a critérios de validade reconhecidos por profissionais: informações confirmadas por fontes válidas e devidamente acompanhadas com referências, contextos significativos, classificação por áreas e subáreas. O acesso à ficha pode ocorrer por todas as chaves indicadas. Em geral, a chave mais utilizada é a unidade terminológica total ou parcial, em cada uma das línguas de trabalho no banco. O acesso por áreas e subáreas é muito útil para realizar investigações temáticas. O acesso por elemento de unidade terminológica é particularmente valioso quando o usuário não encontra diretamente a solução do seu problema.

Segundo o pensamento de Dubuc (1999, p.209), antes de tomar a decisão de automatizar um fichário terminológico, é necessário identificar claramente seus usuários e definir quanto se quer usar e que tipo de uso se deseja fazer dele. Destas explanações depende a seleção do programa e dos recursos que serão empregados.

Um modelo de ficha proposto por Currás (1995, p.60), para registro terminológico de solicitação e inclusão de termo, poderia ter a seguinte estrutura:

- número de identificação do registro;

- unidade léxica;
- documento(s)¹ onde foi encontrado;
- definição;
- origem da definição;
- campo temático;
- designação do conceito;
- forma léxica de onde deriva;
- raiz e origem da palavra utilizada;
- designação do termo;
- forma léxica do termo;
- documento(s) onde o termo foi lido;
- data deste(s) documento(s);
- equivalente(s) em outro(s) idioma(s);
- data de anotação
- comentário(s) de modificações ou atualizações;
- nome da pessoa que realizou as anotações.

A elaboração de fichas terminológicas para a coleta de dados é uma tarefa muito difícil; será muito mais produtivo se criar um banco de dados onde devam constar os dados de que se necessita para o registro do termo na base terminológica. Esse banco de dados, com os novos termos cadastrados pelos usuários como solicitação de inclusão do termo na base terminológica, deverá passar por uma triagem crítica de um profissional da área de terminologia, para análise dos termos antes da sua efetiva inclusão na base terminológica. "O desenvolvimento da microinformática já permite a automatização destas fichas, abrindo, assim, melhores possibilidades de aproveitamento dos fichários pessoais e da sua pasta em comum". (DUBUC, 1999, p.209)

O relacionamento dos dados deve ser simples. Quanto mais simples, menor a possibilidade de erro, desde que não perca a complexa relação de dependência para contextualizar o termo no

¹ Estes documentos podem ser livros, dicionários, vocabulários, artigos de revistas etc.



ato da sua recuperação. Além disso, a simplicidade do protocolo de ingresso dos dados permite confiar a operação a um pessoal especializado de baixo controle do terminólogo responsável. Esta possibilidade é particularmente útil quando o fichário que será automatizado já contenha um número importante de fichas que se consideram válidas. O programa de relacionamento também deve permitir a verificação ortográfica das fichas, de modo a evitar aqueles erros que podem fazer com que a informação útil não seja encontrada. Por outro lado, deve garantir um controle de validade, confirmando que cada ficha armazenada contenha todos os elementos indispensáveis, definidos pelo terminólogo responsável.

A simplicidade no relacionamento dos dados deve ter unidade e facilidade de importação, de modo que se tenha a possibilidade de transportar diretamente o fichário das informações que venham de outro sistema (fichários, processamento de texto, etc.).

Esta simplicidade no relacionamento dos dados deverá ocorrer nos processos de mediação dos usuários com o sistema, pois, na modelagem do sistema o relacionamento não ocorre de forma tão simples, já que se trata de sistematizar a língua, considerando-se diversos fatores complexos para o armazenamento e recuperação da informação.

Segundo Dubuc (1999, p.210), o acesso à informação deve permitir retirar o melhor proveito possível do material armazenado. Deve-se poder editar as fichas pelas unidades terminológicas declaradas nas línguas de trabalho do fichário. Incluem-se como unidades as abreviações e os sinônimos. Também se deve poder editar os elementos de uma unidade terminológica completa. O uso temático de um fichário exige poder acessar por áreas e subáreas temáticas que figurem nos fichários.

Preocupado com o desenvolvimento de sistemas conceituais, Motta (1987, p.31) ressalta em sua obra o uso da Teoria do Conceito ou "Teoria Analítica do Conceito", que se destina a servir de fundamento para análises conceituais de toda e qualquer iniciativa que diga respeito ao estudo e à produção de termos. Ela implica em que cada conceito tem um referente (seja este um conjunto de objetos, um único objeto, uma atividade, um fato, um tópico etc.), sobre o qual

afirmações verificáveis podem ser feitas. Todas essas afirmações podem ser sumarizadas e/ou sintetizadas por um termo que, então, representará um conceito em qualquer processo de comunicação.

De acordo com as informações de Currás (1995, p.61), praticamente a maioria dos países industrializados e os que se encontram em vias de industrialização dispõem, pelo menos, de uma base terminológica onde se compilam suas terminologias e onde se podem buscar informações sobre estes assuntos, como a Eurodicautom, da Unidade Européia. No número 1, volume 12, da *Revista Española de Documentación Científica*, foi publicado um artigo de autoria de A. Irazazábal, S. Alvarez e J. Zarco, no qual se explica detalhadamente a composição desse banco de dados.

Outros bancos de dados terminológicos conhecidos e úteis são o de INFORTERM, AFTERM, NORDTERM, TERMIUM e vários outros. Para consultar pesquisas atualizadas sobre bancos de dados terminológicos, acessar a *Internet* no endereço <http://orion.ufrgs.br/iletras/termisul/index.html>.

Um dos motivos principais pelo qual se iniciaram os estudos terminológicos, os quais adquiriram tanta importância nestes últimos tempos, foi a necessidade de normalizar as linguagens especializadas, próprias de cada uma das diferentes disciplinas e subdisciplinas dos distintos campos do conhecimento. Todos sabem que uma mesma ferramenta, por exemplo, recebe numerosos nomes nas diferentes regiões de um mesmo país.

A Associação Brasileira de Normas Técnicas (ABNT) é o órgão responsável pela normalização técnica no Brasil, tendo sido fundada em 1940 para fornecer a base necessária ao desenvolvimento tecnológico brasileiro.

No que diz respeito à área terminológica, a ABNT tem atuado por etapas. Em 1990, o IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia) sediou o II Simpósio Ibero-Americano de Terminologia e I Encontro Brasileiro de Terminologia Tecno-Científica. Dois anos

depois, acatando idéias manifestadas durante a realização do Simpósio, o IBICT implantou - em conjunto com a ABNT - a CEETT (Comissão de Estudo Especial Temporária de Terminologia), as normas ISO 639:1991, 10241:1992 e 860:1996 que foram traduzidas e adaptadas ao contexto brasileiro e que, reunindo diferentes profissionais ligados à terminologia (tradutores, documentalistas, professores e pós-graduandos), já apresentaram seus primeiros resultados.

No âmbito do desenvolvimento da terminologia no Brasil, a ABNT coopera com o IBICT e com a União Latina na elaboração de Guias de Fontes de Informação em Rede, projeto que visa, por meio da RNP (Rede Nacional de Pesquisas)/*Internet*, colocar à disposição da comunidade científica e técnica um instrumento visando à alimentação descentralizada dos dados via *Internet*. Esta coleta de dados será utilizada para alimentar o BrasilTerm, - uma experiência brasileira em base terminológica - , e aproveitada para atualizar os dados coletados para o projeto RITerm-BD (feita em 1995 pela União Latina, com apoio da UNESCO, e do IBICT no Brasil).

Neste capítulo pesquisou-se a literatura sobre os fundamentos da terminologia como campo de estudo, com o intuito de encontrar subsídios importantes para a configuração do comportamento do sistema de mediação na recuperação da informação, considerando que a terminologia complementa a área de Linguagens Documentárias. Ressalta-se através da evolução histórica a base conceitual das técnicas de tratamento do termo, descrevendo sua disseminação em escolas de várias continentes, tornando o estudo da terminologia uma disciplina.

Destacou-se como o campo da terminologia ganhou importância e sua contribuição na organização de documentos, atribuindo aos termos as propriedades conceituais e de significados no processo de organização e identificação de documentos para sua recuperação. No desenvolvimento da terminologia, pela Teoria Geral da Terminologia, criada por Wüster, evidenciaram-se algumas tendências que seguem se alimentando da Teoria Wüsteriana.

Elaboraram-se estudos das características da terminologia e relataram-se, com base nas bibliografias analisadas, a função da terminologia, as relações entre conceito e termo, os termos e



as unidades sistemáticas, esses itens compõem a estrutura conceitual da ciência terminológica como referência para ser utilizada com uma característica interdisciplinar.

O processo mercadológico está contextualizado neste trabalho, ao se fazer referência às indústrias da língua, em se tratando de desenvolvimento tecnológico e aumento da importância da informação, tanto pela comunidade científica como por outras áreas de interesses. Com a necessidade do estudo da aplicação da tecnologia nessa área, elaboraram-se pesquisas literárias sobre a construção da base terminológica – considerada como base elementar para o sistema de mediação -, bem como, os tipos de bancos de dados para conhecimento dos esquemas e procedimentos, para a construção da base terminológica do sistema de mediação.

Com a base de fundamentação teórica apresentada neste capítulo e no capítulo anterior e para fazer o uso dos conceitos abordados nesta pesquisa, no próximo capítulo retrata-se a análise de um ambiente dotado de sistemas de informações computadorizados, ressaltando a necessidade de recuperação e os fatores favoráveis para desenvolvimento de ferramentas auxiliares de recuperação de informações empresariais em base de dados informatizados.

CAPÍTULO 4

RECUPERAÇÃO DE INFORMAÇÕES EMPRESARIAIS EM SISTEMAS DE *DATA WAREHOUSE*

Mais importante do que a informação é ter acesso a ela no menor espaço de tempo, o resultado dessa acessibilidade pode significar avanços sociais e econômicos valiosos. Desta forma, as empresas criam estratégias individuais em suas diversas áreas, para diferenciarem-se no mercado onde cada dia mais os sistemas e a tecnologia da informação ganham um lugar de destaque.

Conhecer a tecnologia e o uso de um *Data Warehouse* pode ajudar os empresários a descobrirem novas formas de competir em uma economia globalizada, trazendo melhores produtos ou serviços para o mercado, mais rápido do que os concorrentes sem aumentar o custo dos produtos ou dos serviços.

Adotou-se como base o conhecimento acumulado a partir da experiência adquirida na elaboração de ferramentas de recuperação para atender necessidades específicas de usuários do ambiente empresarial, com o intuito de continuidade nas pesquisas de elaboração de ferramentas de recuperação de informações mais dinâmicas e flexíveis.

Com a evolução da tecnologia de informação e o crescimento da utilização de computadores interligados, praticamente todas as empresas de médio e grande porte utilizam sistemas informatizados nos setores operacionais, gerando com o passar do tempo uma enorme quantidade de dados relacionados aos negócios.

Os sistemas de informações, desenvolvidos para atender e armazenar as informações sobre as atividades dos setores operacionais exercendo como coletor dos dados e oferecendo informações nas consultas realizadas pelos usuários destes setores, são considerados sistemas transacionais, ou

seja, sistemas que registram os dados relativos às transações ocorridas nos setores operacionais das empresas.

Segundo Henry C. Lucas (1987), citado por Silva (2002, p.2), entende por sistema de informação o "conjunto organizado de procedimentos, que, quando executados, produzem informação para apoio à tomada de decisão e ao controle das organizações".

Segundo Stair (1999), citado por Canoglia (2002, p.8) os sistemas de informação baseado em computador são "sistemas que empregam hardware, software, bancos de dados, telecomunicações, procedimentos e pessoas para a coleta, armazenamento, transformação de dado em informações e disseminação".

Os dados armazenados por sistemas transacionais das empresas são a base para a construção de *Data Warehouse*, pois, de modo geral, raramente servem como recurso estratégico no seu estado original. Os sistemas transacionais convencionais não são projetados para gerar e armazenar as informações estratégicas em estruturas dinâmicas, o que torna os dados vagos e sem valor para apoiar processos das tomadas de decisões nas organizações.

As decisões normalmente são tomadas com base na experiência dos administradores e acontecimentos no mercado, enquanto podem ser baseadas em fatos históricos que foram armazenados pelos sistemas transacionais utilizados pelas organizações.

O armazenamento das informações em grandes bases de dados – incluindo as bases de *Data Warehouse* organizadas dinamicamente e possibilitando a recuperação ou consolidação por diferentes elementos portadores da mesma informação, contribuiu para a aquisição de ferramentas de recuperação, principalmente aquelas voltadas para a gestão de negócios, havia a análise dinâmica de resultados e de tendências, controle dos fatos e eventos passa a significar não só a possibilidade de uma maior competitividade como, em muitos casos, a própria sobrevivência das empresas que atuam neste mercado.

A compreensão da definição do conceito dessa técnica de armazenamento não é fácil para quem não é da área de tecnologia, bem como não é fácil a construção da estrutura de armazenamento, pois se trata de um processo complexo.

Segundo Willian H. Inmon (2002, p.1), o "*Data Warehouse* é um banco de dados orientado por assunto, integrado, não volátil e histórico, criado para suportar o processo de tomada de decisão", desta forma, permitir múltiplas formas de acesso às informações sobre o andamento dos negócios através da compilação dos dados em uma estrutura apropriada.

Do ponto de vista de Adriano Dal'Alba (2002, p.1), um *Data Warehouse* pode ser definido como um banco de dados especializado, o qual integra e gerencia o fluxo de informações a partir dos bancos de dados corporativos e fontes de dados externos à empresa.

Segundo Marempudi (2002, p.1), *Data Warehousing* é uma técnica que liga componentes de *software* e *hardware* usados para melhorar a tomada de decisões em empresas a partir da análise de grandes volumes de dados cumulativos.

As informações que alimentam os armazens de dados são originários dos sistemas de informações transacionais das empresas, Kimball (2002, p.1) caracteriza um *data warehouse* como "uma cópia das informações transacionais estruturada especificamente para consultas e análises".

Os *Data Warehouse* são construídos para que as informações introduzidas e geradas nos sistemas transacionais possam ser armazenadas em uma estrutura dinâmica, permitindo o acesso por diversos enfoques, não permanecendo limitadas por tabelas e linhas estritamente relacionais.

O surgimento dessa técnica se deu com a evolução tecnológica mundial, dessa forma, viabilizou a utilização de computadores que oferecem recursos de armazenamento das informações em discos magnéticos com grande velocidade de recuperação.



As empresas, até início da década de 80, consideradas de pequeno e médio porte, não tinham condições de investir em tecnologia de informações, pois existiam somente computadores de grande porte, cuja aquisição exigia um investimento tecnológico considerado elevado, muito acima das condições econômicas das empresas de médio e pequeno porte.

De modo geral, a década de 80 foi um período de grande crescimento econômico para o país, em particular para o setor industrial. O crescente desenvolvimento tecnológico na década de 1980 contribuiu para o surgimento dos primeiros microcomputadores a um custo acessível às empresas de pequeno e médio porte, oferecendo baixo investimento na aquisição dos equipamentos, baixo custo na infra-estrutura para instalação dos microcomputadores, por se tratar de equipamentos de fácil manutenção.

Devido ao grande crescimento das atividades industriais e comerciais e ao aumento considerável do volume de informações entre os setores e das empresas com o mercado através da troca de informações com os seus clientes e fornecedores, fez-se necessária a aquisição de tecnologia. O aumento da oferta de tecnologia no mercado possibilitou as empresas fazerem a aquisição de diversos microcomputadores, para informatizar os setores mais críticos das empresas, considerando o custo acessível dos computadores.

As informações de experiências de sucesso e fracasso na aplicação das estratégias adotadas pelas empresas estão armazenadas em banco de dados, em diversos arquivos e tabelas criados para serem alimentados a partir das transações e ações executadas no dia-a-dia das empresas; para obter os resultados dessas experiências são necessários o uso de técnicas de extração de dados e a inteligência do usuário da informação para formatar a apresentação das informações para análise de comportamento dos negócios da empresa.

A formatação da apresentação das informações exige habilidades de um profissional com conhecimento nas duas áreas, ou seja, na área de negócios da empresa e do conhecimento do banco de dados; com a união do conhecimento dessas duas áreas, é possível obter informações valiosas para análise dos negócios da empresa.



Devido à complexidade de uso das ferramentas de extração de informações, torna-se difícil um profissional da área de negócios fazer a extração de dados com eficiência; por isso a cada dia busca-se mais o desenvolvimento de ferramentas de garimpagem e compilação dos dados cada vez menos complexos para possibilitar o acesso a essas técnicas de garimpagem.

As técnicas de garimpagem continuam evoluindo e ainda há um longo caminho a percorrer, pois as ferramentas disponíveis no mercado fazem com que os usuários da área de negócios ainda dependam de profissionais da área técnica para garimpar as informações necessárias para elaborar novas estratégias ou analisar e verificar os resultados das estratégias aplicadas.

Para maior dinamismo dos usuários das informações, várias ferramentas de consulta e garimpagem de dados estão sendo desenvolvidas, com o objetivo de oferecer recursos de recuperação das informações pelos usuários, possibilitando, com o uso das ferramentas, elaborar pesquisas e criar, como resultado da pesquisa, consultas tanto na tela do computador como em relatórios.

Essas ferramentas de garimpagem de informações permitem aos usuários definir regras de filtragem dos dados necessários para que se apresentem os resultados da situação desejada e se permita formatar o *lay-out* de apresentação da consulta, bem como a ordem de apresentação, a posição das colunas, totalizadores, etc.

Para os usuários começarem a dominar essas ferramentas, faz-se necessário o uso intenso das ferramentas para que adquiram conhecimento e prática na sua aplicação. Para incentivar e até mesmo facilitar o uso das ferramentas, é necessário que os usuários detenham conhecimento prévio de como estão estruturadas as bases de dados da empresa e de que forma os dados estão armazenados e quais os nomes das tabelas e dos arquivos de dados.

A obtenção de conhecimento da base de dados das empresas é uma tarefa que exige tempo e paciência dos usuários e isso é o que eles menos têm; precisam, pois, de rapidez na tomada de decisões, já que o mercado se torna cada vez mais exigente.

A linguagem utilizada na alimentação dos dicionários de dados é geralmente a dos técnicos em informática que, via de regra, não utilizam os mesmos termos dos usuários; isso dificulta a recuperação da informação pelos mesmos, por não saberem em qual estrutura está armazenada a informação, ou pela falta de recursos para descobrir com qual termo foi definida — quando o termo é diferente do procurado pelo usuário.

Dessa forma, existe a necessidade de se encontrar uma solução a fim de que os usuários possam obter respostas para as suas dúvidas a respeito da informação que desejam recuperar, sem ter que recorrer ao profissional de suporte ou aos analistas de sistemas, os quais, na maioria dos casos, estão sobrecarregados com outras tarefas, ficando os usuários limitados em suas atividades e na dependência da informação.

Como as ferramentas de consulta para recuperação dos dados utilizam naturalmente o nome nativo das tabelas e campos de dados, a dificuldade dos usuários é localizar no dicionário em qual tabela está a informação desejada.

Faz-se necessário conhecer o ambiente de processamento das informações pelos sistemas informatizados de uma empresa e o alto grau de complexidade para que essas informações possam ser recuperadas pelos usuários e para saber o envolvimento dos usuários com os sistemas. No capítulo cinco diagnosticar-se-á o perfil dos usuários que operam complexos sistemas de informações para saber o quanto é necessário o uso de ferramenta de auxiliar de recuperação.

Considerando o projeto de “Alternativa Metodológica para Design e Desenvolvimento de *Data Warehouse* com base em *Data Mart*”, implementado em uma empresa como análise complementar pode-se concluir que uma das características de que as ferramentas necessitam é uma interface amigável para que sua utilização pelos usuários seja de grande eficiência e total independência dos profissionais da área de tecnologia da informação.

O uso das ferramentas de consulta ocasiona um grande impacto no mundo dos negócios, principalmente com possibilidade de fazer análises das informações recuperadas nos diagnósticos

dos resultados e previsão de tendências, gerando a necessidade imperiosa de se ter, por um lado, uma metodologia mais eficiente de armazenamento, de relacionamento e de limpeza das bases em função do volume de dados gerados ao longo do tempo e, por outro, o desenvolvimento de uma ferramenta com apresentação amigável aos usuários.

Um *Data Warehouse* oferece os fundamentos e os recursos necessários para os Sistemas de Apoio à Decisão (SAD) eficiente, fornecendo dados integrados e históricos que servem desde a alta direção, que necessita de informações mais resumidas, até as gerências de baixo nível, onde os dados detalhados ajudam a observar aspectos mais táticos da empresa. Nele os executivos podem obter de modo imediato as respostas para perguntas que normalmente não possuem respostas em seus sistemas operacionais e, com isso, tomar decisões com base em fatos, não com intuições ou especulações.

Conforme Sant'Ana (et. al., 1997, p.2), destacam-se a seguir dois fatos importantes para base de domínio tecnológico, quando se desenvolve um projeto de *Data Warehouse*:

- obtém-se a oportunidade de avaliar, em diversas bibliografias pesquisadas e da análise de algumas ferramentas, uma metodologia própria de desenvolvimento de sistema, que atenda às necessidades de um SAD, na solução de problemas de recuperação de informações nas bases de dados;
- obtém-se a convicção do alto valor agregado aos negócios corporativos através da implementação de sistemas de *Data Warehouse* que permita às empresas apresentar diferenciais eficientes dentro de um mercado extremamente competitivo.

Duas grandes necessidades são evidenciadas nos projetos de *Data Warehouse*, a primeira está atrelada à gestão da base dos dados derivados e a segunda às ferramentas de acesso e manipulação dos dados derivados ou dados do ambiente SAD.

Estudos sobre a construção de um *Data Warehouse* mostraram que além de atender às necessidades futuras, faz-se necessária uma mudança de arquitetura. No núcleo do projeto do

armazém de dados deve-se abranger fundamentalmente duas espécies de dados: dados primitivos e dados derivados.

Os dados primitivos são informações detalhadas geradas e armazenadas pelos sistemas transacionais das empresas e os dados derivados são informações sintéticas ou resultados de cálculos produzidos para atender às necessidades das gerências das empresas.

É importante destacar que um *Data Warehouse* possui dois componentes básicos: a informação de eventos históricos e as ferramentas de análise de acompanhamento de estratégias a partir da compilação das informações armazenadas. A necessidade fundamental para qualquer armazém de dados é que seja flexivelmente seguro para suportar mudanças e análises da organização.

Portanto, na elaboração de um *Data Warehouse* se faz necessário o uso de alta tecnologia para possibilitar a implementação de recursos para gestão e ferramentas de consulta da base de dados da organização.

O *Data Warehouse* deve ser capaz de responder a consultas avançadas de maneira rápida, sem deixar de mostrar detalhes relevantes à resposta. Para isso ele deve possuir uma arquitetura que lhe permita coletar, manipular e apresentar os dados de forma eficiente e rápida.

Estruturalmente um *Data Warehouse* deve oferecer diversas formas de acesso à informação. Para exemplificar, pode-se considerar que uma informação é o núcleo acessível por várias vias, sendo que, para isso, essas vias devem estar relacionadas diretamente com o contexto em que o núcleo foi gerado.

Oferecer o total relacionamento entre as informações acessadas em um modelo dimensional por cada via, quando a estrutura dos arquivos estiver organizada de tal forma que possa ser representado por um cubo, é a principal característica de um sistema de apoio da decisão.



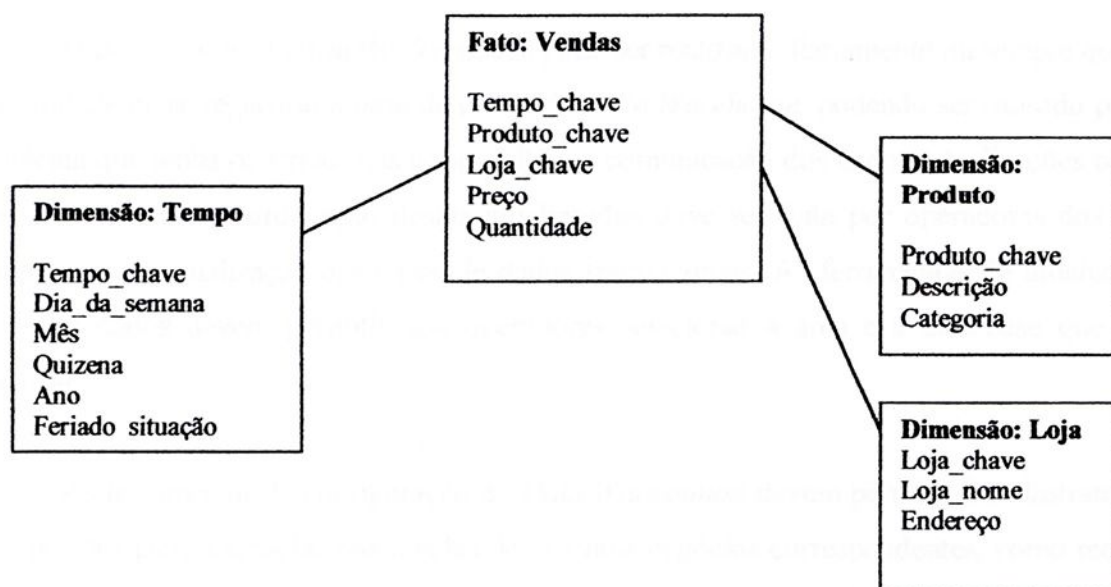


Figura 1 – Modelo dimensional típico

Fonte : KIMBALL, R. *The data warehouse toolKit - practical techniques for building dimensional data warehouse*, John Wiley & Sons, Inc., p. 11, 1997.

Um *Data Warehouse* contém em sua estrutura um modelo dimensional para cada entidade-fato, abstraída dos processos existentes nas empresas, uma função para cada área de negócio. Cita-se como exemplo, a entidade-venda que pode estar representada no *Data Warehouse* pelas dimensões tempo, produto e loja; a entidade-venda é considerada neste caso da a tabela central. As dimensões são tabelas simples que são relacionadas à tabela-fato, fornecendo várias vias de às informações armazenadas na tabela central.

Os ambientes de *Data Warehouse* exigem um cuidado muito especial na gestão da base de dados derivados, em função das transações executadas nos sistemas de informações transacionais e pela atualização dos dados para que as diversas dimensões, principalmente a *dimensão tempo*, estejam retratadas corretamente.

Para construir um *Data Warehouse* eficiente, deve-se considerar que os dados provenientes de vários sistemas podem conter redundâncias e diferenças, e não simplesmente descarregar ou copiar os dados dos sistemas atuais para um banco de dados maior.

O processo de atualização dos dados pode ser realizado diariamente ou sempre que houver necessidade de se regenerar a base de dados do *Data Warehouse*, podendo ser causado por algum problema que tenha ocorrido, tais como falha na comunicação dos dados, atualizações com datas retroativas, etc. A coordenação destas atualizações deve ser feita por operadores dos sistemas sempre após a atualização das bases de dados transacionais. As ferramentas de atualização das bases de dados devem permitir aos operadores selecionar a área e a data-base que desejam atualizar.

As ferramentas de configuração do *Data Warehouse* devem permitir o cadastramento das informações parametrizadas nas tabelas de assuntos/negócios correspondentes, como recurso aos analistas de SAD, com a finalidade de formar um repositório de informações que auxiliem na análise dos dados gerados em um determinado período e que possam ter impacto nos negócios e nas decisões a ser tomadas.

Como exemplo cita-se: se o gráfico da curva de faturamento de um determinado produto, em um determinado período e em uma determinada região, teve um incremento de x%, as razões de tal oscilação poderão estar muito provavelmente atreladas ao evento “prom_fim_ano” (promoção de fim de ano com redução de preço de venda) ou “desc_regional” (desconto pela quantidade vendida), que foi inserido no cadastro de eventos.

Tal necessidade se prende ao fato histórico, isto é, por ser o *Data Warehouse* uma base de dados que retrata a dimensão tempo sob vários ângulos e períodos, torna-se fundamental que haja uma correlação de eventos com os fatos para que as análises dos gestores do SAD estejam fundamentadas em informações que minimizem suas margens de erro.

Após a inserção dos eventos na tabela, os analistas de SAD permitem fazer o vínculo dos mesmos aos fatos (compras, vendas, produção, etc.), para que no momento das consultas em uma determinada área, exista a possibilidade de acesso ao repositório de informações cadastradas nos eventos, para que possa auxiliar os usuários tanto na análise dos resultados passados como na possibilidade de se definirem tendências.

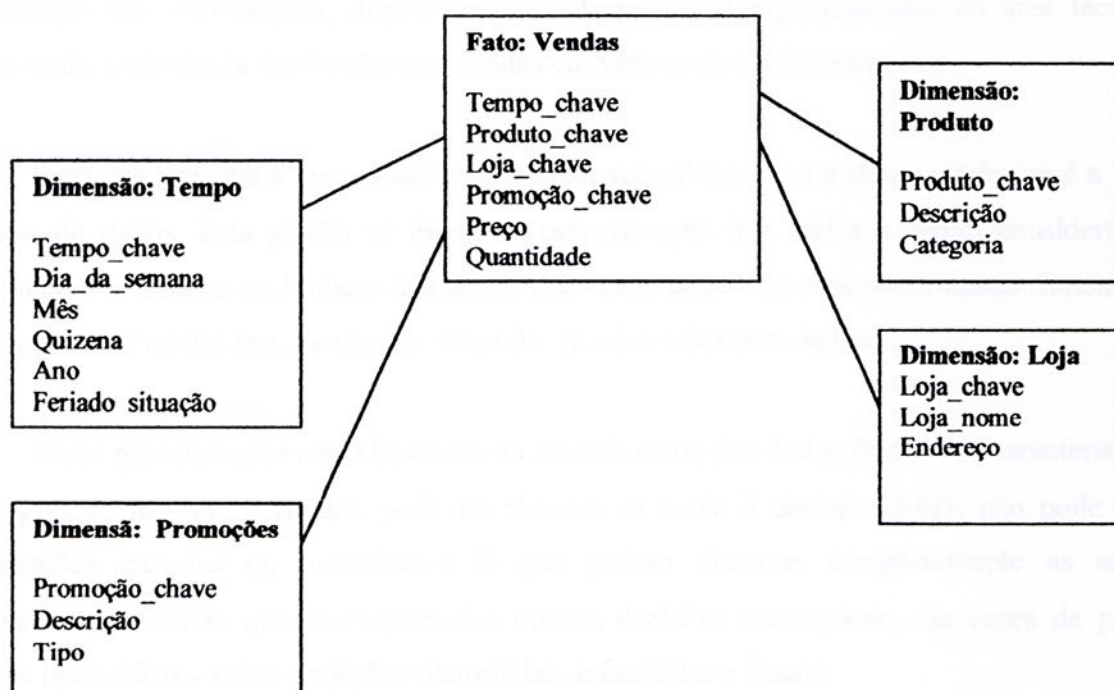


Figura 2 – Modelo dimensional com entidades eventuais

Fonte : KIMBALL, R. *The data warehouse toolKit - practical techniques for buildin dimensional data warehouse*, John Wiley & Sons, Inc., p. 27, 1997.

À medida que o usuário faz as suas opções, as ferramentas de interface de *Data Warehouse* geram internamente códigos SQL (Struct Query Language) os quais acessarão as tabelas selecionadas, retornando uma listagem com os dados definidos e filtrados, disponibilizados por recursos da própria ferramenta, permitindo que o usuário possa escolher também uma “visão geral” em forma de gráfico.

Uma característica importante desta consulta é que os usuários, através de um relacionamento previamente definido entre os fatos e eventos, podem pesquisar sobre, por exemplo, um determinado ponto da curva dentro da dimensão tempo e escolher quais os eventos empresariais ou de mercado que ocorreram no período e que poderiam auxiliar o analista na apuração dos resultados do período.

As ferramentas de consulta permitem uma certa independência dos usuários na recuperação das informações, dispensando a intervenção dos profissionais da área técnica e aumentando a eficiência das tarefas executadas em vários setores das empresas.

Uma das principais preocupações dos desenvolvedores de um *data warehouse* é a gestão da base de dados. Esta gestão se inicia na especificação dos dados a serem considerados e manipulados e termina na limpeza dos dados derivados dentro de uma determinada dimensão de tempo, passando pelas ferramentas de carga das tabelas e dos metadados.

Essas preocupações estão baseadas na autenticidade dos dados dentro das características e das regras de negócio definidas, pois um sistema de apoio à decisão (SAD) não pode conter informações inexatas ou inexistentes já que podem distorcer completamente as análises realizadas, permitindo que as corporações tomem decisões estratégicas - às vezes de grandes valores financeiros - sobre projeções distorcidas, infundadas e irreais.

Do ponto de vista de Sant'Ana (et. al., 1997, p.6), o módulo de consulta do *Data Warehouse* é considerado como gestor da base de dados e tem como característica principal, garantir que :

- os dados primários oriundos das tabelas dos sistemas transacionais devam ser devidamente consistidos, filtrados e formatados pelas ferramentas de extração do *Data Warehouse*, de acordo com o requisitos definidos para as tabelas de dados derivados dentro de cada área de negócio;
- a limpeza dos dados das tabelas do *Data Warehouse* deva ser executada dentro dos requisitos estabelecidos pelos usuários, considerando a dimensão tempo definida para que não haja inconsistências entre as escalas de tempo das várias tabelas relacionadas pelas ferramentas de consulta.;

- os fatos estejam relacionados aos eventos em uma tabela que serve de repositório de informações adicionais para que os resultados das pesquisas possam estar vinculados aos fatos de negócio de forma a elucidar os dados processados;
- as tabelas definidas no *Data Warehouse* apresentem um relacionamento eficaz entre elas, através de uma modelagem e plataforma adequadas que estejam compatíveis com o volume de informações armazenadas, sem perda da eficiência requerida pela ferramenta de consulta - garimpagem – as quais são utilizadas pelos níveis gerenciais;
- as regras formuladas para tratamento dos dados de exceção estejam eficientes e seguras para que a atualização das tabelas definidas no *Data Warehouse* espelhe os dados da forma desejados;
- os níveis de agregação dos dados e as suas dimensões estejam atendendo os requisitos definidos pelos analistas de negócio.

O *Data Warehouse* deve oferecer ainda um nível adequado da granularidade de dados para que as bases de dados tenham, por um lado, as dimensões suportáveis pela tecnologia aplicada e, por outro, atendam aos requisitos dos usuários nas consultas a serem executadas.

A estrutura das tabelas deve ser normalizada de forma a abranger as dimensões que compartilham o mesmo fato para permitir pontos de vista de acordo com a dimensão selecionada para compilação e filtragem dos dados.

Toma-se, como exemplo, a "entidade fato" que resulta de um evento da atividade diária de uma empresa qualquer, o evento "vendedor efetuar uma venda" envolverá o registro dos dados desse evento nos sistemas transacionais da empresa para atender a suas necessidades básicas o qual posteriormente será importado para os armazéns de dados em uma estrutura dinâmica; nesse caso, as propriedades do evento no armazém de dados são compostas de: quantidade, valor, valor da comissão, peso. Vejam-se as definições da estrutura das demais tabela no anexo "A".

Cada registro depositado nessa tabela está ligado simultaneamente às dimensões compostas por entidades de cadastro que se relacionam com o processamento do fluxo da venda em todas as fases para atender o cliente; citam-se como dimensões do evento: tempo, filial, localidade, vendedor e produto. Vejam-se as definições dos atributos das dimensões de cada entidade no anexo "B".

Evidentemente que as tabelas inseridas no *Data Warehouse* são modeladas e estruturadas dentro dos princípios de afinidades com os negócios, isto é, somente fazem ou farão parte da base de dados as informações oriundas dos sistemas transacionais ou mesmo de dados externos que possam agregar valor aos resultados corporativos dentro dos conceitos de um ambiente SAD.

As consultas são realizadas por uma interface pré-definida que explora os dados armazenados, utilizando funções de seleção e filtragem para cada uma das entidades definidas. Como regra de gerenciamento da base de dados está o período histórico das informações disponíveis relacionadas à dimensão tempo.

Evidentemente que inúmeros problemas são encontrados, tanto na base de dados dos sistemas informacionais, como na base de dados do ambiente SAD, que devem ser analisados para adotar uma solução como o *Data Warehouse*. Os problemas mais comuns são os seguintes:

- a) Normalização e Estruturação da Base de Dados Informacional;
- b) Ferramentas de Desenvolvimento;
- c) Tipos de Arquivos da Base de Dados;
- d) Plataforma de Hardware;
- e) Comunicação de Dados;
- f) Qualificação dos Usuários.

Para atender às necessidades dos usuários de recuperação das informações em um ambiente de SAD, a ferramenta e a base de dados do *Data Warehouse* devem estar modeladas para preencher os seguintes requisitos:

- possibilitar aos gestores das empresas acessarem “uma base de dados” que contenha dados integrados, com granularidade móvel, considerando as dimensões tempo, cidade, região, estado, produto, etc., que estejam relacionadas aos negócios das empresas e que possam auxiliar os executivos nas diversas análises sobre resultados passados ou sirvam de parâmetros para projeções e tendências;
- permitir uma manuseabilidade amigável da base de dados, em função das características dos próprios usuários e do acultramento deles em relação ao ambiente SAD, exigindo novas habilidades no perfil profissional dos usuários;
- oferecer maior independência em relação ao desenvolvimento dos seus trabalhos de pesquisa à base de dados SAD, prevendo na curva de aprendizagem uma exigência de ferramentas de pesquisas mais sofisticadas como relacionamentos, agregações, seleção e filtros para sustentar as pesquisas aleatórias definidas *ad hoc*.

Os bancos de dados oferecidos por vários revendedores no Brasil visam tornar compatível o acesso de ferramentas de consulta e de relatórios de fornecedores diferentes, oferecendo maior flexibilidade ao acesso das informações contidas no *Data Warehouse*.

Para o controle da base de dados de um sistema de *Data Warehouse*, a fim de garantir a eficiência do acesso às informações, devem fazer parte de suas funcionalidades os objetivos do sistema as seguintes atividades:

- importação dos dados dos sistemas transacionais, armazenando os fatos consolidados com dimensões relacionadas;
- gestão da base de dados;
- integridade dos dados para acesso de ferramentas de consultas e relatórios;
- fornecer de forma descritiva os eventos ou atividades que causam impacto nos valores armazenados;
- objetivar a implementação de novos fatos de negócios a serem incorporados ao *Data Warehouse*;

- limpeza das bases de dados histórica.

As ferramentas de consulta se baseiam em acesso à base de dados com conceito de *Data Warehouse* para pesquisa e análise de histórico e tendência dos negócios e atividades da empresa.

As ferramentas de consulta para acesso a qualquer base de dados do *Data Warehouse*, utiliza o conceito multidimensionalidade, ou seja, tabela dos fatos consolidados e as tabelas de dimensões relacionadas aos fatos.

Como características os sistemas de *Data Warehouse* devem apresentar de forma amigável aos usuários os seguintes recursos:

- facilidade de acesso para consulta a dados contidas no *Data Warehouse*;
- permitir a seleção de colunas de dados para análise;
- permitir a seleção de filtros de dados para análise;
- visualização do resultado da pesquisa em forma de relatório e de gráfico;

Existem vários níveis de usuários que têm a função de registrar e utilizar as informações que circulam nas empresas para os sistemas de informações cabe aqui destacar que a maioria dos usuários não está devidamente qualificada, para trabalhar em sistemas de informações empresariais, por não lhe darem a devida importância ou pela falta de investimento adequado por parte das empresas de médio e pequeno porte em treinamento e cursos de aperfeiçoamento aos profissionais, tais atitudes prejudicam a qualidade e a confiabilidade das informações armazenadas nos bancos de dados das empresas.

Este trabalho está fundamentado em acontecimentos presenciados em uma empresa de porte médio, assistida por profissionais que trabalham com tecnologia da informação com a função de dar suporte técnico, operacional e treinamento aos usuários.

Os sistemas de uma forma geral estão em constante evolução e cada vez com mais recursos, oferecendo aos usuários maior facilidade e praticidade na sua utilização, aumentando-

lhes o conhecimento dos negócios das empresas e permitindo-lhes sugerir idéias para melhorar e aprimorar os sistemas, objetivando a redução de tarefas redundantes realizadas por pessoas distintas com finalidades diferentes, deixando eles de ser simples usuários para se tornarem usuários-colaboradores, no exercício contínuo da melhoria dos processos empresariais, para aumento da qualidade dos serviços e de seus produtos.

O mercado cada vez mais competitivo exige que as empresas tenham um bom conhecimento de seus concorrentes, mais que isso, pleno conhecimento dos resultados de suas atividades internas, esta tarefa somente é possível através da exploração das informações armazenadas nos computadores das empresas.

Desta forma, evidenciou-se a necessidade de ferramentas de garimpagem das informações armazenadas. Um dos obstáculos a ser vencido é a adaptação dos usuários a essas novas tecnologias, por não apresentarem ainda uma interface de fácil utilização, assim, os usuários necessitam de treinamento especializado para uso das ferramentas, devem também conhecer as estruturas de armazenamento das informações para obter êxito no uso das ferramentas.

Baseando-se em um panorama de complexidades, destacou-se, ainda, o desenvolvimento de ferramentas de consulta menos complexas, com a apresentação dos dados de forma padronizada previamente, com o objetivo de suprir em primeira instância as necessidades dos usuários de acesso a informações compiladas e consolidadas, através da implementação de *dat mart*. Embora esses recursos tenham contribuído de forma significativa, ainda assim, há necessidade de novas compilações e novas consolidações dos dados para análise de resultados das atividades, sendo necessária uma ferramenta de acesso às informações armazenadas nas bases de dados com a perspectiva de atingir a eficácia no uso das informações, disponibilizando ferramentas de consulta com acesso aos bancos de dados de forma nativa.

Muitos tipos de recurso exigem um novo perfil dos profissionais no que diz respeito às formas de recuperação da informação, para obterem informações de que precisam. Enumeram-se neste texto alguns dos requisitos necessários ao novo perfil desses profissionais.



Oportunamente, o diagnóstico dos usuários atuantes frente aos sistemas de informações empresariais, caracterizado no capítulo cinco, apontou as suas virtudes e deficiências e evidenciando a necessidade de investimento em treinamento de pessoal por parte da empresa.

Como o processo de evolução tecnológica é constante, existe a necessidade de um programa de aprendizagem e treinamento contínuo, para que as empresas tenham um produto intelectual apropriado e atualizado para continuarem competitivas em novo mercado o que exige cada vez mais a velocidade de acesso à informação, um dos mais importantes fatores para garantir a sobrevivência das empresas no mercado.

Apostar na criatividade e no intelecto dos usuários tende a ser uma característica fundamental para a sobrevivência e o sucesso das estratégias nos negócios empresariais, para isso os usuários devem ter à disposição grandes variedades de recursos, e informações acessíveis de forma rápida e dinâmica, para concentrar-se mais no poder de criação do que de exploração de informações.

Portanto o grau de sofisticação das ferramentas de recuperação permitindo acesso de forma direta faz com que os usuários conheçam a base de dados. No capítulo cinco detalhou-se a necessidade e dificuldade dos usuários em pesquisarem informações armazenadas nas bases de dados, tornando-se necessária a elaboração de uma ferramenta de mediação entre a linguagem dos usuários e a linguagem das bases de dados, considerando que os esquemas de *Data Warehouse* não prevêem a diversidade lingüística e terminológica.



CAPÍTULO 5

A TERMINOLOGIA DE USUÁRIOS DE BANCOS DE DADOS

DATA WAREHOUSE: RELATO DE EXPERIÊNCIA

Este capítulo baseia-se na experiência adquirida pelo pesquisador, na vivência em um ambiente empresarial, participando ativamente como desenvolvedor de soluções computacionais, implementadas nos processos dos setores operacionais das empresas, objetivando a otimização das atividades setoriais e armazenamento das informações para atender necessidades básicas, como listagem das informações armazenadas para análises superficiais.

Com a evolução dos computadores, conquistou-se de maior capacidade de armazenamento de dados e maior velocidade de processamento na recuperação das informações, na última década o tratamento das informações depositadas nos computadores ganhou grande importância para os profissionais de tecnologia da informação e para a subsistência das empresas.

A evolução tecnológica e o estudo da necessidade de novas técnicas de armazenagem e acesso às informações, introduziram novos conceitos de tratamento da informação nas empresas, com intuito de atender às necessidades dos usuários para análises de resultados e aplicação de estratégias, para alavancar os negócios das empresas no mercado. Também nesse processo o pesquisador atuou inteiramente no desenvolvimento de soluções para atender às novas características, modelando e construindo estruturas dinâmicas de armazenamento para ferramentas de recuperação de informações de apoio empresarial.

Neste capítulo relatam-se algumas das dificuldades enfrentadas pelos usuários na recuperação de informações em sistemas informatizados, enfocando principalmente os problemas lingüísticos detectados no processo de recuperação.



Para melhor entendimento da solução apresentada no capítulo seis, destacam-se neste capítulo características parciais existentes em ambientes empresariais, o perfil dos usuários e as diversas situações que evidenciam as dificuldades de acesso às informações armazenadas nas bases de dados, um problema que necessita de uma solução científica eficiente.

Para maior desempenho das atividades dos usuários no processo de recuperação, ferramentas de consulta e garimpagem de dados são adquiridas e instaladas nas redes de computadores das empresas, para serem utilizadas pelos usuários, permitindo, no uso das ferramentas, criar consultas tanto na tela do computador como em relatórios.

Para abreviar o acesso dos usuários às informações, implementam-se nas empresas novas técnicas de consulta que consistem na preparação dos dados, na formatação da apresentação dos dados pelos analistas de sistemas para atender às necessidades de consulta dos usuários de diversas áreas de negócios das empresas.

A implantação de um *Data Warehouse* é uma tecnologia que exige grande investimento em tempo e em equipamentos, por esse motivo normalmente os projetos de implantação de *Data Warehouse* são divididos em módulos para serem implantados em etapas, procurando atender aos setores considerados de maior volume de processos críticos, favorecendo a melhoria do desempenho com a obtenção de resultado a curto e médio prazo.

As informações enviadas para base de dados *Data Warehouse*, originam-se dos sistemas transacionais, cujo os dados são introduzidos pelos usuários dos setores operacionais das empresas; a confiabilidade das informações enviadas para o *Data Warehouse* é o resultado das informações armazenadas nos sistemas transacionais. Por tanto, a exatidão das informações armazenadas depende da utilização e introdução correta dos usuários dos sistemas transacionais.

Infelizmente não é sempre que isso ocorre, atualmente, pode-se verificar isso através das atitudes tomadas pelos usuários na tentativa de solucionar pequenos problemas no dia-a-dia; muitas vezes, essas soluções encontradas e executadas são inadequadas, pois podem resolver de maneira imediata o problema, prejudicando a tarefa de outros usuários. O usuário pode também

não estar devidamente preparado para solucionar o problema, agindo dessa forma, pode causar outros problemas para a cadeia de usuários, provocando uma seqüência de erros e correções que prejudica o bom andamento dos processos.

Essas atitudes inadequadas são tomadas por falta de treinamento e, muitas vezes, por falta de orientação adequada; os usuários, ao se depararem com o problema, deixam de recorrer ao seu superior, devido a barreiras físicas, profissionais ou psicológicas, os quais acabam tomando atitudes incoerentes para solucionar os problemas.

Há algum tempo, muitas empresas estão preocupadas com a organização e normalização dos processos, procurando mantê-los mais racionais e coerentes, buscando, assim, uma padronização de procedimentos para que as tarefas básicas não se percam ao longo do tempo. Implantaram os departamentos de organização e métodos com a finalidade de estudar e avaliar os processos das empresas, buscando melhorar a normalização das atividades internas, criando para esse fim as normas de procedimentos que funcionam como manuais de instruções para seus usuários, procurando orientá-los de forma organizada na execução das tarefas.

Ocasionalmente, os usuários deparam-se com tarefas que não fazem parte do seu cotidiano. Esses tipos de tarefas são enfrentados muitas vezes como um problema a ser solucionado pelos usuários e não como um procedimento de pouca ocorrência; nessas ocasiões podem tomar decisões não adequadas para a execução da tarefa e tratadas como a solução de um problema; nesses casos, por desconhecimento do processo, os usuários devem fazer uso das normas de procedimentos, o que normalmente não ocorre.

Essas normas de procedimentos acabaram se tornando alvo de críticas dos usuários, por não participarem, em muitos casos, das mudanças de procedimento impostas pelos departamentos de organização e métodos. O volume das informações contidas nas normas dificulta a consulta rápida para esclarecer dúvidas de procedimentos de uma determinada tarefa; a falta de clareza dos dispositivos mencionados nas normas de procedimentos utiliza terminologia de conhecimento do profissional de Organização e Métodos diferente da linguagem utilizada pelo

usuário final, apresentando um problema de questão da linguagem, fazendo com que os usuários ao longo do tempo percam o interesse em consultar as normas.

Nos últimos anos, o trabalho dos profissionais da área de Organização e métodos deixou de trazer bons resultados devido à velocidade das mudanças que as empresas estão tendo que fazer para atender à dinâmica do mercado. A aplicação de novas estratégias para a sobrevivência da empresa da integração do mercado mundial através da globalização vem dificultando a vida das empresas de médio e pequeno porte, com o aumento da concorrência no mercado interno e a entrada de produtos internacionais a um custo menor produzido por grandes corporações internacionais.

Em função desse dinamismo do mercado, surgiu a necessidade da criação de manuais cuja atualização acompanhe as mudanças dos processos e cuja utilização possa ser de forma prática, rápida e dinâmica; surgiram, então, os manuais eletrônicos chamados de auxiliares que podem ser consultados pelos usuários dos sistemas de informações através de uma tecla padronizada, de qualquer ponto do sistema. Esses manuais auxiliares ao serem acionados trazem as informações sobre os recursos existentes, além, de como e o que o usuário pode fazer naquele ponto do sistema para atingir o objetivo daquela fase do processo.

Os manuais auxiliares são mais um dos recursos tecnológicos que fazem parte da evolução dos sistemas de informações, construída pelas empresas ou departamentos de tecnologia e que exigem investimentos para sua elaboração, mas que ainda não conseguiram atingir o elemento principal dos sistemas de informação que é a evolução cultural dos usuários finais na utilização dos sistemas de informação.

Outro fator que influencia na utilização dos sistemas de informações das empresas é a alta rotatividade de profissionais. Profissionais que saem das empresas em busca de melhores oportunidades ou por qualquer outra razão e que, na fase de desligamento, são obrigados a passar, em pouco espaço de tempo, o conhecimento adquirido nas tarefas executadas para outros profissionais que assumirão a vaga.

Essas transições não são bem sucedidas na maioria dos casos, não só pela falta de empenho do profissional que está se desligando da empresa ou do departamento, mas também por várias outras razões, entre elas o tempo disponível para que o usuário transmissor passe os seus conhecimentos ao usuário receptor, tempo esse nem sempre suficiente para transmitir o que o usuário transmissor aprendeu em anos de atividades.

Enquanto essa transição ocorre, ambos os usuários são obrigados a manter as atividades cotidianas funcionando normalmente, muitas vezes, não permitindo fazer uma pausa ou diminuição do ritmo para a transição de tarefas, isso torna as atividades do usuário receptor muito mais difíceis, tendo que absorver conhecimentos práticos sem muita base teórica das tarefas.

Muitas tarefas, que ocorrem esporadicamente, como estorno de algum lançamento de débito ou crédito, ou cancelamento de um processo de pagamento ou recebimento, ou ainda alteração da alíquota padrão de imposto de uma determinada transação, entre outras, não são levadas ao conhecimento do usuário aprendiz ou receptor; quando essas tarefas ocorrem o usuário transmissor não está mais no setor e acabam se tornando um problema a ser solucionado pelo novo usuário, pois até então ele desconhecia a existência deste tipo de tarefa e desconhecia também a forma de tratamento que deve ser dada a essa atividade.

Normalmente, nestes casos, os usuários acabam criando alternativas não muito adequadas, ou, sem saber, podem ter recebido do usuário que se desligou da empresa alguma instrução de uma alternativa que foi criada por ele e que passa de usuário para usuário como sendo o procedimento correto.

Às vezes o usuário que está saindo da empresa deixa de levar ao conhecimento do usuário receptor qual a importância do seu trabalho dentro de todo o processo e o que representam suas tarefas para a empresa e o que a empresa espera dele.

As empresas estão substituindo funcionários recém-formados com conhecimento tecnológico acadêmico e nenhuma experiência profissional no mercado empresarial, com a perspectiva de que aumente a qualidade das informações e, em consequência, se tornem

informações mais confiáveis. Isso é um fator positivo, mas deixa a desejar pelo pouco conhecimento em negócios empresariais.

Com a quantidade de profissionais com conhecimento em tecnologia da informação disponíveis no mercado, as empresas estão substituindo, também, os antigos profissionais com anos experiência pelos formados em tecnologia, o que contribui mais diretamente para o bom andamento dos sistemas. Isso facilita a implantação de sistemas de informações nos setores das empresas que ainda não estão informatizados; mas de qualquer forma é necessário um período maior de adaptação para os usuários tomarem conhecimento dos negócios e objetivos das empresas.

As atitudes dos usuários despreparados podem causar os mais variados tipos de problemas nos sistemas de informações das empresas, pois as informações passadas por esses usuários são utilizadas por vários usuários. Cada vez mais os sistemas de informações tornam-se integrados e uma informação que normalmente é utilizada por vários usuários com objetivos diferentes pode prejudicar o desempenho dos demais usuários, gerando trabalho adicional para atender a uma outra necessidade e tendo que fazer a correção da informação para que a sua tarefa seja executada corretamente e seus resultados sejam atingidos, quando isso acontece, ocorre a desconfiança da veracidade das informações dos sistemas.

As empresas fazem investimento para construir sistemas de informações cada vez mais seguros, tornando-os mais fechados e limitando a intervenção dos usuários; procuram automatizar todos os itens possíveis de acordo com os negócios praticados pelas empresas, buscando tornar os sistemas de informações cada vez mais inteligentes para gerar informações seguras e confiáveis aos usuários dos sistemas.

A atitude de alguns usuários, de forma inconsciente, pode causar danos financeiros de pequenas ou grandes proporções, sendo que os sistemas de informações das empresas processam esses dados para a geração de informações comerciais que impactam diretamente no dia-a-dia das empresas e nas relações com outras empresas que fazem transações comerciais eletrônicas.

No entanto, não são somente os processos existentes nos sistemas os prejudicados por essas informações indevidas, a implementação de recursos nos sistemas para atender às novas necessidades dos usuários ou de outros usuários das empresas em determinadas áreas ou departamentos acaba identificando a existência de informações inadequadas.

Ocorre que, muitas vezes, algum tempo depois da implantação desses recursos, os profissionais de tecnologia executam tarefas corretivas para a reconstituição ou recuperação das informações, através de processamentos internos de correção de forma automática ou manual, ao analisar o impacto dessas informações indevidas nos sistemas e nos negócios das empresas e ao buscar alguma forma de recuperar os possíveis danos que estas informações possam ter causado.

É importante instruir os usuários sobre o procedimento correto, além de analisar qual foi a necessidade que os levou a tomarem aquela atitude, para que o profissional tecnológico faça o ajuste do sistema para atender à necessidade do usuário de maneira adequada.

Usuários não são somente aqueles que têm a função de entrar com as informações nos sistemas, estes normalmente não fazem o uso do resultado de suas informações mas também aqueles que utilizam o resultado das informações de forma analítica ou sintética para analisar e acompanhar o andamento dos acontecimentos através das informações armazenadas nos sistemas.

Essas tarefas estão sendo incorporadas por profissionais que gerenciam os negócios das empresas, passando a fazer parte do novo perfil de profissional com a chegada da tecnologia cada vez mais presente no dia-a-dia das empresas; essas análises podem ser realizadas em forma de relatórios pré-configurados, recuperados pelos sistemas, criados no momento do levantamento de necessidades.

Faz-se necessário que as empresas passem a dar maior importância à área de recursos humanos e incorporem o treinamento e aperfeiçoamento dos profissionais que trabalham ou irão trabalhar com as informações que circulam nas empresas; é preciso, também, procurar incluir na consciência de cada um o significado da informação e qual o tratamento que um sistema de informação realiza em cada dado armazenado, além de levar ao conhecimento dos usuários os

resultados finais de seu trabalho, mesmo que de forma simbólicas e mostrar-lhes de forma panorâmica o fluxo das informações dentro da empresa, destacando os acontecimentos dentro desse fluxo antes e depois da participação do usuário no processo, bem como, quando se dá a seqüência no fluxo para completar o processo.

Acredita-se que esta é a forma eficiente de levar a conhecimento dos usuários o grau de importância da informação introduzida por eles nos sistemas e nos processos das empresas.

Os sistemas de informações tornam-se a cada dia o mais importante recurso para produção do conhecimento, contudo recuperar informações de forma eficiente está sendo o grande desafio de cientistas e profissionais da informação de todo o mundo. Como produto dos sistemas de informações, empresas especializadas em tecnologia criaram sistemas de armazenamento de informações para tornar mais eficiente a sua recuperação.

Os *data warehouse*, como são chamados os sistemas de recuperação eficiente, amplamente definido no capítulo quatro, contribuem de forma significativa nesta recuperação, mas isso acontece somente dentro de uma estrutura padronizada, com formas de busca e recuperação previamente definidas. Em uma pesquisa com dados novos é necessário recorrer ao profissional da informação, limitando os usuários na criação de suas próprias pesquisas para recuperar informações que atendam às suas necessidades.

Para que a exploração das informações depositadas no banco de dados seja realizada com maior velocidade na elaboração de rotinas de pesquisa, existem no mercado produtos que permitem ao usuário da informação fazer suas próprias pesquisas, sem necessidade de profundos conhecimentos tecnológicos, bastando aos usuários conhecerem a estrutura de como as informações estão armazenadas. Esse é outro desafio, pois envolve a definição da estrutura de armazenamento de informações, elaborada por profissionais da informação com diferentes vocabulários, o que contribui para a pluralidade de termos.

A necessidade de identificar essa pluralidade terminológica em sistemas de informação empresariais torna-se evidente para viabilizar a elaboração de uma linguagem intermediária que

funcione como um direcionador de significados para o sistema de recuperação de informações. A identificação da pluralidade terminológica encaminha para elaboração de uma solução prática de ferramenta de mediação para tratamento de uma linguagem intermediária.

Utilizando esse método e os recursos tecnológicos, é possível sugerir para os sistemas de *Data Warehouse* uma ferramenta de mediação, dotada de linguagem intermediária que realizará a ligação da linguagem natural do usuário e dos bancos de dados para facilitar a recuperação da informação com maior eficiência.

Pode-se idealizar uma estrutura de armazenamento dos termos coletados e identificar as propriedades básicas necessárias na definição da lógica no relacionamento dos termos e seus significados comuns, desta forma, acredita-se que elementos significativos estão sendo reunidos para a estruturação da base terminológica, para armazenagem da terminologia dos usuários e dos sistemas de informações empresariais, tornando sua existência essencial para a definição do sistema de mediação para recuperação de informações.

Os termos coletados da linguagem natural dos usuários poderão ser inseridos na base de dados da ferramenta de mediação de recuperação ou como prova circunstancial da existência das diferenças terminológicas.

A busca de um método para obter dos usuários os termos usados em sua linguagem natural foi a maior dificuldade a ser superada, pois o pesquisador não conhecia nenhum método eficiente que permitisse aos usuários total liberdade para exteriorizar seus pensamentos, sem perda da qualidade de identificação da linguagem.

O método de coleta de dados através do uso do protocolo verbal permite ao pesquisador direcionar o pensamento dos usuários para o assunto de interesse da pesquisa, obtendo dessa forma um grau altíssimo de informações relevantes, tornando o trabalho de coleta muito mais abrangente.

Para obter mais informações sobre técnicas e procedimentos de aplicação do "Protocolo Verbal", consultar a Tese de Doutorado da Nardi (1999, p. 116-156), aí se pode verificar a forma prática de uso dessa técnica.

Através da identificação da terminologia utilizada pelos usuários e o conhecimento quantitativo dessa amostragem, pretende-se propor um modelo de mediação de recuperação de informações como um dos objetivos da pesquisa mencionado anteriormente, que sirva como linguagem de mediação da ferramenta de recuperação das informações empresariais e a base terminológica com os termos usuais e técnicos do ambiente de pesquisa.

A finalidade dessa etapa da pesquisa é a coleta de termos dos usuários mediante aplicação do protocolo verbal, para coleta de termos dos usuários dos sistemas de informações da uma empresa, a fim de compor uma amostra de termos utilizados pelos usuários.

O ambiente empresarial utilizado como coleta de termos de usuários é o mesmo vivenciado pelo pesquisador, destacado no início do capítulo quatro. Trata-se de uma empresa privada de manufaturas, portadora de vários setores informatizados e grandes volumes de operações, com necessidade de recuperação e compilação de informações constantes para análise de resultados.

Realizou-se análise de um processo no ambiente empresarial e pôde-se verificar claramente as informações identificadas pelos usuários por termos diferentes. Usou-se como método de coleta de termos a técnica introspectiva, denominado protocolo verbal, o qual permitiu ao pesquisador direcionar o pensamento dos informantes do sistema para o assunto de interesse da pesquisa, tornando o trabalho de coleta abrangente e obtendo, dessa forma, um conjunto de termos relevantes para uma linguagem intermediária.

Os termos mais utilizados pelos usuários, nas tarefas diárias que dizem respeito às informações armazenadas nos bancos de dados da empresa, têm como objetivo tornar a aplicação de técnicas de recuperação da informação uma operação simples e eficiente para os usuários dos

sistemas de informação; todavia, para que isso ocorra, é necessário conhecer os termos que os usuários utilizam e agregá-los aos dicionários de dados para recuperação das informações.

O foco inicial de análise, nesta pesquisa, é uma situação social bastante específica: funcionários do departamento de vendas de uma empresa que utilizam os sistemas de informações para recuperar informações e analisá-las. Com esse propósito, solicitou-se aos colaboradores a criação de um padrão textual que contivesse as informações para atender a uma necessidade concreta.

O observador-participante atuou com nível de participação moderada, tentando alternar o papel de observador e de participante; também interagiu com os usuários na sessão de gravação de protocolos individuais interativos para enriquecer os dados. Ainda, como participante, atuou como analista e usuário para oferecer apoio aos usuários na criação do formulário.

Por ser a coleta de termos o resultado do processo de criação do padrão textual, o pesquisador atuou com o mínimo de interação na aplicação do protocolo verbal.

Para conseguir as informações necessárias na aplicação do protocolo verbal, será necessária a execução de algumas tarefas que envolvem a elaboração de passos; através de estímulos de palavras, obter-se-ão as respostas desejadas.

Nesta pesquisa, a gravação em vídeo não foi necessária, pois observa-se que sinais não-verbais não faziam parte do objetivo, que é a coleta das terminologias usadas para recuperação de informações.

Na primeira fase, como procedimento anterior à coleta de termos, fez-se a seleção dos usuários informantes, como segue:

- a primeira tarefa para a pesquisa foi a seleção dos usuários informantes; o foco da pesquisa concentrou-se em registrar as diferentes terminologias para a mesma informação; para a seleção dos usuários considerou-se que todos participassem de tarefas com informações comuns;

- foram selecionados dois usuários, com a mesma função, porém com formação diferentes, únicos usuários envolvidos diretamente com a recuperação das informações do armazenados pelo sistema de vendas, que trabalham com informações do setor de vendas da empresa.

O mesmo sistema de informações é utilizado por vários usuários de diferentes regiões do país; optou-se por selecionar usuários de uma mesma filial, em virtude da localização geográfica das demais filiais e pelo bom relacionamento que se tem com esses usuários; os usuários selecionados trabalham com as mesmas informações com objetivos diferentes. Se esse trabalho fosse elaborado com usuários de diferentes filiais poder-se-ia obter maior diversidade de termos.

Os usuários escolhidos demonstraram pleno interesse em fazer parte da tarefa da pesquisa; um deles preferiu desenvolver essa tarefa em um horário fora do expediente normal de trabalho, pois dessa forma estaria mais à vontade para pensar, sem os desgastes e as interrupções das atividades diárias de trabalho, enquanto que o outro não impôs nenhum problema para a realização da tarefa.

Cada participante tem níveis diferentes, relacionados a diferentes fatores; o usuário informante 1 tem 2º grau completo e 2 tem grau superior completo, cursando uma especialização; ambos possuem conhecimento das informações necessárias ao ato do cadastramento de um pedido de venda.

Como essa pesquisa pretende explorar a linguagem usual de um setor específico da empresa, a seleção dos usuários teve que levar em consideração o contexto em que cada usuário opera.

Na segunda fase, durante o processo de coleta de termos procedeu-se como segue:

- antes das sessões de gravações dos protocolos verbais, foi realizada uma sessão preliminar prática usando o protocolo verbal; neste ensaio, foi passado aos usuários informantes um contexto simples, de forma que o conteúdo não interferisse no contexto final, objetivo desta

coleta; demonstrou-se como aplicar a técnica produzindo um pequeno esboço de um formulário e exteriorizando o raciocínio para compô-lo;

- logo após a sessão de como utilizar a técnica do “pensar alto”, revelou-se aos usuários informantes qual era o contexto do padrão textual que deveriam criar e solicitou-se que, durante a elaboração do padrão, descrevessem as colunas em uma folha de papel para facilitar o processo de organização mental das informações e a exteriorização do pensamento; foi permitido aos usuários informantes revisar e verificar as informações inseridas para que não se esquecessem ou voltassem a pensar na informação que tinha sido mencionada; solicitou-se que utilizassem o processo “pensando alto” o tempo todo;

- apresentaram-se aos usuários informantes as instruções da técnica de coleta de termos do protocolo verbal, conforme descrito no apêndice “A”, para melhor compreensão do processo entre os participantes.

Na terceira etapa, apresentaram-se os procedimentos posteriores à coleta de termos dos usuários informantes, como segue:

- após as sessões de gravação do protocolo verbal dos usuários, foi feita a coleta do padrão textual gerado pelos usuários; o objetivo era utilizar esse documento como ponto de apoio para o protocolo verbal;

- transcrição do *lay-out* gerado pelos usuários, como ponto de apoio para facilitar a análise do protocolo de cada usuário, na pesquisa de possíveis diferenças terminológicas, entre o *lay-out*, o protocolo e os termos utilizados no sistema de informações da empresa.

O uso do protocolo verbal ressaltou os elementos que comprovam as diferenças existentes entre usuários e sistemas de informações. O objetivo de sua aplicação é apresentar os resultados finais dessa técnica no que diz respeito aos termos diferentes utilizados pelos usuários dentro do contexto lingüístico do tema que foi explorado.

Para a transcrição do protocolo verbal dos usuários informantes, adotou-se como forma de evidenciar as informações relevantes na identificação de termos da linguagem dos usuários informantes uma convenção para representação escrita dos termos considerados importantes no ambiente de pesquisa, como segue:

Após a coleta das informações dos usuários informantes, pôde-se identificar com clareza as diferenças existentes de denominação da informação entre os usuários e os sistemas da empresa, comprovando na prática as diferenças como um dos fatores que dificulta a recuperação da informação quando solicitado pelo usuário do banco de dados, como podemos observar nos quadros que seguem:

Quadro A - Análise comparativa dos termos diferentes utilizados entre o padrão textual elaborado pelos usuários informantes:

Termos do usuário informante 1	Termos do usuário informante 2
Dados do Cliente	Código e Nome do Cliente
Data de inclusão do Pedido	Data do Pedido
Itens que compõem o pedido	Produto, Preço e Quantidade
Bloqueio por motivo de crédito	Conceito, Limite de Crédito
Conceito de Crédito do Cliente	Conceito
Data Limite para Faturamento	Limite de Entrega

Quadro B - Análise comparativa dos termos diferentes constantes nos sistemas de informação e nos documentos gerados pelos participantes:

Termos no sistema de Informação da Empresa	Termos do usuário informante 1	Termos do usuário informante 2
Código Representante, Nome do Representante	Representante	Representante
Código, Nome do Cliente	Dados do Cliente	Código e Nome do Cliente
Data da Digitação do Pedido	Data de Inclusão do Pedido	Data do Pedido
Condição de Pagamento	Condição de Pagamento	Prazo de Pagamento
Valor Unitário	Preço	Preço
Motivo do Bloqueio para formação de carga	Bloqueio por motivo de crédito	Conceito ; Limite excedido

Aprofundando ainda mais a análise dos protocolos verbais dos entrevistados conforme descrito nos apêndice “B” e “C”, pode-se observar que existem informações pensadas de uma forma e descritas no padrão textual nos apêndice “D” e “E”, de outra forma; isso é fruto do nosso pensamento que relaciona automaticamente um termo a vários outros semelhantes. Ter-se-á opção de escolha para definir qual é o melhor termo para aquele momento, mas isso não ocorre desta forma todas às vezes; acredita-se que o cérebro é capaz de recuperar mais ou menos quantidade de termos relacionados, em função do estado psicológico do usuário, o que significa que há variações nesse processo, por isso, às vezes, não se recorda de uma palavra que se sabe que existe, mas simplesmente não se lembra.

Fatores como os expostos anteriormente dificultam a criação de ferramentas para a recuperação de informações pelos usuários comuns, por não existir o tratamento lingüístico nas ferramentas de recuperação, o que dificulta a utilização dessas ferramentas por usuários sem conhecimentos técnicos. Um usuário, ao tentar encontrar um termo a partir do próprio vocabulário, não o encontra, por este estar identificado pelo analista de sistema de uma outra forma, dificultando a recuperação das informações desejadas.

Neste capítulo elaborou-se o estudo da terminologia dos usuários que utilizam os sistemas de informações da empresa e destacou-se, através do levantamento bibliográfico na área de Representação Documentária, a aplicação de técnicas de coleta de informações da linguagem natural dos usuários. Durante a pesquisa, a maior preocupação foi encontrar técnicas para se obterem informações que identificassem os termos utilizados na indicação da informação a ser recuperada; adotou-se nessa obra a aplicação da técnica de coleta de termos chamada "Protocolo Verbal", neste capítulo relata-se o procedimento e a aplicação do processo de coleta, foram entrevistados dois usuários do setor empresarial. Com o resultado da aplicação da técnica de coleta se, evidenciaram as diferenças existentes nos termos coletados dos usuários, sendo elas o enfoque para elaboração desse projeto de pesquisa, foram ainda expostos, na forma de quadros, os diferentes termos que denominam a mesma informação.



CAPÍTULO 6

MODELAGEM DA FERRAMENTA DE MEDIAÇÃO

Com base nas experiências de recuperação de informações dos sistemas informatizados, adquiridas pelo pesquisador no ambiente empresarial, neste capítulo relata-se primeiramente o funcionamento dos processos de recuperação de informações executadas nas empresas atualmente e quais são os problemas existentes; no segundo momento desenvolvem-se as características do modelo da ferramenta que servirá como interface auxiliar para recuperação de informações e que é objeto dessa pesquisa e, por último, destacam-se algumas melhorias que o uso da ferramenta poderá fornecer, beneficiando as empresas e seus usuários.

No mundo empresarial brasileiro, os valores relativos de compra e de venda dos produtos resultam de diversos fatores que influenciam nos valores mercadológicos dos bens de consumo, exigindo de cada corporação buscas individuais de fatos relevantes, com o objetivo de explorar esses fatos relevantes, diferenciando-se dos demais concorrentes, tornando as corporações mais competitivas, já que a homogeneidade dos custos e dos preços não consegue, por si só, garantir a eficiência e os resultados estabelecidos.

Se por um lado a padronização das principais características deste mercado gera uma equalização virtual das empresas concorrentes dentro de patamares similares, cria-se, por outro lado, uma grande oportunidade para que essas busquem diferenças através dos recursos “inteligentes”, elaborando estratégias com o uso da massa de dados e informações que contém o seu “background”.

Neste cenário, têm importância estratégica os Sistemas de Apoio à Decisão (SAD) que passaram do estágio de geradores de agregados de controle e de eficiência para o nível de geradores de diferenciais competitivos, podendo significar muitas vezes o sucesso das corporações.

As empresas estão abastecidas com sistemas de informações modulares de entrada e saída de dados, desenvolvidos por equipes internas e ou adquiridos de empresas especializadas na fabricação de *software*, para atender às necessidades básicas de armazenamento e recuperação das informações pelos usuários de forma padronizada e automatizada.

A necessidade de recuperar informações como apoio a tomada de decisões, para elaboração e execução de estratégias de melhorias internas ou mercadológicas, tornou necessário o desenvolvimento, a aquisição e o uso de ferramentas de recuperação de informações, possibilitando o acesso por intermédio de consultas às informações específicas para atender às necessidades dos usuários das empresas, auxiliando-os em suas tarefas de análise das informações para tomadas de decisões.

Algumas das ferramentas, conhecidas como *data mart*, foram desenvolvidas para atender às necessidades das áreas que possibilitam retorno em curto prazo. Neste contexto, as áreas mais enfocadas pelas empresas na aquisição de *data mart* geralmente são as áreas de vendas e contas a receber, considerando que as relações de geração de recursos para subsistência das empresas são mais significativas.

Embora esses *data mart* sejam alimentados por informações originadas da mesma base de dados, o desenvolvimento e implantação de um outro *data mart* consistem em criar uma nova estrutura de armazenamento das informações para agilizar as rotinas de consulta e recuperação da informação, pois cada área de negócios tem seus critérios de análise e isso inviabiliza a padronização dos módulos e bancos de dados para armazenamento das informações, contribuindo para a criação de mais tabelas, mais campos e mais informações compiladas.

Dessa forma, cada *data mart* considera-se parte integrante da base dados do *Data Warehouse*. Embora conceitualmente e estruturalmente tenham a mesma finalidade, os *data mart* atendem às necessidades por áreas de negócios e assuntos, enquanto que o *Data Warehouse* liga todos os *data mart* com uma única base de dados com várias tabelas relacionadas entre si,

permitindo maior poder para elaboração de recuperação de dados para consultas de dados complexos.

Os sistemas permitem compilar e recuperar informações de diversas maneiras, desde que estejam previamente formatados os resultados a serem exibidos. Como exemplo, podemos considerar que as áreas de vendas e contas a receber têm como tarefa básica projetar: a quantidade de produtos vendidos no período, tomando como base os dados históricos e as quantidades de produtos efetivamente vendidas no mesmo período; análise das metas e da margem de contribuição dos produtos na semana; ferramentas de gestão e negociação da carteira de títulos de cobrança na situação de inadimplência.

Muitas mudanças mercadológicas impactam diretamente na estrutura de apresentação dos resultados das informações fornecidas pelos *data mart*, ocasionando a necessidade de ajustes rápidos nas ferramentas para atender às novas características, exigindo dos analistas de sistemas a reanálise de processos, para que mudanças semelhantes no futuro possam ser previamente parametrizadas.

Mas todas as soluções tecnológicas disponíveis não são suficientes, pois ainda continua existindo a necessidade por parte dos usuários de explorar informações analíticas nas bases de dados e de novas compilações dos dados, para atender às suas necessidades, como pesquisas não padronizadas *ad hoc*, para recuperar informações das bases de dados dos *Data Warehouse*.

Para suprir a necessidade de realizar pesquisas não padronizadas, empresas fornecedoras de soluções construíram ferramentas que permitem obter informações das bases de dados dotadas de recursos flexíveis para montagem da apresentação das informações a ser recuperadas, tal como o recurso de seleção das colunas do resultado da pesquisa, recurso de ordenação, recurso de filtragem (ou seja, quais informações se desejam que façam parte do resultado da pesquisa), recurso de agrupamento, recurso de totalizadores automáticos das colunas de valores, recurso de montagem de resumos automáticos dos dados recuperados.

Poderia dizer-se que ferramentas como citadas no parágrafo anterior devem ser utilizadas por usuários das empresas que necessitam recuperar informações não padronizadas para análises peculiares. Assim, reconhece-se que a interface das ferramentas tem a finalidade de facilitar ao máximo sua utilização possibilitando que usuários finais usem as ferramentas e seus recursos de forma eficiente.

As ferramentas de pesquisas não padronizadas acessam as bases de dados dos sistemas informatizados ou as bases de *Data Warehouse* da forma nativa, ou seja, os recursos de seleção das tabelas de dados e respectivos campos de dados são apresentados na nomenclatura em que as tabelas e os campos foram criados na estrutura que foi definida pelos profissionais de tecnologia.

Dessa forma, exigem-se dos usuários finais conhecimentos das tabelas das bases de dados para selecionar a tabela que contém a informação que desejam recuperar e o significado de cada campo da estrutura das tabelas para que possam utilizar todos os recursos que as ferramentas oferecem. É justamente nessas questões que começam os problemas de recuperação, como descrito no capítulo cinco.

Com as necessidades de disponibilidade de informações, as empresas estão analisando, há algum tempo, os diversos conceitos, metodologias e ferramentas envolvidas com os Sistemas de Apoio à Decisão (SAD) e implementando aplicações nesta área para os níveis de gestão da empresa.

Observa-se que as ferramentas técnicas possuem vários recursos que permitem compilar e recuperar as informações dos bancos de dados, mas sua utilização está restrita a uma pequena parte dos usuários por exigir conhecimento técnico, principalmente a familiarização do usuário com a linguagem apresentada na interface da ferramenta.

Atualmente, por não conhecerem a nomenclatura das tabelas e campos, quando surgem necessidades de recuperar informações, os usuários recorrem aos profissionais de tecnologia para fazer a recuperação da informação no formato desenhado por eles. Isso ocorre porque

difícilmente existem usuários com níveis de conhecimento técnico e familiarizado com as estruturas das tabelas de dados das bases de *Data Warehouse*.

A aplicação da técnica de coleta de dados, adotada no capítulo cinco desta pesquisa, como prova da existência de diferença lingüística entre os indivíduos integrantes de um mesmo ambiente empresarial, permite observar uma das causas da dificuldade encontrada pelos usuários no uso de ferramentas de pesquisa para recuperação de informações.

No capítulo cinco desta pesquisa, está relatada a aplicação da técnica de coleta de termos dos usuários, porém com o objetivo de colher termo da linguagem dos usuários no que diz respeito ao conhecimento adquirido no ambiente de trabalho, comprovando a pluralidade terminológica e a dificuldade de entendimento da linguagem adotada pelos sistemas empresariais.

Baseando-se no tratamento das informações armazenadas nos bancos de dados das empresas, pôde-se observar a variação de termos usados para a recuperação das informações que não estava prevista nas ferramentas de busca.

A necessidade de recuperação de informações existentes em um ambiente empresarial e o contato com os conceitos de representação do conhecimento através de instrumento da área de Análise Documentária trabalhou-se na revisão de literatura da Linguagem Documentária e no processo de Representação Documentária no capítulo dois desta pesquisa, e no capítulo três, os métodos de tratamento terminológico, para subsidiar soluções para os problemas da pluralidade terminológica existente nos sistemas de informações empresariais.

A necessidade de maior interação entre os setores e os diversos departamentos faz da tecnologia uma das mais importantes ferramentas para a sobrevivência das empresas em um mundo cada vez mais competitivo, através do desenvolvimento e utilização de sistemas de informações.

Portanto, para atingir maior interação entre os setores no uso das informações para o conhecimento coletivo dos usuários no ambiente empresarial, vários problemas devem ser

solucionados. Este projeto enfoca um dos principais obstáculos ao processo de recuperação das informações, que é a dificuldade de representação de um vocabulário específico para identificar as informações empresariais, pois a nomeação e definição dos termos técnicos ou abreviaturas não convencionais por profissionais da área tecnológica dificultam a interpretação da estrutura dos bancos de dados.

No estudo de algumas das alternativas para solucionar os problemas de recuperação, selecionou-se para estudo a que demonstraria melhores resultados e beneficiaria os usuários dos sistemas de informação da empresa; a solução escolhida para estudo foi a implementação de ferramentas de consulta acessíveis aos usuários, permitindo-lhes identificar o nome técnico ou representativo e a localização da informação desejada usando vocabulário próprio.

Sendo assim, sugere-se a necessidade de elaboração de uma ferramenta de mediação, servindo de apoio lingüístico no uso das ferramentas de garimpagem de informações, que possibilite aos usuários finais pesquisarem as informações armazenadas nos bancos de dados.

Essa necessidade identificou de forma preliminar a realização de uma pesquisa exploratória mediante a revisão de literatura, para investigar a possibilidade de elaboração da interface do sistema de recuperação.

O trabalho de pesquisa inicia-se pelo processo de levantamento bibliográfico sobre a construção de linguagens documentárias e controle de vocabulários, para obtenção de fundamentos e introdução de conceitos que deverão ser incorporados no processo de recuperação em Sistemas de Informação e Sistemas de *Data Warehouse* no ambiente empresarial. Na seqüência, faz-se o estudo teórico e metodológico de construção de linguagem para conhecimento das estruturas e procedimentos de recuperação.

A necessidade de ressaltar o problema tratado nesta pesquisa usou métodos de coleta de termos dos usuários obtendo-se uma amostra que permitiu identificar a diversidade existente na linguagem adotada pelos usuários ao acessar os sistemas de informações empresariais.

No capítulo três, elaboraram-se estudos das características da terminologia, e se relataram, com base nas bibliografias analisadas sobre a Terminologia, as relações entre conceito e termo, e entre os termos e as unidades sistemáticas; esses itens compõem a estrutura conceitual da ciência terminológica como referência para ser utilizada com uma característica interdisciplinar.

Como resultado, desenvolveu-se o modelo da ferramenta de mediação; com a incorporação dos conceitos da lingüística da área de Representação da Informação tratada no capítulo dois e os conceitos de recuperação do significado da área da Terminologia tratados no capítulo três, surgiram características importantes para modelagem da ferramenta de mediação; tal modelo inspirou-se nos problemas de recuperação apresentados no capítulo cinco desta pesquisa, com base em ambientes empresariais informatizados expostos no capítulo quatro.

A ferramenta de mediação deverá facilitar o acesso às informações, diminuindo o tempo despendido pelos usuários na busca de conhecimento dos processos existentes nas empresas em que atuam, principalmente na utilização dos sistemas de informações, em que os usuários têm dúvidas constantes sobre onde obter a informação de que precisam.

Na revisão de literatura, relatou-se sobre os conceitos de estrutura e funcionalidade de Linguagens Documentárias como os tesauros e os sistemas de classificação documentária.

Ressalta-se que o estudo científico revela caminhos importantes de caráter interdisciplinar como a introdução da terminologia que complementa a área de Linguagens Documentárias, através da análise da evolução histórica de tratamento da informação.

Ao adotar-se a terminologia como método que ampara o desenvolvimento da ferramenta de mediação para a recuperação de informações dos sistemas empresariais, pode-se criar um repositório de termos da linguagem natural que, por se tratar de uma comunidade de indivíduos com objetivo comum, pode ser considerado uma linguagem especializada empregada no ambiente empresarial, pois tende a solucionar os problemas de comunicação e recuperação de informações na corporação.

O estudo da lingüística desenvolveu nos últimos anos técnicas para analisar o processo do conhecimento humano durante a leitura, produzindo de forma direta e indireta subsídios significantes para o tratamento da informação, possibilitando novos estudos na forma de armazenar informações e criar novas de técnicas e ferramentas de recuperação de informações mais eficazes.

Portanto, fazendo-se uso desses subsídios teóricos e metodológicos de caráter interdisciplinar, pôde-se adotar conceitos para o mundo digital e modelar ferramentas de recuperação das informações que interajam com a linguagem dos usuários.

Na evidência de diferenças lingüísticas existentes no ambiente adotado para a pesquisa e como resultado da elaboração do padrão textual elaborado pelos usuários, apresentado no capítulo cinco, houve motivação para continuar modelando a ferramenta de mediação.

Desta forma, a solução deverá oferecer rapidez e facilidade na tarefa de recuperação das informações armazenadas nos bancos de dados dos sistemas das empresas, unindo-se os conceitos de tratamento da informação resgatada na investigação dos métodos de organização do conhecimento na Representação Documentária, Linguagem Documentária, e a Terminologia.

Na implementação do instrumento de mediação, deve-se realizar a importação dos termos compilados para uma base terminológica os quais, depois de importados, deverão ser analisados pelo profissional de tratamento da informação, para receberem as representações e relações necessárias.

O profissional de tratamento da informação participará ativamente do processo de organização da base de dados terminológica, pois será responsável em determinar quais termos dos sistemas de informações empresariais devem ser registrados na base de dados terminológica, realizando a mesma tarefa com os termos dos usuários dos sistemas de informações empresariais.

Embora moroso, o trabalho de identificar e representar todos os termos definidos pelos profissionais da área tecnológica é de vital importância para registro na base terminológica, pois



o instrumento de mediação já estará sendo implementado com uma quantidade de termos bastante representativos, o que estimulará a sua utilização.

É importante ressaltar que cada termo codificado estará relacionado ao metadado da estrutura dos sistemas da empresa. Portanto, os termos deverão estar relacionados a termos mestres, classificados pelo profissional de tratamento da informação e o termo mestre conterá o endereço físico da informação na base de dados da empresa.

Considerando que a mesma informação poderá estar replicada em várias tabelas da base de dados, para isso, deverão ser relacionados todos os endereços para que os usuários possam saber selecionar de qual tabela recuperar a informação, pois embora replicadas com o mesmo significado, as tabelas normalmente espelham contextos e períodos diferentes de informações históricas armazenadas.

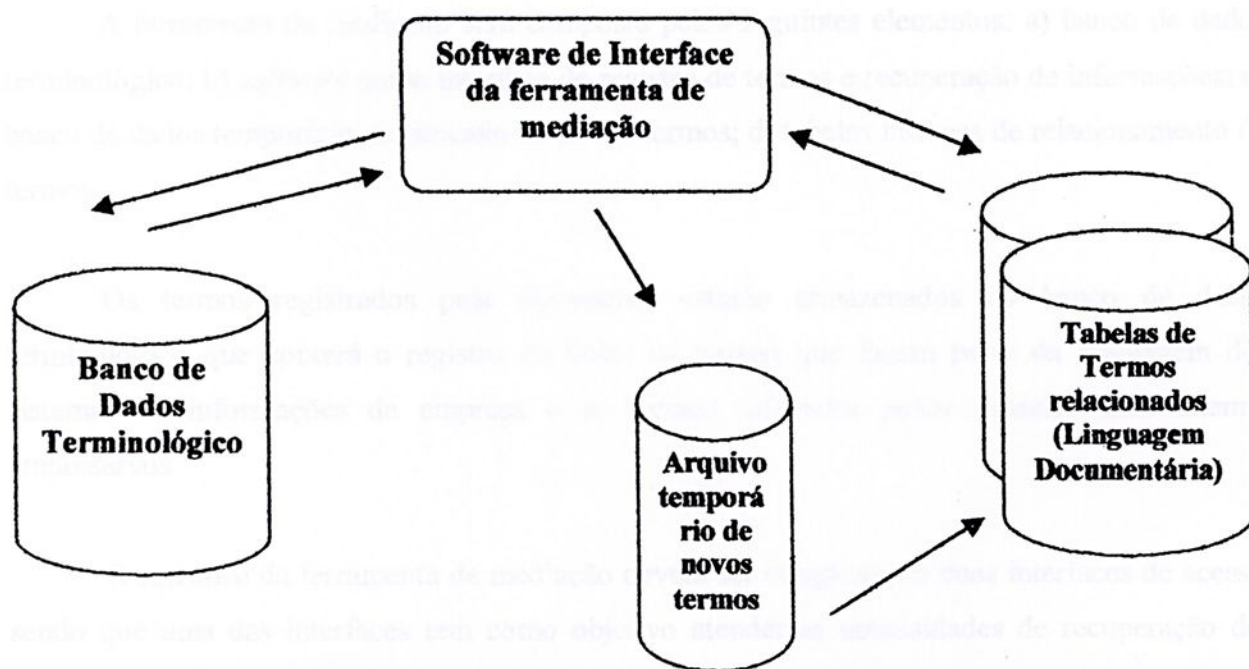


Figura 3 : Fluxograma lógico da ferramenta de mediação

Fonte : Silva, J.M.P. da, Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de *Data Warehouse*, Dissertação de Mestrado, p.101, 2002.

A ferramenta deverá contemplar as palavras com variações gráficas, o que normalmente ocorre por questões geográficas, costumes sociais e evolução lingüística. Dessa forma será ampliado o tempo de utilização da ferramenta, retratando as diferentes formas de escrita de um termo. Considerando a idéia de que a pesquisa solicitada e a recuperação das informações funcionarão por intermédio de um mecanismo que explorará os termos relacionados, a ferramenta se tornará mais eficiente.

O modelo da solução proposto deve oferecer eficiência e dinamismo, como ferramenta que possui uma linguagem de mediação entre usuários e bases de dados empresariais, para facilitar o processo de recuperação das informações armazenadas em arquivos magnéticos. A ferramenta é composta por duas características importantes que são: fornecer suporte lingüístico exibindo listas de termos relacionados considerando a linguagem dos usuários; servi-se de listas de endereços eletrônicos com a localização da informação armazenada fisicamente na base de dados empresariais.

A ferramenta de mediação será composta pelos seguintes elementos: a) banco de dados terminológico; b) *software* como interface de registro de termos e recuperação de informações; c) banco de dados temporário de inclusão de novos termos; d) tabelas internas de relacionamento de termos.

Os termos registrados pela ferramenta estarão armazenados no banco de dados terminológico que conterà o registro de todos os termos que fazem parte da linguagem dos sistemas de informações da empresa e os termos utilizados pelos usuários dos sistemas empresariais.

O *software* da ferramenta de mediação deverá ser composta de duas interfaces de acesso, sendo que uma das interfaces tem como objetivo atender às necessidades de recuperação dos usuários dos sistemas de informações empresariais, a princípio conterà as seguintes opções: a) entrada de termos para pesquisa; b) lista de termos relacionados e c) solicitação de registro de termos. A outra interface do *software* da ferramenta com as características de configuração do processo de mediação de uso exclusivo do profissional de tratamento da informação inicialmente

conterá as seguintes opções: a) manutenção de termos (com inclusão, alteração e exclusão); b) manutenção da tabela de relacionamento de termos (com inclusão, alteração e exclusão); c) listagem de solicitação de inclusão de novos termos e d) autorização de inclusão de novos termos solicitados pelos usuários.

Dessa forma, uma das interfaces de mediação deve ser elaborada para acesso dos usuários das empresas com a finalidade específica de atender às necessidades de consultas terminológicas às informações contidas no banco de dados terminológico da ferramenta, para conhecer estruturalmente o modelo de armazenamento das informações nas bases de dados.

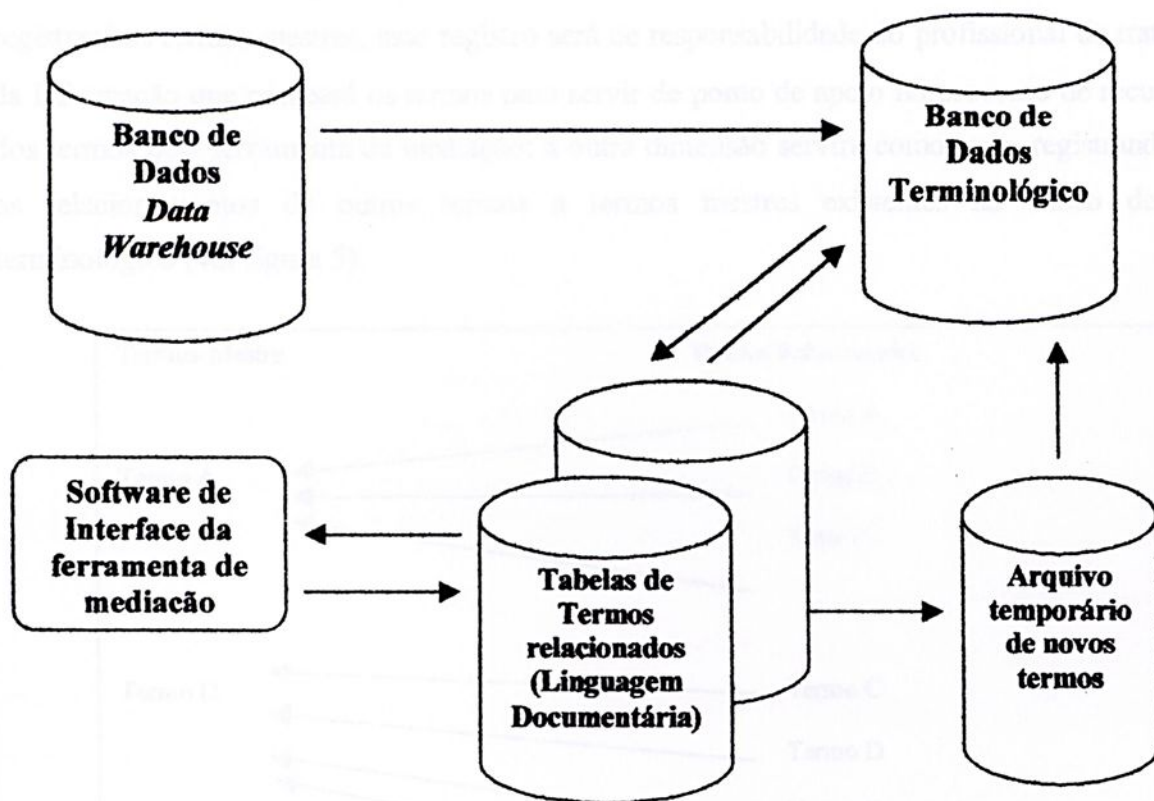


Figura 4: Fluxograma físico da ferramenta de mediação

Fonte: Silva, J.M.P. da, Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de Data Warehouse, Dissertação de Mestrado, p.103, 2002.

A interface dos usuários deverá atender às solicitações de consultas no uso da linguagem natural, e os termos não registrados no banco de dados terminológico da ferramenta devem ser enviados para o repositório de termos para serem analisados e tratados por profissionais de tratamento da informação na outra interface do *software* da ferramenta.

Para que a ferramenta de mediação possa exercer o seu papel de tratamento lingüístico, além do banco de dados terminológico, conterà as tabelas de relacionamentos entre termos, essenciais nos processos de recuperação da ferramenta, possibilitando ao profissional de tratamento da informação fazer a parametrização dos elos de parentesco entre os termos.

A tabela de relações entre termos deverá conter duas dimensões: uma das dimensões registrará os termos mestres, esse registro será de responsabilidade do profissional de tratamento da informação que nomeará os termos para servir de ponto de apoio no processo de recuperação dos termos pela ferramenta de mediação; a outra dimensão servirá como guia, registrando todos os relacionamentos de outros termos a termos mestres existentes na banco de dados terminológico (ver figura 5).

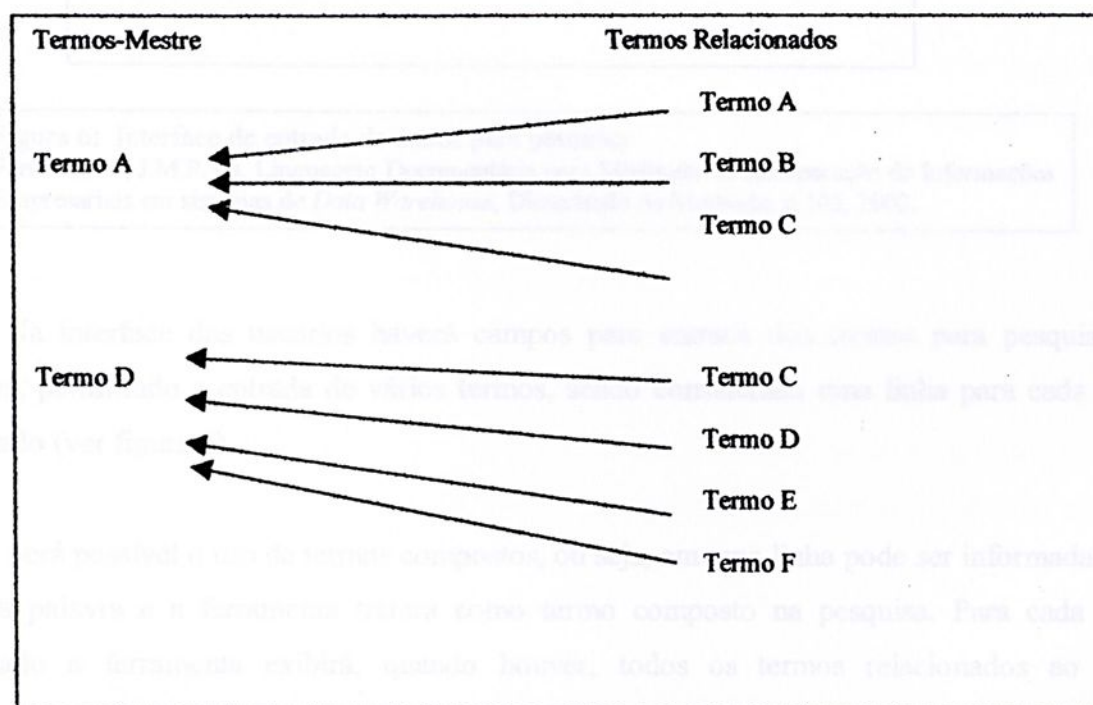


Figura 5: Gráfico de relacionamento de termos

Fonte: Silva, J.M.P. da, Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de *Data Warehouse*, Dissertação de Mestrado, p.104, 2002.

Quando o usuário fornecer o termo para pesquisa, a ferramenta de mediação percorrerá a dimensão definida como guia do processo de recuperação e localizará o termo mestre do termo fornecido, e a partir do termo-mestre a ferramenta deverá exibir todos os termos relacionados ao termo-mestre, na interface de termos relacionados modelados na figura quatro deste capítulo.

Estruturalmente o banco de dados terminológico adotará alguns princípios de construção de tesauros. De acordo com Curras (1995, p.102-103), é importante ressaltar que palavras-chave no vocabulário especializado expressam conceitos concretos.

Informe os termos para pesquisa	
1	
2	
3	
4	

Figura 6: Interface de entrada de dados para pesquisa

Fonte: Silva, J.M.P. da, *Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de Data Warehouse*, Dissertação de Mestrado, p.105, 2002.

Na interface dos usuários haverá campos para entrada dos termos para pesquisa dos usuários, permitindo a entrada de vários termos, sendo considerada uma linha para cada termo informado (ver figura 6).

Será possível o uso de termos compostos, ou seja, em uma linha pode ser informadas mais de uma palavra e a ferramenta tratará como termo composto na pesquisa. Para cada termo informado a ferramenta exibirá, quando houver, todos os termos relacionados ao termo informado. Portanto, o usuário deverá selecionar dos termos relacionados quais devem fazer parte da pesquisa (ver figura 7).

Termos Relacionados			
Termo informado : _____			
Termo(s) Relacionados:			
<u>Descrição</u>	<u>Significado</u>	<u>Contexto aplicado</u>	<u>Selecione</u>

Figura 7: Interface com a listagem dos termos relacionados ao termo pesquisado
 Fonte: Silva, J.M.P. da, *Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de Data Warehouse*, Dissertação de Mestrado, p.106, 2002.

Caso existam termos informados que não façam parte do banco de dados terminológico da ferramenta, a interface comunicará ao usuário que o termo não está cadastrado e procederá ao registro do termo em um arquivo temporário; os termos registrados no arquivo temporário será analisado pelo profissional de tratamento da informação para confirmar se o termo deverá ser cadastrado no banco de dados terminológico; ao se registrar o termo no arquivo temporário a ferramenta fornecerá ao usuário uma caixa de diálogo ao usuário para fazer a descrição resumida do significado do termo e em que contexto está sendo empregado. Portanto, as informações fornecidas nesse processo são enviadas para o repositório de termos aguardando registro, os quais serão posteriormente analisados por profissionais de tratamento da informação (ver figura 6).

Após concluída a etapa de entrada e seleção dos termos para consulta, a ferramenta deverá, com base nos termos informados, fazer a busca no dicionário de endereços das bases de dados e das tabelas, para exibir como resultado das pesquisas os endereços eletrônicos do local em que a informação está disponível (ver figura 9).

Solicitação de Registro de Termo		
Termo : _____		
Significado do termo : _____ _____		
Contexto empregado : _____		
Nome usuário solicitante: _____	Data: _____	Hora : _____

Figura 8: Interface de solicitação de registro de termos

Fonte: Silva, J.M.P. da, *Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de Data Warehouse*, Dissertação de Mestrado, p.107, 2002.

A outra interface da ferramenta de mediação deve ser elaborada para os acessos exclusivos dos profissionais de tratamento das informações, pois nesta interface será permitido realizar alterações de ajustes dos dados a todas as tabelas de configuração de relacionamento de termo.

	Resultado da Consulta					
Nome / descrição do Termo	Base de dados Nome		Nome da Tabela		Nome do Campo na Tabela	
	Físico	Lógico	Físico	Lógico	Físico	Lógico

Figura 9: Interface do resultado da consulta de termos

Fonte: Silva, J.M.P. da, *Linguagem Documentária para Mediação de Recuperação de Informações Empresariais em sistemas de Data Warehouse*, Dissertação de Mestrado, p.107, 2002.

Nessa interface de uso técnico de configuração das relações dos termos, com base nos conceitos lingüísticos de representação da informação, estará disponível o acesso ao banco de dados terminológico, para fazer inclusões de termos novos ou alterações das informações agregadas aos termos, como seu significado e a descrição de qual contexto o termo é empregado no ambiente empresarial que está sendo pesquisado.

Na opção de acesso à Linguagem Documentária (ferramenta de mediação) a tabela de relacionamento de termos permitirá inserir ou excluir termos que possam compor a relação de abrangência de significados dos termos; nesta opção define-se o grau de relação entre os termos para facilitar sua identificação, como grau de hierarquia, associação e de equivalência, tratado no capítulo dois.

O profissional de tratamento da informação deve adotar os princípios da especificidade, para realizar o relacionamento de termos. Determinar o grau de hierarquia entre os termos significa ter termos relevantes e menos relevantes dependendo do nível de importância no ambiente em que o termo está sendo empregado.

Na opção de acesso à tabela de endereço físico de armazenamento das informações, permite-se inserir ou excluir termos e as definições do endereço lógico e físico dos termos pesquisados; esta tabela sempre será alterada quando houver a criação de uma nova tabela ou base de dados nos sistemas de informações empresarias ou quando houver a exclusão ou a desativação de uma tabela ou base de dados nos sistemas de informações das empresas.

E, por fim, a opção de acesso ao arquivo temporário que conterà as solicitações de registro de termos que não fazem parte do banco de dados terminológico feitas pelos usuários, através da descrição do significado do termo, o profissional de tratamento da informação deverá efetuar as análises, e proceder de acordo com os conceitos lingüísticos, determinar a inclusão ou não do novo termo no banco de dados terminológico, ao incluir deverá providenciar o cadastramento do termo também na tabela de termos relacionados.

Segundo Dubuc (1999, p.211), "A estrutura deve permitir a seleção dos dados com alguns parâmetros e uma ordem definida. E o sistema tem que oferecer ao usuário possibilidade de decodificar todas as abreviaturas e símbolos utilizados".

Pode-se adicionar na elaboração da ferramenta a possibilidade dos profissionais de tratamento da informação receberem mensagens eletrônicas instantâneas à medida em que os usuários realizem solicitações de novos termos, tornando a análise e inclusão de novos mais rápidas e eficientes e melhorando o desempenho da ferramenta.

O uso das aplicações de recuperação de informações pode ser realizado diretamente pelos usuários das empresas; a necessidade de obter a localização e como está estruturada a base de dados da empresas podem ser conseguidas pelos próprios usuários, acessando a ferramenta de mediação que conterà um banco de dados construído para o modelo proposto nesta pesquisa.

As solicitações e suporte para recuperação de informações realizadas ao pessoal técnico serão significativamente reduzidos à medida que a ferramenta e o dicionário terminológico passarem a atender complementemente as perguntas dos usuários.

A melhora dos meios de acesso às informações armazenadas nas bases de dados empresariais passa por revisões constantes e reavaliações para encontrar soluções que contribuam de maneira significativa no processo de obtenção de informações para análise do perfil das empresas, comportamento do mercado, tendências dos negócios, a análise das informações armazenadas nos bancos de dados das empresas viabiliza projetos de estratégias para crescimento empresarial e social.

Esse modelo poderá atender a necessidades não no ambiente empresarial, mas também em instituições com objetivos sociais, pois pode ser implantado em uma base de dados de uma rede hospitalar, ou ainda em repartições públicas. O modelo trata da linguagem natural como linguagem de especialidade por ambiente, e nesses locais a linguagem utilizada, quando trata de assuntos relacionados diretamente ao ambiente, também pode ser considerada como linguagem de especialidade.

Pode-se ampliar a utilização da ferramenta, incorporando à base de dados da ferramenta de interface os chamados termos técnicos, expressões mneumônicas, ou formas sintetizadas de representação utilizadas pelos profissionais de tecnologia da informação, para identificarem ou referenciarem as informações nas tabelas de dados, cuja referência deve corresponder ao endereçamento físico da informação. Essas referências técnicas que fazem parte da ferramenta de mediação favorecerão a sua utilização também pelos profissionais de tecnologia da informação, podendo até ser utilizadas como ferramentas de administração de novos termos técnicos implementados nas bases de dados.

Neste capítulo destacaram-se inicialmente os mecanismos de recuperação das informações armazenadas em bancos de dados empresariais, com a finalidade de deixar evidentes as dificuldades dos usuários dos sistemas de informações em recuperar informações com aplicações de alto nível técnico de acessibilidade, mas que não oferece nenhum tratamento lingüístico, dificultando o uso das aplicações, exploração das bases de dados e conseqüentemente a recuperação da informação.

Como propósito dessa pesquisa modelaram-se neste capítulo as características essenciais de uma ferramenta de mediação entre a linguagem dos usuários e os sistemas de informações empresariais, como recurso auxiliar de acesso às informações armazenadas nas bases de dados empresariais e no uso de ferramentas de garimpagem de dados.

Por fim, relatam-se os benefícios a serem alcançados ao se inserir a ferramenta modelada nesse capítulo na rotina de recuperação de informações em um ambiente empresarial, resultando no maior volume de informações recuperadas, maior volume de informações processadas e analisadas pelos usuários, desvinculando a dependência da necessidade de recuperação dos usuários do pessoal de tecnologia.

Acredita-se que essas melhorias de recuperação no ambiente empresarial deverão favorecer substancialmente o poder de reação das empresas na análise dos fatores de sucesso, as estratégias podem ser mantidas ou modificadas pelos acontecimentos internos e mercadológicos.

Com a liberdade na exploração da base de dados, os usuários ganharão maior velocidade na análise e reação aos resultados das informações recuperadas, conseguindo com mais agilidade elaborar novas estratégias de trabalho.

CONCLUSÃO

A proposição desta pesquisa foi o estudo teórico e metodológico de Representação Documentária para elaboração do modelo de mediação de recuperação de sistemas de *Data Warehouse* de sistemas de informações empresariais, servindo como linguagem de mediação entre o usuário e os bancos de dados empresariais, cuja elaboração pelo profissional da informação deverá ser permanentemente subsidiada por uma base terminológica das áreas de especialidades.

Buscou-se como resultado prático os elementos que permitam a independência dos usuários dos sistemas de informações das empresas em localizar e recuperar as informações desejadas, diminuindo assim, as solicitações de suporte dos profissionais da área de tecnologia da informação, possibilitando que cada usuário possa buscar por si e, de forma direta, as informações de que necessita e incorpora-las às suas atividades.

A recuperação das informações armazenadas nos bancos de dados dos sistemas de informações empresariais é dificultada pelas limitações técnicas ou funcionais de acesso às informações depositadas nos computadores. Torna-se cada vez mais necessário criar novos recursos tecnológicos, para fornecer aos usuários que trabalham com as informações empresariais, meios de facilitar a recuperação das informações.

Trabalhou-se no sentido de encontrar soluções para que os usuários possam obter respostas para as suas dúvidas a respeito da informação que desejam recuperar, sem ter que recorrerem ao profissional de suporte ou aos analistas de sistemas, os quais, na maioria dos casos, estão sobrecarregados com outras tarefas, ficando os usuários limitados em suas atividades e na dependência da informação.

Dessa forma, concentrou-se os estudos na necessidade de elaboração de uma ferramenta de mediação baseado na constatação teórica que confirma a interdisciplinaridade entre a Análise Documentária e a Terminologia e que possibilite ao usuário realizar pesquisas usando a própria linguagem. Os estudos realizados buscaram encontrar conceitos e métodos para viabilizar a elaboração da ferramenta de mediação no sentido da informação estar representada adequadamente.

A interação entre a Análise Documentária e a Terminologia existe na busca pela organização do conhecimento, embora tratem de problemas diferentes e paralelos, pois a união das duas áreas responde pelos conceitos de armazenamento e recuperação da informação com enfoques diferentes não conflitantes, mas complementares.

Como instrumento de investigação do processo de representação da informação para alcançar a solução delineada nesta pesquisa, os objetivos foram: investigar a interdisciplinaridade entre as áreas de Representação Documentária e Terminologia mediante referencial teórico publicado; comprovar a pluralidade lingüística no ambiente especializado através da coleta de linguagem dos usuários; elaborar modelo da ferramenta de mediação para recuperação de informações armazenadas em bancos de dados empresariais.

O estudo de literatura sobre o tratamento da informação, no que diz respeito às relações entre conceito e termo e os termos e as unidades sistemáticas, compõem uma estrutura conceitual que caracteriza a interdisciplinaridade entre a Representação Documentária e a Terminologia.

A comprovação da pluralidade lingüística existente no ambiente empresarial obteve-se com a aplicação de entrevista para coleta dos dados. Dessa forma, pode-se evidenciar os diferentes termos para a mesma informação utilizada, permitindo observar como a variedade lingüística prejudica o processo de recuperação.

Os problemas de recuperação da informação causada pela pluralidade lingüística geram a necessidade de reunir recursos metodológicos e tecnologicos para serem solucionados. Neste sentido desenvolve-se nesta pesquisa um modelo de uma ferramenta de mediação que possibilite

aos usuários realizarem pesquisas usando a linguagem natural do ambiente de trabalho; para a concretização do modelo da ferramenta de mediação foi necessário investigar a possibilidade de interdisciplinaridade entre áreas de organização do conhecimento, a Análise Documentária e a Terminologia são áreas que preocupam-se em organizar, armazenar, recuperar informações cuidando para seja manter no processo de recuperação o resgate do conhecimento.

Nos estudos preliminares para suprir as necessidades dos usuários na busca de informações nos bancos de dados dos sistemas de informações empresariais foram analisadas as seguintes alternativas: programa de treinamento dos usuários; sistemática de suporte de atendimento aos usuários; elaboração de uma ferramenta de consulta para uso dos usuários.

A alternativa escolhida nesta pesquisa para solucionar os problemas de recuperação das informações armazenadas nos bancos de dados empresariais com maior eficiência, foi a implementação de uma ferramenta de consulta acessível aos usuários, permitindo-lhes identificar o nome técnico ou representativo e a localização da informação desejada com base no próprio vocabulário.

Estudos identificaram que elementos necessários para o desenvolvimento da alternativa mencionada no parágrafo anterior revelaram a necessidade de realizar uma pesquisa exploratória, mediante a revisão de literatura para investigar a possibilidade de elaboração da interface do sistema de recuperação.

Por se tratar de uma solução de recuperação de informações armazenadas em bancos de dados empresariais fez-se o estudo de conceitos de recuperação, resgatando métodos de tratamento lingüísticos para viabilizar a interação da linguagem dos usuários com a linguagem dos sistemas de informações empresariais.

Realizaram-se estudos de fundamentação teórica de organização do conhecimento na área de Representação Documentária. Coletando, assim, as características identificadas de tratamento da informação que serviram como diretrizes para elaboração do modelo de sistema de mediação.

Abordou-se no estudo da Representação Documentária a necessidade de garantir o significado do termo quando isolado do texto, o que proporcionou a investigação sobre os fundamentos da Terminologia, encontrando-se subsídios importantes para a configuração do comportamento do sistema de mediação na recuperação da informação. Observou-se neste estudo que a Terminologia complementa a área de Representação Documentária.

Destacaram-se, nesta pesquisa, as funcionalidades dos mecanismos de recuperação nos sistemas de informações empresariais e qual a importância que as ferramentas de recuperação têm para auxiliar os usuários na elaboração de estratégias e tomadas de decisões. Evidenciou-se assim, a necessidade de maior independência dos usuários no acesso às informações armazenadas nas bases de dados.

Ao analisar as dificuldades dos usuários em recuperar informações armazenadas nos bancos de dados empresariais, evidenciou-se que a pluralidade lingüística existente entre os usuários faz parte de um vocabulário que permite diversas formas de representação, que as ferramentas de recuperação dos sistemas de informações empresariais atuais não contemplam as diferentes maneiras de representação da informação.

Sendo assim, constatou-se a necessidade de uma ferramenta de mediação que permita os usuários acessarem as informações armazenadas nos bancos de dados das empresas usando sua linguagem natural e confirmando o pensamento inicial dessa pesquisa que propôs a elaboração de proposta de modelo de uma ferramenta de mediação para recuperação de informações armazenadas nos bancos de dados das empresas.

O modelo da ferramenta de mediação teve como base na experiência adquirida pelo pesquisador em fornecer soluções para recuperação de informações armazenadas em bancos de dados empresariais. Desta forma, uniram-se os conceitos tecnológicos e os conceitos de Representação Documentária e da Terminologia como elementos fundamentais para o desenvolvimento de uma ferramenta de mediação eficiente.



Do ponto de vista do autor, o modelo da ferramenta de mediação proposto nesta pesquisa possui características que poderão facilitar sua implementação e incentivará os usuários a acessarem as informações dos bancos de dados empresariais. Além disso, o modelo poderá tornar-se um recurso importante tecnológico importante para recuperação de informações armazenadas nos sistemas de computadorizados.

O modelo da ferramenta de mediação teorizado nesta pesquisa está composto pelos seguintes elementos: banco de dados terminológico; *software* como interface de registro de termos e recuperação de informações; banco de dados temporário de inclusão de novos termos; tabelas internas de relacionamento de termos.

O *software* da ferramenta de mediação, delineado nesta pesquisa, oferecerá duas interfaces de acesso, sendo uma delas disponíveis aos usuários dos sistemas de informações para realizarem: entrada de termos; lista de termos relacionados e solicitação de registro de termos. A outra interface do *software* da ferramenta terá como características a configuração dos processos de mediação e será de uso exclusivo do profissional de tratamento da informação, inicialmente esta interface conterá as seguintes opções: manutenção de termos; manutenção da tabela de relacionamento de termos; listagem de solicitação de inclusão de novos termos e autorização de inclusão de novos termos solicitados pelos usuários.

Os termos registrados pela ferramenta estarão armazenados em um banco de dados terminológico que conterá o registro de todos os termos que fazem parte da linguagem dos sistemas de informações da empresa e os termos utilizados pelos usuários dos sistemas empresariais. Os novos termos para inclusão serão armazenados no banco de dados temporário aguardando avaliação do profissional de tratamento da informação.

As comunidades de usuários que utilizam os sistemas de informatizados, normalmente, têm a necessidade de fazer análise das informações armazenadas nos bancos de dados empresariais, embora existam muitas ferramentas de recuperação de informações para bancos de dados, não são disponíveis aos usuários pelo fato de que as ferramentas fazem o acesso a

estrutura dos bancos de dados de forma nativa, ou seja, em uma estrutura com definições técnicas e uma linguagem de difícil compreensão por usuários comuns.

Com esse enfoque a ferramenta de mediação permite-se ser utilizada em qualquer grupo de usuários com objetivos comuns, pois se considera que a linguagem veiculada entre os usuários uma linguagem de especialidade, por exemplo: usuários de uma empresa de calçados; usuários de uma empresa têxtil; usuários de uma empresa de alimentos; usuários de um hospital, esses grupos têm objetivos comuns e usam uma linguagem típica para obter as informações que necessitam no ambiente em que trabalham.

Porém não se deve aplicar a ferramentas de mediação como solução global em ambientes cuja comunidade de usuários tenham objetivos diferentes, pois quando os assuntos são tratados em diferentes enfoques, a linguagem empregada também poderá conter diferentes significados em função do contexto, o que poderia prejudicar a eficiência do processo de recuperação da ferramenta. O profissional de tratamento da informação terá grande dificuldade em fazer o relacionamento dos termos, podendo também as diferenças de significas causar um relacionamento com assuntos de áreas que não fazem parte do contexto.

A tecnologia a ser adotada para construção dos bancos de dados e do *software* que fará a leitura independe da tecnologia de banco de dados empregada no sistema de informações da empresa ou local em que se deseja implementar a ferramenta. A ferramenta pode ser construída com qualquer tecnologia, sendo de fácil implementação permite o acesso a tecnologias mais acessíveis.

Por tratar essa solução de um modelo para construção de uma ferramenta de mediação foi o primeiro passo para o tratamento das diferenças lingüística existente no processo de recuperação de informações armazenadas em uma estrutura computacional, existe a necessidade de implementar a solução proposta em um ambiente com as características recomendadas nesta pesquisa.

Pode-se estudar o uso dessa solução em portais da *WWW*, os portais na internet têm por principio estar direcionados para um determinado contexto, com objetivo de facilitar o acesso dos

usuários, oferecendo o máximo de informações sobre o tema trata no portal, obtendo também de forma geral uma comunidade de usuários com objetivos comuns, porém com diferenças lingüísticas, dessa forma pode-se fazer o uso da solução modelada nesta pesquisa.

Estudos podem ser feitos com o propósito de analisar a eficácia da ferramenta de mediação no sentido de ser incorporado essa técnica nas ferramentas de recuperação de informações existente no mercado ou em construção, servindo recurso metodológico importantes no processo de recuperação podendo estar disponível em uma interface *Data Warehouse*.



REFERÊNCIAS BIBLIOGRÁFICAS

ALONSO, D.V. Lenguaje e información. In: _____. *Lenguajes Documentarios*. [S.l.]: Nueve Parhadigma, 1997. p. 23-44.

ARAÚJO, V. M. R. H. de. Sistemas de informação: nova abordagem teórico-conceitual. *Ciência da Informação*, Brasília, v.24, n.1, p.54-76, jan./abr. 1995.

AUGER, P. et al. Automatisation des procédures de travail en terminologie dans meta. *Journal des traducteurs*, Montreal, v. 35, n. 1, p. 121-127, mar. 1991. Apud CABRÉ, M. T. *La terminologia: teoria, metodologia, aplicaciones*. Barcelona: Antártida, 1993. 529 p.

BARRETO, A. de A. A questão da informação. *São Paulo em Perspectiva: revista da Fundação SEADE*, v.8, n. 4, p.3-8 In: GINEZ DE LARA, M. L. *Representação e linguagens documentárias: bases teórico-metodológicas*. São Paulo: Universidade de São Paulo, 1999. 208 f. Tese (Doutorado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo.

BATTAGLIA, M. da G. B. A inteligência competitiva modelando o sistema de informação de clientes – FINEP. *Ciência da Informação*, Brasília, v.28, n.2, p.200-214, maio/ago. 1999.

BOULANGER, J. C. Alguns componentes lingüísticos no ensino da terminologia. *Ciência da Informação*, Brasília, v.24, n.3, p.313-318, set./dez. 1995.

CABRÉ, M. T. La terminologia hoy: concepciones, tendencias y aplicaciones. *Ciência da Informação*, Brasília, v.24, n.3, p.289-298, set./dez. 1995.

_____. *La terminologia: teoria, metodologia, aplicaciones*. Barcelona: Antártida, 1993. 529 p.

CAMPOS, M. L. A. *Linguagem Documentária: teorias que fundamentam sua elaboração*. Niterói: EdUFF, 2001. 133 p.

_____. *Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesauro-baseado-em-conceito e a teoria geral da terminologia*. Rio de Janeiro, 1994. 196 f. Dissertação (Mestrado em Ciência da Informação), Universidade Federal do Rio de Janeiro.

CANOGIA, C. Convergência da Inteligência Competitiva com Construção de Visão de Futuro: proposta metodológica de Sistema de Informação Estratégico (SIE) Disponível em: < _____. Acessado em: 08 de mar. 2002.

CINTRA, A. M. et al. *Para entender as linguagens documentárias*. São Paulo: APB, 1994. 72 p.

_____. Elementos de lingüística para estudos de indexação. *Ciência da Informação*, Brasília, v.12, n.1, p.5-22, 1983.

CURRÁS, E. *Tesauros: linguagens terminológicas*. Trad. Antônio Felipe Corrêa da Costa. Brasília: IBICT, 1995. 285 p.

DAL'ALBA, A. O que é um *data warehouse* Disponível em: <
>. Acessado em: 10 de fev. 2002.

DUBUC, R. Qu'est-ce que la terminologie ?. *Le banque des mots*, Paris, n.13, p. 3-14. 1977. Apud GINEZ DE LARA, M. L. *Representação e linguagens documentárias: bases teórico-metodológicas*. São Paulo, 1999a. 208 f. Tese (Doutorado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo.

_____. *Manuel pratique de terminologie*. Trad. Ileana Cabrera. Chile: Ril, [1999b]. 236 p.

FELBER, H. *Termologia manual*. Paris: UNESCO, 1984. 234 P. Apud CAMPOS, M. L. A. *Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia*. Rio de Janeiro, 1994. 196 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação da Universidade Federal do Rio de Janeiro, Universidade Federal do Rio de Janeiro.

_____. The Vienna School of terminology: fundamentals and its theory. In: INTERNATIONAL SYMPOSIUM ON THEORETICAL AND METHODOLOGICAL PROBLEMS OF TERMINOLOGY, *Procendins...* Moscow, nov. 2-7, 1979. p. 69-86. Apud CAMPOS, M. L. A. *Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia*. Rio de Janeiro, 1994. 196 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação da Universidade Federal do Rio de Janeiro, Universidade Federal do Rio de Janeiro.

FUJITA, M. S. L. *Linguagem documentária em odontologia: uma aplicação do sistema de indexação PRECIS*. São Paulo, 1992. Tese (Doutorado em Ciências da Comunicação)- Escola de Comunicação e Artes, Universidade de São Paulo.

GARDIN, J.C. et. al. *L'automatisation des recherches documentaires: un modèle general "Le SYNTOL"*. 2.ed. revue et augmentée. Paris: Gauthier-Villars. 1968. Apud CINTRA, A. M. et al. *Para entender as linguagens documentárias*. São Paulo: APB, 1994. 72 p.

_____. Document analysis and linguistic theory. *Journal of Documentation*, London, v. 29, n.2, 137-168, Jun. 1973.



GARVIN, P. L. The definitional model of language. In: _____. *Natural language and computer*. New York: McGraw-Hill, 1963. p.3-22. Apud ALONSO, D.V. *Lenguajes Documentarios*. [S.l.]: Nueve Parhadigma, 1997. 159 p.

GINEZ DE LARA, M. L. *Representação e linguagens documentárias: bases teórico-metodológicas*. São Paulo: Universidade de São Paulo, 1999. 208 f. Tese (Doutorado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo.

_____, *A Representação Documentária: em jogo a significação*. São Paulo: Universidade de São Paulo, 1993. 133 f. Dissertação (Mestrado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade Paulista.

_____; et all. *Manual de elaboração de tesouros monolíngües*. Brasília: CNPq/PNBU, 1990. 78 p. FELBER, H. *Termologia manual*. Paris: UNESCO, 1984. 234 P. Apud CAMPOS, M. L. A. *Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia*. Rio de Janeiro, 1994. 196 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação da Universidade Federal do Rio de Janeiro, Universidade Federal do Rio de Janeiro.

_____, Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. *Ciência da Informação*, Brasília, v.22, n.3, p.223-226, set./dez. 1993.

IGLESIAS MATURANA, M. T. *Análisis y recuperacion de informacion documental*. Santiago: I.P.S. 1992. 237p.

INMON, W. H. O que é um *Data Warehouse*. Disponível em: < >. Acessado em: 05 fev. 2002.

_____. *The data warehouse toolkit: practical technique for building dimensional data warehouses*. 2. ed. New York: Jonh Wiley & Sons, [1996]. 388 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION - Terminologia - vocabulário. Proposta de versão brasileira. (ISO 1087 – 1994). Apud GINEZ DE LARA, M. L. *Representação e linguagens documentárias: bases teórico-metodológicas*. São Paulo: Universidade de São Paulo, 1999. 208 f. Tese (Doutorado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo.

KIMBALL, R. *A Definition of Data Warehousing*. Disponível em: < >. Acessado em: 08 de mar. 2002.

_____, *The data warehouse toolKit - pratical techiques for buildin dimensional data warehouse*, John Wiley & Sons, Inc., 1997.

KOBASHI, N. Y. Análise documentária e representação da informação. *Informare*, Rio de Janeiro, v.2, n.2, p.5-27, jul./dez. 1996.

_____. *A elaboração de informações documentárias: em busca de uma metodologia*. São Paulo, 1994. 195 f. Tese (Doutorado em Ciência da Comunicação) – Escola de Comunicação e Arte, Universidade de São Paulo.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Brasília: Brinquet de Lemos, 1993. 347 p.

_____. *El control del vocabulario en la recuperación de información*. 2. ed. Valencia: Universitat de València, 1992. 286 p.

LUCAS, Henry C. — *Information system. Concepts for management*. Singapura: McGraw Hill International Editions, 1987 (3ª ed.). In: SILVA, A. B. M. da, A Gestão da Informação Arquivística e suas repercussões na produção do conhecimento científico. Disponível em < >. Acessado em 08 de mar. 2002.

LYONS, J. *Semântica*. Lisboa: Martins Fontes, 1977. v.1. Apud CINTRA, A. M. et al. *Para entender as linguagens documentárias*. São Paulo: APB, 1994. 72 p.

MAREMPUDI, A. *Data Warehousing - For Better Business Decisions* Disponível em: <http://www.intellibusiness.com/DWWP_anju.htm>. Acessado em: 08 de mar. 2002.

MARI, H. 1996. Dos fundamentos da significação à produção do sentido. *Perspectivas em Ciência da Informação*, Belo Horizonte, v.1, n.1, p.93-109, jan./jun. 1996.

MOTTA, D. F. *Método relacional como nova abordagem para a construção de tesouros*. Rio de Janeiro: SENAI, 1987. p. 29-45.

NARDI, M. I. A. *A metáfora e a leitura como evento social: instrumentos do pensar a Biblioteconomia do futuro*. São Paulo, 1999. 254 f. Tese (Doutorado em Lingüística Aplicada e Estudos da Linguagem) – Pontifícia Universidade Católica, São Paulo.

NATALI, J. W. Documentação e lingüística: inter-relação e campos de pesquisa. *Rev. Brás. Bibliotecon. Doc*; São Paulo, v. 11. n. 1/2, p. 33-42, jan./jun. 1978.

NOVELLINO, M. S. F. Organização do conhecimento: as contribuições da ciência da informação e da lingüística no contexto de uma ação comunicativa. IN: SEMINÁRIO DE ESTUDOS DE INFORMAÇÃO, 1, 1997, Niterói. *Anais...* Niterói: EDUFF, 1997. 105-110p.

PICHT, H. La terminologia como factor de desarrollo. *Revista Interamericana de Bibliotecologia*, Medellin, Colombia, v.19, n.1, p.7-30, ene./jun. 1996.

PINTO, M. C. M. F. Análise e Representação de assuntos em sistemas de recuperação da informação: linguagens de indexação. *Rev. da Escola Biblioteconomia da UFMG*, Belo Horizonte, v. 14 n. 2, p. 169-186, set. 1985.

QUONIAN, L.; TARAPANOFF, K.; ARAÚJO JÚNIOR, R. H. de; ALVARES, L. Inteligência obtida pela aplicação de *data mining* em bases de teses francesas sobre o Brasil. Ciência da Informação, Brasília, v.30, n.2, p.20-28, maio/ago. 2001.

SAGER, J. C. Prólogo: la terminología, puente entre varios mundos. 1993. In: SLODZIAN, M. Transfert de connaissances nouvelles et aménagement terminologique. *Meta*, v.40, n.2, p.238-243. 1995.

SANTOS, R. N. M. dos. Sistemas de informações estratégicas para a vitalidade da empresa. Ciência da Informação, Brasília, v.25, n.1, p.12-14, jan./abr. 1996.

SILVA, A. B. M. da, A Gestão da Informação Arquivística e suas repercussões na produção do conhecimento científico. Disponível em < >. Acessado em 08 de mar. 2002.

SMIT, J. W., Documentação e lingüística: interrelação e campos de pesquisa. Revista Brasileira de Biblioteconomia e Documentação, v. 11, n. 1/2, 1978.

TÁLAMO, M.F.G.M. et. al. Contribuição da terminologia para a elaboração de tesouros, *Revista Ciência da Informação*, Brasília, v. 21, n. 3, p.197-200, set./dez., 1992.

_____. Informação: organização e comunicação. IN: SEMINÁRIO DE ESTUDOS DE INFORMAÇÃO, 1, 1997, Niterói. Anais... Niterói: EDUFF, 1997a. 11-14p.

_____. Linguagem documentária. São Paulo: APB, 1997b. 12p. (Ensaio APB, 45)



APÊNDICE A

Instruções aos participantes da pesquisa sobre a técnica de coleta de dados do protocolo verbal ou “pensar alto”

Fazer-se uso de uma técnica de coleta de dados na minha pesquisa e para isso solicita-se que criem um padrão textual, com os campos ou colunas que devem constar no documento para que possam executar suas tarefas diárias; precisa-se que vocês criem esse documento, falando-se tudo que estão pensando durante o processo de criação, para que se possa registrar o que pensaram na definição das informações descritas no do documento criado.

Tente imaginar-se sozinho no departamento em que trabalha e que necessita do relatório dos pedidos em carteira no sistema de informação da empresa; que precisa de uma solução para poder administrá-los com maior eficiência, e nesse documento deverá conter as informações necessárias.

Para que os participantes não se sintam inibidos, fez-se um pequeno exercício, criando um relatório de um cadastro de clientes, utilizando a técnica de “pensar alto” e ao mesmo tempo, escreveram no papel os campos de informações da forma mais adequados, a uma necessidade imaginária.

Como usuários e trabalhadores no setor de vendas, solicita-se que criem um relatório de pedidos em carteira com no mínimo doze colunas e ao fazê-lo, procurem exteriorizar o pensamento durante a realização da tarefa, com um processo de pensar diretamente ligado à fala.

Não devem preocupar-se com a presença do pesquisador e muito menos com o gravador; o conteúdo dessa coleta será para fins de pesquisa científica e não serão divulgados os nomes reais dos participaram da pesquisa.

APÊNDICE B

TRANSCRIÇÃO DO PROTOCOLO VERBAL COM AS NOTAÇÕES DO USUÁRIO INFORMANTE 1:

Bom, nós vamos é ... *representante* que é para você saber a composição da sua carteira, quem está compondo a sua carteira, distribuição geográfica, visualizando sabe o representante tal os pedidos vão estar dentro de uma certa região, representante, os dados cadastrais, (o quê do representante), o *código do representante e o nome do representante*, CÓDIGO E O NOME DO REPRESENTANTE, é ... agora os *dados do cliente*, CLIENTE, que seriam os dados cadastrais, que seriam a razão social e o local ... *código, nome, cidade*, é ... depois disso ... *número do pedido*, PEDIDO, para facilitar uma consulta que queira fazer de tela essas coisas, NÚMERO DO PEDIDO, os *itens que compõem o pedido*, COMPÕE O PEDIDO, itens daí já *descrição, preço, quantidade*, já mata tudo no item, (se você quiser detalhar pode, porque aí você está falando do pedido propriamente dito), itens que compõem o pedido ... e aí ... aqui seria um sub itens, sub itens, é ... itens preço, eu preciso disso, é... *condição de pagamento*, PREÇO CONDIÇÃO DE PAGAMENTO, se ... *prazo de entrega*, ele pode não estar numa condição de não estar disponível para entrega imediata, então é uma informação importante, PRAZO DE ENTREGA, prazo de entrega, (.....) que mais poderia entrar aí, (.....) seria interessante por representante, (você não pôs por exemplo a descrição do produto), itens que compõem o pedido, esses itens são a descrição do produto,(descrição quantidade), é quantidade , preço né, é que em itens eu já considerei o próprio produto, então são itens que compõem o produto, descrição, preço, condição de pagamento prazo / uma informação que não tem em nossa carteira que seria interessante é ... totalizador por peso, peso por representante, REPRESENTANTE, porque como você trabalha, [você tem representantes que trabalham numa área fechada, você sabe se a carteira dele comporta uma carga, sem que você tenha que ficar somando as quantidades para saber se já tem uma carga para a região dele, é uma informação que não tem hoje e que seria interessante], então abrir no final de cada representante abrir um campo lá onde se fala total de peso que você tem, é ... esse [total de peso aí se você tiver condição de prazos de entrega diferenciados, se já criaria um outro ... condições diferentes, você vai ter mercadoria se está disponível para entrega imediata e tem mercadoria que você só vai poder faturar somente daqui uma semana, por exemplo, a condição

negociada foi, (então tem que levar em consideração prazo de entrega), prazo de entrega que é esse totalizador de peso aí, (seria um item também do relatório para você ver qual que é o prazo do entrega), há porque o prazo de entrega está aqui, onde coloquei que itens que compõem o pedido são ... descrição, preço, condição de pagamento e prazo de entrega, ou seja são condições negociadas na venda, preço como condição de pagamento e prazo de entrega são condições negociadas na venda], é ... *totalizador de peso por representante* é ... aí você poderia abrir para embarque imediato. imediato, ... como se não fosse para, para embarque imediato seria , dependendo da negociação, você ter um pedido na carteira para faturamento semana que vem e outro pra depois, esse seria outros e tal, outros /, seriam outros totalizadores, ou seja, (.....) você tem um caminhão para embarque imediato e tem mais duas cargas com prazo diversos, que dizer você tem que entrar no pedido para saber se você compõe a carga, ou você não vai ter que ficar né (sem essa visão) sem essa visão, é ... que mais que mais poderíamos, (bom se é uma entrega CIF ou FOB, aquelas coisas que têm que ter) , há sim (.....) *prazo de entrega* e condição de entrega e condição de entrega (.....) é.... *condição de entrega CIF ou FOB* (.....), [pensar em mais informações que possam ter no pedido que tenha que ser ... (ligada a situação financeira do cliente na empresa), há exatamente], é ... se ele tem algum bloqueio de crédito né / *bloqueio por motivo de crédito*, é ... TIPO DE BLOQUEIO né, porque por motivos de crédito ele pode ser por inadimplência, por ser por há / ele tem um limite de crédito está extrapolando isso (.....), (você gostaria que aparecesse a duplicata que é a causa da inadimplência para ser analisada, se é problema de baixa, se o cliente não pagou mesmo), é ... [é interessante essa informação né porque ... uma duplicata que esteja bloqueando esse pedido pode ser por exemplo uma baixa ocorrida ontem], (.....) (aqui no relatório como você chamaria isso data da ...), é ... (.....) *duplicata vencida e não baixada* (.....) e *data do vencimento* (.....), seria um complemento dessa duplicata vencida e não baixada é *média de atraso do cliente no semestre*, se o cara tem uma média de sete dias de atraso, você já tem uma noção de que aquela duplicata está vencida e não paga porque ele já é um atrasildo (.....), [o conceito dele, conceito porque são coisas distintas aqui você pode ter é no caso de uma duplicata paga e não baixada ou em atraso até um certo um certo período ele tem como conceito suspenso, certo, o crédito dele está suspenso, só que ele pode estar em outra condição ele pode estar no conceito de rejeitado, ou seja, já não cabe análise de pagou ou não pagou e tal e o próprio conceito dele já ... bloqueia *a priori* o pedido dele tal, você olhando o conceito você já irá saber que esse cliente vai está bloqueado] (.....) *conceito de crédito do cliente*,

bom, vou usar os termos que nós usamos aqui que é, rejeitado (....) [prá que você não descuide do período do tempo em que o pedido está em carteira (.....) é uma informação da data do ... cadastro do pedido lá né], *data de inclusão do pedido* né, (....) [aqui estabelece um sistema (.....) que estabelece uma data limite para faturamento, e partir dali ele já fica com data excedida, aparece uma observação que ele está com data excedida, então você já batendo o olho já sabe que esse pedido é antigo] (.....) *data limite para faturamento* (.....).



APÊNDICE C

TRANSCRIÇÃO DO PROTOCOLO VERBAL COM AS NOTAÇÕES DO USUÁRIO INFORMANTE 2:

Bom, gostaria de um relatório em que constassem os pedidos em carteira né, então vejamos, relatório que separasse primeiro os *representantes*, nesses representantes teria que aparecer o *número do pedido*, a *data do pedido* ... e também ... *PRODUTO* ... nesse produto tem que constar o *preço* ... *quantidade*, [*número do pedido*, *data do pedido*, *preço* e *quantidade*], ... nome do cliente ... e o código, primeiramente o código do cliente e o nome do cliente ... aí o sistema teria que pegar REPRESENTANTE POR REPRESENTANTE, tá, [cada representante tem um código, representante um, representante dois, aí apareceria o número do pedido, data do pedido, produto, preço e a quantidade], ... (que mais informações você acha relevante), bom aí no teria que aparecer ... (.....) também tem que ter o prazo que ele comprou, quantidade além de quantidade tem que ter o *prazo*, [você comprou com quantos dias né, (prazo de entrega ?) não prazo de compra, ele comprou com quatorze dias, (ah tá), você comprou com quantos dias, quatorze ou vinte e oito dias, (como ele vai pagar isso), isso (são prazos de), CONDIÇÃO DE PAGAMENTO né], ... essa coluna é óbvia que tem que constar no relatório, representante tal, número pedido (.....), (alguma sobre a ficha financeira do cliente), ficha financeira ... não digo a ficha financeira mas sim ... o conceito dele ... conceito, se ele é bom, se é rejeitado, ruim, e o LIMITE, *limite de crédito* também, [é obvio que se o cara estiver rejeitado, vai aparecer lá na frente como rejeitado,... baseado no limite de crédito, se o pedido vai ultrapassar o limite de crédito, aparecesse lá excedido ... (pode relacionar embaixo) (.....), acho que o que tiver dentro do padrão não vai aparecer nada], (.....) (esse relatório é quase um espelho do seu trabalho hoje, tem alguma coisa que poderia te ajudar a melhorar o seu trabalho ?) ... deixa eu ver ... (.....) acredito que esteja tudo aqui já, (.....) eu acho que é isso aí mesmo.

APÊNDICE D

PADRÃO TEXTUAL PARA LAY-OUT DO DOCUMENTO DO USUÁRIO INFORMANTE 1:

1. Representante (Código e Nome);
2. Dados do Cliente (Código/Nome/Cidade);
3. Número do Pedido;
4. Itens que compõem o pedido;
5. Descrição;
6. Preço;
7. Condição de Pagamento;
8. Prazo de entrega;
9. Totalizadores : de peso por representante;
10. Para entrega imediata;
11. Outros;
12. Condição de entrega (CIF/FOB);
13. Tipo de bloqueio por motivo de crédito;
14. Duplicatas vencidas e não baixadas e data do vencimento;
15. Média de atraso do cliente no semestre;
16. Conceito do crédito do cliente (bom/ruim/rejeitado);
17. Data de inclusão do pedido;
18. Data limite para faturamento.

APÊNDICE E

PADRÃO TEXTUAL PARA LAY-OUT DO DOCUMENTO DO USUÁRIO INFORMANTE 2:

1. Representante;
2. Número do pedido;
3. Código e Nome do Cliente;
4. Data do pedido;
5. Produto;
6. Preço
7. Quantidade;
8. Prazo (Condição de Pagamento);
9. Conceito;
10. Limite de entrega;
11. Limite excedido.

ANEXO A - ESTRUTURA DAS TABELAS DE DADOS DO DATA WAREHOUSE

Entidade	Atributo	Tp/Tam	Descrição
AgenteCobrador	IdAgente	Numeric(3)	Código de identificação do agente cobrador
	Nome	Character(5)	Nome do agente cobrador
Classificacao	IdClassificacao	Numeric(3)	Código de identificação da classificação
	Descricao	Character(30)	Descrição da classificação
Comprador	IdComprador	Numeric(3)	Código de identificação do cobrador
	Nome	Character(40)	Nome do comprador
Compras	Quantidade	Numeric(14,3)	Quantidade a ser comprada do produto
	ValorTotal	Numeric(14,2)	Valor total da compra
	Descontos	Numeric(14,2)	Descontos efetuados na compra
Consumo	Quantidade	Numeric(14,3)	Quantidade consumida do produto
Destino	IdDestino	Numeric(5)	Código de identificação do destino
	Cidade	Character(20)	Cidade destino
	Estado	Character(2)	Estado destino
Filial	IdFilial	Numeric(2)	Código de identificação da filial
	Nome	Character(15)	Nome da filial
Fornecedor	IdFornecedor	Numeric(5)	Código de identificação do fornecedor
	Cidade	Character(20)	Cidade do fornecedor
	Estado	Character(2)	Estado do fornecedor
Frete	ValorFrete	Numeric(14,2)	Valor cobrado para o frete
	Peso	Numeric(14,3)	Peso transportado
	Quebra	Numeric(14,3)	Quebra ocorrida
Linha	IdLinha	Numeric(3)	Código de identificação da linha de produção
	Descricao	Character(30)	Descricao da linha
Localizacao	IdCidade	Numeric(5)	Código de identificação da cidade
	Cidade	Character(20)	Nome da cidade
	Regiao	Character(30)	Nome da regiao que a cidade pertence
	Estado	Character(2)	Estado em que a cidade se localiza
Origem	IdOrigem	Numeric(5)	Código de identificação da origem
	Cidade	Character(20)	Nome da cidade que a origem se estabelece
	Estado	Character(2)	Estado que a origem se estabelece
Pagamentos	Valor	Numeric(14,2)	Valor do pagamento efetuado
	Juros	Numeric(14,2)	Total de juros cobrado
	Desconto	Numeric(14,2)	Descontos efetuados no pagamento
	DiasAtraso	Numeric(6)	Número de dias de atraso

ANEXO B - LISTA DAS TABELAS COM OS ATRIBUTOS DAS DIMENSÕES

Entidade	Atributo	Tp/Tam	Descrição
Produto	IdProduto	Numeric(6)	Código de identificação do produto
	Descricao	Character(50)	Descrição do produto
	Marca	Character(10)	Marca do produto
	Grupo	Character(25)	Grupo que o produto pertence
Recebimentos	ValorFreteRecebido	Numeric(14,2)	Valor recebido no frete
	JurosRecebido	Numeric(14,2)	Juros aplicados no valor do frete
	Descontos	Numeric(14,2)	Descontos efetuados no valor do frete
	DiasAtraso	Numeric(6)	Número de dias de atraso
Setor	IdSetor	Numeric(5)	Código de identificação do setor
	Descricao	Character(30)	Descrição do setor
Tempo	IdTempo	Numeric(4)	Código de identificação do tempo
	Data	Date	Data do tempo
	Dia	Character(7)	Nome do dia na semana
	Semana	Numeric(2)	Numero da semana no ano
	Mes	Character(8)	Numero do mês no ano
	Trimestre	Numeric(1)	Número do trimestre no ano
	Ano	Numeric(4)	Número do ano
TipoCarga	IdTipoCarga	Numeric(2)	Código de identificação do tipo de carga
	Descricao	Character(30)	Descrição do tipo de carga
Transportador	IdTransportador	Numeric(7)	Código de identificação do transportador
	Nome	Character(40)	Nome do transportador
	Cidade	Character(20)	Cidade do transportador
	Estado	Character(2)	Estado do transportador
Turno	IdTurno	Numeric(2)	Código de identificação do turno
	Inicio	Date	Data de início
	Termino	Date	Data de término
Venda	Quantidade	Numeric(14,3)	Quantidade de produtos vendidos
	Valor	Numeric(15,2)	Valor da venda
	ValorComissao	Numeric(15,2)	Valor pago de comissão
	Peso	Numeric(15,3)	Peso transportado
Vendedor	IdVendedor	Numeric(3)	Código de identificação do vendedor
	Nome	Character(30)	Nome do vendedor
ViaTransporte	IdVia	Numeric(1)	Código de identificação da via de transporte
	Descricao	Character(15)	Descrição da via de transporte

ANEXO C

Notações da transcrição:

Itálico ----- verbalização de termos exatamente como foram transcritos no documento gerado.

MAIÚSCULAS --- termos relevantes baseados na conexão mental dos usuários.

() ----- comentários de auxílio do pesquisador.

[] ----- esquemas e subesquemas associados ao referente de um termo

Reticências ----- pausas na fala (sem considerar tempo de duração)

/ ----- auto-interrupção de pensamento / fala

// ----- interrupção de pensamento/fala pelo interlocutor

(.....) ----- omissão de trecho não relevante da transcrição

Obs.: Notações adaptados, Nardi (1999, p.138)

GLOSSÁRIO

background - fundo, regras de tratamento da informação que não visíveis na interface.

Commodities - é uma linha de crédito que envolve o empréstimo de recursos de bancos estrangeiros para a produção rural destinada à exportação. Fortuna E., Mercado Financeiro Produtos e Serviços, Ed. Qualitymark, 1996.

Data Mart - sistema de armazenamento de dados setorial para suporte a tomada de decisão.

Data Mining - ferramenta de mineração de dados em bancos de dados relacional.

Data Warehouse - sistema de armazenamento de dados empresarial para suporte a tomada de decisão.

Delphi - Nome de ferramenta técnica para desenvolvimento de sistema em linguagem de programação.

front-end - Interface de apresentação do sistema para seus usuários.

insights - idéias - utilizado no surgimento de novas idéias.

lay-out - Formato de apresentação das informações.

Novell - Nome do sistema operacional para computadores ligados em rede.

PARADOX - Nome de ferramenta técnica para desenvolvimento de sistema em linguagem de programação.



cm

1

2

3

4

5

6

7

8

unesp

9

13

14

15

16

17

18

19

20