“Práticas Participativas na Elicitação de Requisitos para *Database Marketing*”

Fernanda Regina Benhami dos Santos

Trabalho Final de Mestrado Profissional em Computação

Instituto de Computação
Universidade Estadual de Campinas

"Práticas Participativas na Elicitação de Requisitos para *Database Marketing*"

Fernanda Regina Benhami dos Santos

Julho de 2006

Banca Examinadora:

- **Profa. Dra. Maria Cecília Calani Baranauskas (Orientadora)**
  Instituto de Computação - UNICAMP
- **Prof. Dr. Hans Liesenberg**
  Instituto de Computação – UNICAMP
- **Dr. Rodrigo Bonacin**
  Centro de Pesquisas Renato Archer - CenPRA

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP
Bibliotecária: Maria Júlia Milani Rodrigues – CRB8a / 2116

Santos, Fernanda Regina Benhami dos

Sa59p Práticas participativas na elicitação de requisitos para *database marketing* / Fernanda Regina Benhami dos Santos -- Campinas, [S.P. :s.n.], 2006.

Orientadora : Maria Cecília Calani Baranauskas

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Computação.

1. Marketing de banco de dados. 2. Engenharia de software. 3. Banco de dados. I. Baranauskas, Maria Cecília Calani, 1954-. II. Universidade Estadual de Campinas. Instituto de Computação. III. Título.

Título em inglês: Participatory design techniques used in requirements elicitation for database marketing.

Palavras-chave em inglês (Keywords): 1. Database marketing. 2. Software engineering. 3. Database.

Área de concentração: Engenharia da Computação

Titulação: Mestre em Computação

Banca examinadora: Prof. Dr. Hans Liesenberg (IC-UNICAMP)
Dr. Rodrigo Bonacin (CenPRA)

Data da defesa: 31-07-2006

Programa de Pós-Graduação: Mestrado em Computação

“Práticas Participativas na Elicitação de Requisitos para *Database Marketing*”

Este exemplar corresponde à redação final do Trabalho Final devidamente corrigido e defendido por Fernanda Regina Benhami dos Santos e aprovado pela Banca Examinadora.

Campinas, 31 de Julho de 2006.

Profa. Dra. Maria Cecília Calani Baranauskas.
(Orientadora)

Trabalho Final apresentado ao Instituto de Computação, UNICAMP, como requisito parcial para a obtenção do título de Mestre em Computação na área de Engenharia de Computação.

TERMO DE APROVAÇÃO

Trabalho Final Escrito defendido e aprovado em 31 de julho de 2006, pela Banca Examinadora composta pelos Professores Doutores:

Prof. Dr. Rodrigo Bonacin
CenPRA

Prof. Dr. Hans Kurt Edmund Liesenberg
IC – UNICAMP

Profª. Drª. Maria Cecília Calani Baranauskas
IC – UNICAMP

© Fernanda Regina Benhami dos Santos, 2006.

Todos os direitos reservados

Ao meu marido Wagner, aos meus pais Elza e José Carlos.

Agradecimentos

A todos os meus amigos, colegas de trabalho, familiares que de alguma forma colaboraram com a realização deste trabalho.

Ao meu marido, por estar sempre ao meu lado a cada conquista.

Aos meus pais por acreditarem e apoiarem todas as realizações da minha vida.

Aos colegas do mestrado, alunos e professores, que agregaram muito ao conhecimento adquirido por mim nesta fase.

Aos colegas do grupo de estudos, coordenado pela Professora Cecília Baranauskas, que apoiaram, comentaram e agregaram seus conhecimentos, para a produção final deste trabalho.

A oportunidade de poder trocar informações e conhecimento com pessoas tão especiais.

Obrigada a todos pela contribuição e apoio na realização deste sonho.

Resumo

O uso de informações pelas organizações para criarem estratégias e estabelecerem longos relacionamentos com seus clientes, cria a necessidade da construção de um *Database Marketing* para dar suporte a estas iniciativas.

Muitas estratégias deste tipo não foram bem sucedidas por não disponibilizarem dados e informações sobre seus clientes, que realmente fossem úteis para as análises realizadas.

Este trabalho propõe uma abordagem que utiliza práticas participativas como *Group Elicitation Method*, *BluePrint Mapping* e PICTIVE para compor uma metodologia aplicada à fase de elicitação de requisitos, a fim de minimizar as falhas ocorridas nesta fase, que poderiam se propagar aos resultados esperados pelas organizações.

Um estudo de caso demonstra o uso destas técnicas em conjunto com os comentários sobre os resultados obtidos apontando os pontos fortes e fracos, além da proposta de trabalhos futuros que podem ser realizados a partir dos resultados apontados.

Abstract

The use of information by organizations to define strategies and establish a long-term relationship with their customers, generates the need of having a Database Marketing to give support to those initiatives.

Many of these strategies did not succeed because of a lack of data and information about their customers, that were not available to be used in the analysis.

This essay suggests an approach based on participatory design (PD) techniques such as *Group Elicitation Method*, *BluePrint Mapping* and PICTIVE to create a methodology applicable to the requirements elicitation phase, in order to minimize missing information that occurs regularly in this phase, and could be propagated to the results expected by organizations.

A case study demonstrate the use of this methodology based on PD techniques with comments about results obtained, including a list of weak and strong points of this methodology, and also a proposal for future work that can be done based on the results listed herein.

Sumário

Resumo ..... *x*

Abstract ..... *xii*

Lista de Figuras ..... *xvi*

Lista de Tabelas ..... *xviii*

Tabela de Siglas ..... *xix*

1 Introdução *1*

2 O Cenário do Trabalho e a Problemática *3*

2.1 O que é um Database Marketing e sua finalidade **3**

2.2 Tipos de informação de um DBM **5**

2.2.1 Informações típicas de um DBM 5

2.3 Cenário usual para obtenção das informações do cliente **6**

2.4 Os *stakeholders* de um DBM e seus papéis **8**

2.5 Estado corrente de elicitação de requisitos para um DBM **8**

2.5.1 Documentação do método utilizado para elicitação de requisitos 9

2.5.2 Participação dos envolvidos 10

2.6 Problemática da elicitação de requisitos de Database Marketing **10**

3 Fundamentos do Domínio de Database Marketing *13*

3.1 Dados e Banco de Dados **13**

3.2 *Data Warehouse* e sua Modelagem **21**

3.2.1 Visão Inmon 24

3.2.2 Visão Kimball 26

3.2.3 Diferença entre as abordagens Inmon e Kimball 27

3.3 Database Marketing **28**

3.4 Mineração de Dados *(Data Mining)* **28**

3.5 Considerações Finais sobre os conceitos apresentados **30**

4 Elicitação de Requisitos para um DBM: Abordagem Proposta *31*

4.1 O Processo de Engenharia de Requisitos **31**

4.2 Práticas Participativas na Engenharia de Requisitos **33**

4.2.1 Group Elicitation Method 34

4.2.2 BluePrint Mapping 35

4.2.3 PICTIVE (Plastic Interface for Collaborative Technology Iniciatives through Vídeo Exploration) 36

4.3 Uma Proposta Participativa para Elicitação de Requisitos de um DBM **37**

5 Estudo de Caso: DBMTech 39

5.1 Cenário atual da DBMTech 39

5.1.1 Equipe de Trabalho 40

5.1.2 Necessidades da DBMTech 41

5.2 Abordagem Proposta Aplicada ao Estudo de Caso 46

5.2.1 O uso de Práticas Participativas 47

5.3 Visão Geral do Processo 52

5.3.1 Primeira Oficina de Trabalho 52

5.3.2 Segunda Oficina de Trabalho 54

5.3.3 Terceira e Quarta Oficinas de Trabalho 59

5.4 Processos e Modelo de Dados proposto (*StarSchema*) para a DBMTech 74

6 Discussão, Conclusão e Trabalhos Futuros 79

6.1 Discussão dos Resultados Obtidos 79

6.1.1 Pontos Fortes no Uso das Práticas Participativas 80

6.1.2 Pontos Fracos no Uso das Práticas Participativas 81

6.1.3 Trabalhos Futuros 82

7 REFERÊNCIAS 83

8 Anexo I 87

Lista de Figuras

Figura 3-1: Modelo Dimensional e Modelo Relacional [24].....15
Figura 3-2: Representação da tabela *lookup* [14]. 17
Figura 3-3: Representação de um *star schema* [17]. 18
Figura 3-4: Representação de um *snowflake schema* [17].....19
Figura 3-5: Arquitetura típica de um *Data Warehouse* [17].....22
Figura 3-6: Uso da Informação em um *Data Warehouse* [17]. 23
Figura 3-7: Arquitetura de um Data Warehouse Corporativo [23].....25
Figura 3-8: Arquitetura de um Data Warehouse baseada em Data Marts [23] 27
Figura 3-9: Ciclo de mineração de dados [16].....29
Figura 4-1: Processo de engenharia de requisitos.....31
Figura 4-2: Proposta Participativa na Engenharia de Requisitos.....37
Figura 5-1: Organograma da DBMTech.....40
Figura 5-2: Ilustração do método criado para elicitação de requisitos com uso das técnicas participativas. 48
Figura 5-3: Diagrama gerado durante as sessões onde o *BluePrint Mapping* foi utilizado e resultado após o término das sessões. 50
Figura 5-4: Quadro de Mapeamento das Entidades e seus atributos. 51
Figura 5-5: Organização das Oficinas de Trabalho 52
Figura 5-6: Artefato produzido na segunda oficina de trabalho por meio da técnica *BluePrint Mapping*. 55
Figura 5-7: Processo 1 - Geração dos Relatórios de Análise de Receita por Cliente. 56
Figura 5-8: Processo 2 - Geração da lista de contatos atualizados baseado nas últimas atividades em que cada contato participou. 57
Figura 5-9: Processo 3 - Geração do Relatório de Custos por Atividade de Marketing. 58
Figura 5-10: Processo 1 - Geração dos Relatórios de Análise de Receita por Cliente com identificação das entidades. 60
Figura 5-11: Processo 2 - Geração da lista de contatos atualizados baseado nas últimas atividades em que cada contato participou, com identificação das entidades. 61
Figura 5-12: Processo 3 - Geração do Relatório de Custos por Atividade de Marketing, com identificação das entidades. 61
Figura 5-13: Relatório A - Evolução da Receita (de 2001 a 2004). 63
Figura 5-14: Relatório B - Clientes por Segmento (de 2001 a 2004). 63
Figura 5-15: Relatório C - Clientes por Produto/Solução. 64
Figura 5-16: Fonte de Dados para o Relatório D - Tabela de Clientes e Produtos.....64
Figura 5-17: Revenue Weathervane – Relatório gerado a partir dos dados apresentados na Figura 5-16.....65
Figura 5-18: Processo 1 – redesenhado com base nas alterações identificadas na terceira oficina de trabalho. 66
Figura 5-19: Relatório Final – baseado em informações dos relatórios B,C e D. 67
Figura 5-20: Relatório F - Participação em Ações de Marketing. 71
Figura 5-21: Diagrama resultado dos trabalhos com PICTIVE, fonte para o StarSchema. 72
Figura 5-22: Processo 1 - Geração dos Relatórios de Análise de Receita por Cliente. 75
Figura 5-23: Processo 2 - Geração da lista de contatos atualizados baseado nas últimas atividades em que cada contato participou. 76

Figura 5-24: Processo 3 - Geração do Relatório de Custos por Atividade de Marketing. 76
Figura 5-25: Modelo de Dados Conceitual (*StarSchema*) de Custos. 77
Figura 5-26: Modelo de Dados Conceitual (*StarSchema*) de Faturamento. 78

Lista de Tabelas

Tabela 5-1: Lista de entidades levantadas inicialmente 42
Tabela 5-2: Análise de Cliente/Receita por período anual 53
Tabela 5-3: Assunto: Análise de Cliente/Receita por trimestre 53
Tabela 5-4: Custos (Trimestral, Mensal e Anual) 53
Tabela 5-5: Atividades 53
Tabela 5-6: Relatórios analisados no Processo 1 56
Tabela 5-7: Relatórios analisados no Processo 2 57
Tabela 5-8: Relatórios analisados no Processo 3 58
Tabela 5-9: Lista Geral de Contatos 69
Tabela 5-10: Relatório “Ações de Marketing” – Lista 69
Tabela 5-11: Relatório E - Contatos por Atividade 69
Tabela 5-12: Relatório F - Custos de Marketing por Ação 70

Tabela de Siglas

Sigla	**Significado**
CRM	*Customer Relationship Management*
DBM	*Database Marketing*
ETL	*Extraction, Transformation and Load*
HTML	*Hyper Text Markup Language*
MDB	*Multidimensional Database*
OLAP	*On-Line Analatyc Processing*
PICTIVE	*Plastic Interface for Collaborative Technology Iniciatives through Vídeo Exploration*
RDBMS	*Relational Database Managament System*
SFA	*Sales Force Automation*
SIM	aplicação de controle de eventos e atividades do *marketing*

1 Introdução

A necessidade de obtenção de informações para tomada de decisão tem levado as empresas a desenvolver diversos projetos para que as informações estejam disponíveis e não tomem tempo para buscá-las; no caso de uma organização, um *Data Warehouse*, no caso do Departamento de *Marketing*, um *Database Marketing*, entre outras opções. Para a criação de qualquer destes repositórios de dados que poderão fornecer rapidamente informações estratégicas e ainda dar suporte a análises de comportamento ou de tendências, as organizações buscam empresas especializadas para a criação destes repositórios. Muitas vezes estas empresas não possuem uma metodologia adequada o que leva à ausência de informações importantes que não foram identificadas na fase de elicitação de requisitos. Esta falha no processo de elicitação de requisitos pode gerar falhas no produto final entregue ao usuário, principalmente:

1. Ausência de informações necessária sobre os clientes
2. Ausência de conhecimento das análises desejadas
3. Dificuldade de obtenção dos dados sobre os clientes

Cada um dos fatores citados levam ao mau entendimento, durante a fase de elicitação de requisitos, que se propaga por todo o projeto, até a sua implementação, caso não seja esclarecido e ajustado na fase de elicitação. Um problema a ser considerado é a ausência de uma metodologia espiral que facilite esta validação durante a fase de elicitação o que minimizaria grande parte do re-trabalho que acontece na fase de implementação, por não satisfazer as necessidades dos usuários. Uma metodologia sistemática que possa oferecer uma forma dinâmica de trabalho entre os *stakeholders* e o analista de requisitos, pode garantir um entendimento claro e de fácil ajuste para qualquer *stakeholder*, resultando um trabalho conjunto com uma visão única para todos aqueles que participaram do processo de elicitação.

Este trabalho propõe o uso de práticas participativas na elicitação de requisitos de sistemas computacionais adaptadas e utilizadas para a elicitação de requisitos de um DBM, promovendo um resultado participativo e consistente com a colaboração de todos os *stakeholders*. As práticas participativas *Group Elicitation Method*, *Blueprint Mapping* e PICTIVE foram utilizadas no estudo de caso que permitiu relatar a aplicação das mesmas na definição do modelo de dados para o DBM da DBMTech.

Este trabalho está organizado em capítulos da seguinte forma:

- Capítulo 2: descreve o cenário atual e a problemática do *Database Marketing*, são determinados os tipos de informações disponíveis para um DBM, os *stakeholders* envolvidos na criação e uso do DBM, e o estado corrente na elicitação de requisitos de um DBM, segundo experiência obtida em campo.

- Capítulo 3: aborda e define os fundamentos do domínio de *Database Marketing*, desde o fundamento base para sua construção, os dados e bancos de dados, assim como conceitos de *Data Warehouse* e suas diferentes abordagens, e também o uso do *Database Marketing* por meio de técnicas de *Data Mining* ou ainda exploração dos dados através de relatórios.

- Capítulo 4: apresenta a abordagem proposta para elicitação de requistos para *Database Marketing* fazendo uso de práticas participativas *Group Elicitation Method*, *Blueprint Mapping* e PICTIVE.

- Capítulo 5: descreve o estudo de caso que foi realizado de forma a aplicar as técnicas participativas e comprovar melhoras no processo de elicitação de requisitos realizados anteriormente pela DBMTech.

- Capítulo 6: apresenta as discussões e conclusões obtidas através do estudo de caso realizado, além da proposta de futuros trabalhos.

2 O Cenário do Trabalho e a Problemática

O cenário atual encontrado nas empresas relativo ao DBM é apresentado neste capítulo, assim como a sua finalidade, as informações contidas no DBM, os *stakeholders* envolvidos e a problemática envolvida na elicitação de requisitos.

2.1 O que é um Database Marketing e sua finalidade

No passado, a necessidade de contatar clientes e potenciais clientes (*prospects)* fazia com que as empresas buscassem listas de nomes no mercado para contatar um determinado público. Atualmente estas listas são enriquecidas para que se tenha informações que permitam um conhecimento mais profundo de seus clientes.

Esta base de dados enriquecida é chamada *database marketing* (DBM), que fornece informações para as ações de *marketing* de uma empresa, sejam ações de ofertas de produtos ou ações de relacionamento com o cliente.

Um DBM tem por finalidade apoiar as decisões da área de *marketing* em relação às ações de *marketing* direto, com índices satisfatórios de resposta.

A existência de um DBM em uma empresa permite a gestão de relacionamento com os clientes, ou CRM (*Customer Relationship Management*), sendo um trabalho contínuo, permanente, com resultados que somente irão surgir a médio prazo, porque demandam um arcabouço mais sofisticado, modelagens estatísticas, tratamento de dados, estabelecimento de pontos, montagem de matrizes etc. [1].

Um exemplo que pode ser citado é a venda de um produto numa instituição financeira. Suponha que um banco deseja oferecer cartões de crédito aos clientes que possuam conta corrente com o objetivo de aumentar o número de serviços ou produtos contratados por cliente e deve eleger aqueles clientes com perfil mais adequado a esta oferta; ou seja, os clientes que tem características similares àqueles que já possuem cartão de crédito, porque isto indica que estes clientes têm maior probabilidade de adquirir a oferta realizada. Para que esta análise possa ser realizada, existe a necessidade de uso de um DBM onde informações demográficas, informações de comportamento, informações psicográficas (aquelas coletadas em pesquisas sobre atitudes dos clientes), informações financeiras, entre outras, estão organizadas de forma que permitam

análises que demonstrem os perfis dos clientes e o analista de DBM possa decidir qual perfil é ideal para a oferta desejada. O uso do DBM nesta situação minimiza os gastos de *marketing* direto por meio de ações focadas nos públicos-alvo mais aderentes à oferta realizada, garantindo melhores índices de respostas.

A necessidade de mudança no foco do *marketing* direto demanda a criação de bases de dados voltadas ao cliente e não ao produto, de maneira que se possa identificar perfis e tendências a fim de atingir o público ideal em ações de *marketing*.

Para que um processo de *marketing* direto seja implementado com sucesso, é desejável [6]:

1. Identificar e coletar dados relevantes dos clientes e potenciais clientes (*prospects).*
2. Usar a tecnologia de banco de dados para transformar dados brutos em informações de *marketing* poderosas e acessíveis.
3. Aplicar técnicas estatísticas aos bancos de dados de clientes e potenciais clientes (*prospects).*
4. Avaliar a economia da coleta, da manipulação e da análise dos dados, e capitalizar a economia do desenvolvimento e da implementação de programas de *marketing* orientados pelos dados.
5. Agir criativamente nas oportunidades de *marketing* que surgem desses processos, para desenvolver relacionamentos com clientes individuais e construir negócios.

A construção de um *Database Marketing* não acontece da noite para o dia; existem processos criados para a obtenção de informações de clientes e potenciais clientes (*prospects*), existem testes de campanhas realizados junto aos clientes potenciais (*prospects*) para medir o suposto retorno de campanhas baseado nas respostas coletadas e, apesar disso, não temos os dados suficientes para o sucesso das campanhas. Esta construção é composta por ciclos onde registramos e enriquecemos os dados a cada campanha realizada ou oportunidade identificada de coleta de informação [3].

É importante que o *Database Marketing* esteja nas mãos da área de *marketing* ou vendas e nunca na área de sistemas. Ele não deve ser parte de outros sistemas, mas deve ser composto a partir da combinação dos dados dos diversos sistemas da empresa em forma de registros utilizáveis pelo *marketing*. Existe um grande incentivo para a construção de bancos de dados deste tipo para: proteger um investimento (os clientes da empresa), controlar a explosão das

informações sumarizando o conhecimento sobre seus clientes obtido de múltiplas fontes, dominar uma compreensão plena das forças e das fraquezas de uma empresa, avaliar completamente as linhas e os catálogos dos produtos atuais e passados, e fazer a empresa crescer [5].

2.2 Tipos de informação de um DBM

Os tipos de informação utilizados para *marketing* direto têm uma diversidade enorme de fontes de dados e uma grande quantidade em termos de volume. Os dados utilizados possuem informações das mais diversas, tais como: dados de comportamento, dados de desempenho do cliente (desenvolvido internamente pela empresa), dados geodemográficos, dados psicográficos, dados de estilo de vida, dados financeiros e dados de pesquisa (adquiridos através de pesquisas solicitadas a empresas especializadas) [6].

É importante salientar a necessidade de obter estas informações de acordo com a lei vigente que protege o consumidor [30], sendo assim, qualquer informação de um cliente obtida por uma empresa só poderá ser divulgada ou repassada a terceiros com a autorização prévia do próprio, além disso qualquer cliente poderá ter acesso as próprias informações que estejam em poder da empresa a qualquer momento e solicitar qualquer correção que julgue necessário.

2.2.1 Informações típicas de um DBM

1. **Dados de Clientes:** dados históricos de vendas e promoções de cada cliente relativo ao comportamento e desempenho junto à empresa [6].
2. **Dados de Pesquisa Primária:** fornecidos diretamente pelos indivíduos (clientes, clientes potenciais e até não-clientes) a respeito de si mesmos e classificados em dois tipos principais [6]:
    a. **Demográficos individuais:** estes dados são baseados em dados coletados em pesquisas realizadas diretamente junto aos indivíduos como, região onde reside, faixa etária, faixa salarial.
    b. **Comportamentais e de atitude:** estes dados são baseados em dados coletados em pesquisas realizadas diretamente junto aos indivíduos como, preferência

por que tipo de literatura, preferência por restaurante, freqüência de visita ao supermercado, entre outros.

3. **Dados de Pesquisa Secundária:** adquiridos de terceiros sobre clientes e potenciais clientes tanto no nível geográfico ou individual, ou seja, sob a visão de cada indivíduo ou de uma determinada região geográfica, classificados em quatro tipos principais:
    a. **Demográficos:** estes dados são baseados em dados do censo produzidos através de uma série de cálculos estatísticos como [6]: número de residências em um estado ou cidade ou bairro, quantidade de famílias, valor médio da renda, idade média dos indivíduos, entre outras.
    b. **Psicográficos:** estes dados tratam das atitudes, costumes e percepções das pessoas e não de características observáveis externamente [6]. Por exemplo, o registro da opinião da pessoa sobre um evento realizado, ou ainda sobre um produto ofertado, através de um questionário respondido pelo cliente.
    c. **Comportamentais e estilo de vida:** estes dados são baseados em pesquisas e tratam das diferenças entre os interesses pessoais dos indivíduos e de suas atividades de lazer.
    d. **Financeiro:** estes dados são baseados em informações financeiras adquiridas pelos *bureaus* de serviços financeiros [6] tais como: compras em cartão de crédito, parcelas de empréstimos, aplicações e histórico de pagamentos em instituições financeiras.

2.3 Cenário usual para obtenção das informações do cliente

Apesar dos muitos exemplos da comunicação e métodos de diálogo entre empresa e cliente, especialmente o *feedback* do cliente para a empresa que vende, não esgotamos este assunto; existem vários métodos para obtenção das valiosas informações de *marketing* provenientes dos clientes, de forma direta ou indireta [3].

Quando um cliente deseja comprar um item caro, ele fará alguma investigação em relação a modelos e preços antes de comprar, o que gera a oportunidade do vendedor descobrir o que realmente o cliente deseja. O vendedor sempre fará algumas perguntas que parecem uma mera conversa trivial mas na verdade são pesquisas, investigações de *marketing* que perguntam ao

cliente: “Diga-me, em que você está interessado e o que deseja? O que o motiva a comprar? O que preciso fazer para lhe vender um automóvel / uma casa / um barco / um anel de brilhantes? O que o levará a comprar de mim?”. As perguntas que parecem ser meramente casuais sobre as preferências do cliente com relação a cor, preço, custos operacionais, características, entrega, garantias, crédito e outros assuntos são a base para definição do perfil do cliente. A coleta destas informações pode não ser somente verbal, mas escrita e, ainda, adiciona informações sobre os desejos e intenções do cliente com base nas expressões e linguagem corporal expressada pelo cliente durante a tentativa de venda [3].

Existem vendedores, por sua vez, que não fazem esta pesquisa a fim de descobrir os desejos de seu cliente, simplesmente vendem aquele produto que o cliente se propôs a comprar sem tentar entender se de fato o cliente está satisfeito com a compra [3].

A decisão verdadeira de se construir um DBM deve ter implícita a completa disposição de fazer o que for necessário para obter os dados adequados à construção do banco de dados [3].

A coleta das informações depende de uma força tarefa muito grande de todos os envolvidos no processo para que os resultados apareçam nos estudos baseados nos dados coletados. Qualquer oportunidade de contato com o cliente deve ser considerada: uma feira, um seminário, uma visita, uma entrevista para a coleta das informações, afinal o cliente é a fonte destas informações.

Existem informações que podem ser obtidas por meio de pesquisas realizadas por empresas especializadas; isto ocorre quando a empresa possui algumas informações básicas de cada cliente e deseja enriquecer estas informações. As informações básicas como nome, endereço, idade, são enviadas a estas empresas especializadas que guardam dentro de seus bancos de dados gigantes todas as informações desejadas pela empresa solicitante. Para que este tipo de enriquecimento seja válido, a área de *marketing* precisa ter definido previamente seus objetivos de análise sobre estes dados; caso não tenha esta definição, os dados solicitados podem ser insuficientes para os resultados desejados das análises. Estas análises podem demonstrar as tendências de compra de um produto, o perfil de compra de um grupo de cliente, qual a penetração que um produto teria dentro de um segmento de clientes, entre outras.

2.4 Os *stakeholders* de um DBM e seus papéis

Stakeholders são as partes interessadas, ou seja, as pessoas ou organizações capazes de influenciar as definições sobre a funcionalidade e as restrições impostas a um sistema a ser desenvolvido. No caso de um DBM incluem usuários, gerentes, patrocinadores entre outros, envolvidos no processo organizacional. A não identificação correta dos *stakeholders* pode levar a mudanças constantes nos requisitos do sistema [4].

Relacionamos a seguir os principais envolvidos usualmente no processo de criação de um DBM:

Gerência de Vendas: diante da estratégia de vendas da empresa, esta área demanda para a área de *marketing* a escolha do melhor público para a venda dos produtos foco, com o objetivo de atingir as metas definidas.

Gerência de *Marketing*: viabilizar a estratégia de vendas por meio de ações de *marketing* direto.

Analista de *Marketing*: realiza as análises necessárias, com base no DBM, para apresentar um estudo que demonstre a viabilidade ou não de uma ação de *marketing*.

Gerência de TI: alinhada às necessidades da área de *marketing,* provê toda a infra-estrutura necessária para a realização de análises e armazenamento dos dados utilizados.

Analista de Sistemas: identifica as necessidades de enriquecimento dos dados junto à área de *marketing* e elabora documentos que permitam a implementação de novos dados no DBM acessado pelos analistas de *marketing*.

Programador: implementa os dados especificados pelos analistas de sistemas a fim de disponibilizar as informações necessárias aos analistas de *marketing*.

2.5 Estado corrente de elicitação de requisitos para um DBM

Utiliza-se usualmente entrevistas e questionários focados no negócio do cliente a fim de identificar os dados realmente necessários para um DBM. O analista de negócios conduz as entrevistas e os analistas de *marketing* e os analistas de sistemas são entrevistados para que as percepções de cada um sejam documentadas.

Os questionários são elaborados a partir de conhecimento adquirido em outros projetos ou de consultores que já atuaram na área de negócio da empresa onde queremos atuar. As questões normalmente são focadas nos objetivos comuns às áreas de *marketing* e nas análises que supostamente possam ser realizadas utilizando ferramentas estatísticas para a produção destes cenários de estudo. Existem também as questões relacionadas à área de tecnologia focando na maneira como estas informações serão obtidas, mantidas e disponibilizadas para a área de *marketing*.

Estes questionários são devolvidos, preenchidos e, após uma avaliação das respostas recebidas, já se pode entender, de forma horizontal, as análises desejadas, os tipos de informações disponíveis na empresa, as ações de *marketing* que serão executadas com base no DBM.

Num próximo passo, uma entrevista é realizada com cada pessoa responsável pelo desenvolvimento interno do DBM, das áreas de *marketing* e de tecnologia. Estas entrevistas são importantes para reforçar as informações prestadas no questionário pelos analistas de *marketing* e analista de sistemas que forem entrevistados para esclarecer os pontos que não tenham ficado claros nas respostas. Procura-se detalhar um pouco mais o que foi questionado a fim de desenhar, de alguma maneira, o modelo de dados preliminar que poderá ser implementado no DBM. Este modelo de dados depende muito do tipo de informação desejada pela área de *marketing* e do tipo de dados obtidos de fontes externas e da própria área de tecnologia da empresa.

Após as entrevistas, atas são geradas relatando o que foi entendido junto às pessoas entrevistadas e um plano de trabalho é elaborado. Basicamente, neste plano de trabalho, são relatadas as fontes de dados identificadas, os tipos de informações necessárias para as análises desejadas, e os tipos de transformações que devem ser feitas com os dados fontes para que se possa disponibilizar o dado utilizável para a área de *marketing*. São também listados os principais objetivos da área de *marketing* e como o DBM pode ajudar a atingir estes objetivos.

2.5.1 Documentação do método utilizado para elicitação de requisitos

A documentação produzida nem sempre é apresentada segundo um padrão; a cada caso reutiliza-se algum documento já existente adaptando-o para as necessidades daquela empresa.

A lista a seguir mostra alguns dos documentos utilizados:

1. Questionário

2. Atas de Reunião
3. Plano de Trabalho

Discutiremos no Capítulo 5.2 desta dissertação uma abordagem proposta para que o resultado da fase de elicitação de requisitos para DBM, de forma que seja objetivo e claro oferecendo os subsídios necessários para a fase de especificação. Atualmente a produção destes documentos não é suficiente para o prosseguimento do projeto, um trabalho de revisão é realizado para minimizar o re-trabalho na fase de implementação.

2.5.2 Participação dos envolvidos

A participação das áreas de *marketing* e tecnologia é imprescindível para um projeto de DBM; a primeira para a definição de suas próprias necessidades, a segunda para a viabilização técnica do projeto.

Em algumas empresas, nem sempre a área de negócio é envolvida, o que prejudica o resultado do trabalho, gerando insatisfação por parte da área usuária por não ter todos os dados que esperam ter disponíveis. Este cenário vem mudando com a necessidade cada vez maior da área de *marketing* de novas informações, muitas vezes a própria área de *marketing* contrata uma consultoria para a realização de um projeto de DBM com o apoio da área de tecnologia, o que faz com que a satisfação e o resultado sejam potencializados.

Para viabilizar tecnicamente um projeto de DBM, a área de tecnologia precisa conhecer a origem dos dados que serão disponibilizados para o *marketing*, a complexidade em extrair estes dados de suas origens e o esforço em consolidar estes dados de maneira a torná-los utilizáveis pelo *marketing*. Além disso, é preciso conhecer onde serão armazenados, qual a política de atualização, como serão disponibilizados os dados para exploração e análise da área de *marketing*.

2.6 Problemática da elicitação de requisitos de Database Marketing

Muitas vezes um projeto de DBM não é bem sucedido por uma série de fatores que enumeramos e discutimos a seguir:

1. **Ausência de informações necessária sobre os clientes** – ao entregar o DBM ao usuário final, este se depara com informações que não atendem suas necessidades de trabalho ou ainda, informações ausentes que impedem a realização das análises desejadas. Por exemplo: quantidade de contatos realizados anualmente por cliente; região de domínio do cliente, criada à partir dos seus hábitos de consumo, entre outras. Esta é uma falha típica da fase de elicitação porque o usuário tem dificuldade em reconhecer as informações que ainda não tem disponível.
2. **Ausência de conhecimento das análises desejadas –** muitas análises que podem ser realizadas com base no DBM acabam não sendo realizadas por falta de conhecimento dos usuários sobre o domínio de um DBM. Durante a fase de elicitação de requisitos o analista de requisitos pode aportar seu conhecimento e garantir um melhor resultado discutindo possibilidades de análises ainda não percebidas pelos usuários. Numa abordagem desestruturada, somente com base em questionários e entrevistas, alguns pontos de vista podem ser esquecidos e não utilizados numa análise geral de todos os usuários potenciais do DBM.
3. **Dificuldade de obtenção dos dados sobre os clientes –** o usuário final, normalmente, deseja ter uma série de informações cadastrais ou informações métricas sobre os clientes e de difícil acesso, o que causa insatisfação no resultado entregue em um DBM. Exemplo: valor das parcelas do financiamento de veículos; para obtenção destes dados a área de tecnologia deve criar um processo de extração de dados que faz a leitura da base de dados enviada pela empresa terceirizada que faz a administração deste financiamento. Neste caso, somente a área de tecnologia pode mostrar a complexidade do processo de obtenção da informação, o que ajuda a definir a expectativa junto ao usuário final em relação ao que pode ser ou não disponibilizado para as análises desejadas. É importante lembrar que envolver a área de tecnologia na fase de elicitação é imprescindível, porque esta é a área que mais conhece o ambiente computacional e pode alertar sobre a inexistência ou dificuldade na obtenção dos dados, o que permite gerar uma expectativa realizável para o usuário final.

Cada um dos fatores discutidos levam ao mau entendimento durante a fase de elicitação de requisitos que se propaga por todo o projeto até a sua implementação caso não seja esclarecido

e ajustado inicialmente. Um dos problemas observados é a ausência de uma metodologia que facilite esta validação durante a fase de elicitação, o que minimizaria grande parte do re-trabalho que acontece na fase de implementação por não satisfazer as necessidades dos usuários. Uma metodologia sistemática que possa oferecer uma forma dinâmica de trabalho entre os *stakeholders* e o analista de requisitos, pode garantir um entendimento mais claro e de fácil ajuste para qualquer *stakeholder*, resultando um trabalho conjunto com uma visão única para todos aqueles que participaram do processo de elicitação.

As práticas participativas aplicadas na elicitação de requisitos de sistemas computacionais podem ser adaptadas e utilizadas para a elicitação de requisitos de um DBM, promovendo um resultado participativo e consistente com a colaboração de todos os *stakeholders*.

O estudo de caso apresentado neste trabalho relata a aplicação de algumas técnicas participativas que permitem uma definição do modelo de dados para o DBM da DBMTech.

3 Fundamentos do Domínio de *Database Marketing*

Este trabalho aborda uma série de conceitos que serão discutidos ao longo desta dissertação e que são definidos brevemente neste capítulo.

Os conceitos abordados neste trabalho são fundamentados sobre repositórios de dados que vão desde o menor nível de informação, os dados, seguindo por banco de dados, modelos de dados lógicos e físicos e, atendendo às necessidades do negócio, os *Data Marts*, os modelos multidimensionais e os *Data Warehouses*, as análises de *Data Mining*.

Atualmente os repositórios de dados são utilizados principalmente nas seguintes aplicações [21]:

- *On-Line Analytic Processing* (OLAP) para suporte à decisão. Usuários realizam "*slice and dice*", conseguem explorar os dados no menor nível de detalhe possível por meio de combinações de filtros da forma como desejam.

- Bases de dados para suporte aos sistemas de informações estratégicas.

- *Data mining*, que utiliza o *Data Warehouse* como fonte de informações para sistemas de *Knowledge Data Discovery* (KDD) por meio de uma série de análises de inteligência artificial e técnicas estatísticas para encontrar associações, seqüências, classificações, segmentos e tendências.

- *Database marketing* (DBM), que utiliza o *Data Warehouse* para prover serviços customizados aos clientes de uma empresa.

Especificamente neste trabalho discutiremos sobre o DBM, os termos acima fornecem a base para conceituar, descrever e especificar um DBM.

Um DBM disponibiliza informações para que a área de *Marketing* realize análises como relatórios OLAP assim como *Data Mining* possibilitando o estudo do perfil dos clientes; estes conceitos também são comentados neste capítulo.

3.1 Dados e Banco de Dados

Um banco de dados é um dos meios de persistência dos sistemas de informações; estes dados são o ponto de partida, a fonte para todos os processos de geração de informação [22].

Um banco de dados estruturado e organizado adequadamente contém e representa os dados de uma organização e suportará grande parte dos processos de geração de informação. Os dados não devem ser simplesmente armazenados e guardados, mas devem ser armazenados para um propósito de uso definido pela organização, afinal dados devem ser utilizados.

Pode-se concluir que a maior razão para se ter um banco de dados é a persistência da informação e que o propósito de estruturar os dados em um banco de dados é a facilidade em organizar os dados agrupando os mesmos em informação significativa que geram novas informações.

Modelos de Dados Conceituais, Lógicos e Físicos

Existem 3 níveis de modelagem de dados: conceitual, lógico e físico [8]. Será descrita a diferença entre os 3 modelos, a ordem em que eles são criados, pois eles são criados de forma evolutiva até a concepção do modelo físico; também como partir de cada nível e evoluir para o próximo.

- **modelo de dados conceitual:** engloba as entidades significativas e os relacionamentos entre elas; não existe nenhum atributo e nem chave primária de tabela especificada. Este nível é o mais alto nível que permite ao modelador de dados identificar o relacionamento entre as entidades.

- **modelo de dados lógico:** engloba todas as entidades e os relacionamentos entre elas, a especificação de todos os atributos, de todas as chaves primárias, de todas as chaves estrangeiras e a normalização dos dados. Este nível é aquele que permite ao modelador de dados especificar o maior nível de detalhe possível, sem se preocupar com como isto será implementado no banco de dados.

- **modelo de dados físico:** engloba a especificação de todas as tabelas e colunas, a identificação de todas as chaves estrangeiras utilizadas para estabelecer os relacionamentos entre as tabelas, a denormalização dos dados conforme as agregações solicitadas pelo usuário e a análise das considerações físicas que podem determinar um modelo físico um pouco diferente do modelo lógico definido. Este nível é aquele que permite ao modelador de dados especificar como o modelo de dados lógico será implementado no banco de dados; este trabalho deve considerar a

conversão de entidades em tabelas, a conversão de relacionamentos em chaves estrangeiras e a conversão de atributos em colunas.

Modelos Dimensionais de Dados

Os modelos dimensionais são frequentemente utilizados em *Data Warehouse*s. Este modelo é diferente de um modelo relacional que utiliza a terceira forma normal (3NF) e é aplicado aos dados dos sistemas transacionais (OLTP). Esta forma de representação é muito detalhada e complexa para usuários de negócios interpretarem, pois exibe muitos detalhes dos relacionamentos entre as tabelas enquanto um modelo dimensional tem estes detalhes suprimidos, conforme observado na Figura 3-1. Neste exemplo a tabela fato é a COMPRA, que possui as dimensões TASK, FORNECEDOR e COMPONENTE associados a COMPRA no *Star Schema.* Podemos observar a mesma representação, porém baseada no modelo relacional de-normalizado pela 3NF, onde são acrescentadas as dimensões DEPARTAMENTO, PROJETO e REGIÃO; no caso do *Star Schema* estas dimensões adicionais não são representadas por terem seus atributos incluídos nas dimensões que estão relacionadas diretamente na tabela fato.

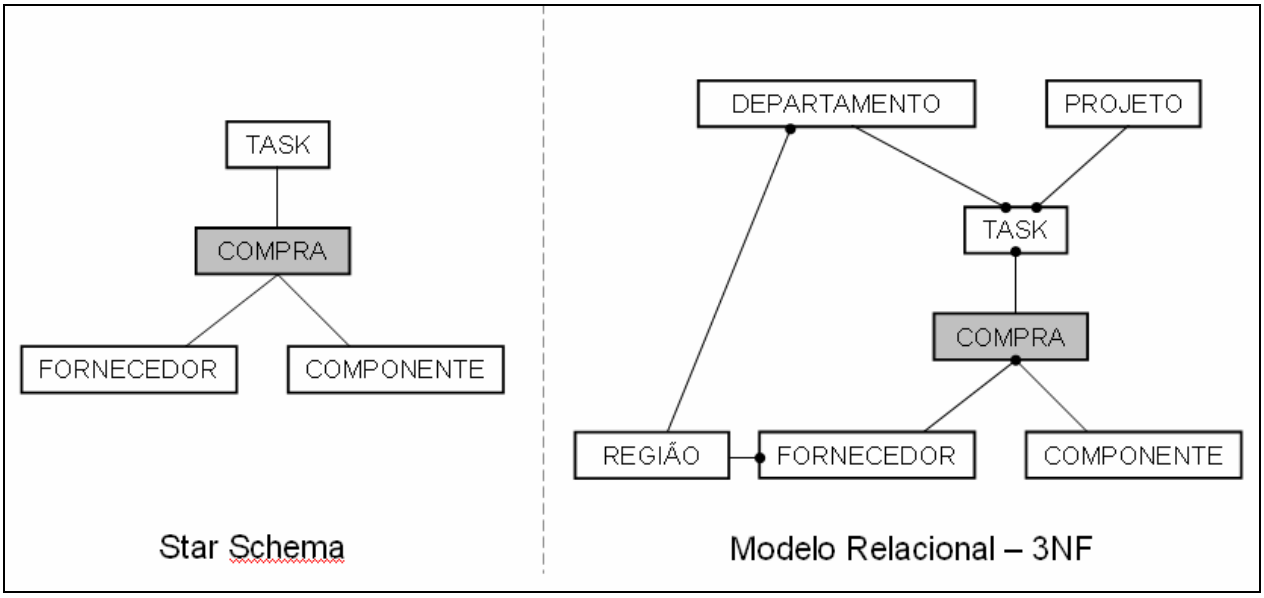


Figura 3-1: Modelo Dimensional e Modelo Relacional [24]

Nos modelos dimensionais, os dados são armazenados de forma diferente e simplificada, em forma de cubos de dados atendendo às necessidades específicas do usuário final [14]. Alguns termos específicos são utilizados neste tipo de modelagem, como segue [9]:

- **dimensão**: armazena as descrições textuais das dimensões de negócios e são regularmente utilizadas como restrições ou cabeçalhos num conjunto de respostas ao usuário [14]. Exemplo: dimensão tempo, dimensão cliente.

As dimensões podem ser classificadas como dimensões de modificação lenta, ou seja, aquelas que são praticamente constantes, e classificadas em 3 tipos [14]:

- as dimensões de Tipo Um são aquelas que têm todos os registros da tabela de dados substituídos a cada nova carga de dados; ou seja, não mantém nenhum histórico;

- as dimensões de Tipo Dois são aquelas que têm sempre um novo registro incluído na tabela de dados a cada valor modificado em relação ao original; ou seja, é possível pesquisar os registros e saber quando uma descrição foi alterada;

- as dimensões de Tipo Três são aquelas que tem um novo campo que representará o novo valor de um atributo existente na tabela de dados, ou seja, é possível pesquisar em cada registro o valor original, o valor anterior e o valor atual.

- **atributo**: único nível de dados numa dimensão. Exemplo, mês é um atributo na dimensão tempo.

- **hierarquia**: a especificação de níveis que representam o relacionamento entre diferentes atributos de uma hierarquia. Exemplo: uma hierarquia possível na dimensão tempo é Ano → Trimestre → Mês → Dia.

- **tabela fato**: é a tabela que contém as medições numéricas do negócio onde cada uma das medições é obtida da intersecção de todas as dimensões [14]. Exemplo: total de vendas poderia ser um cálculo. Este cálculo é armazenado em uma tabela fato com a granularidade apropriada. Para se obter o total de vendas por loja por dia, a tabela fato deverá ter 3 atributos: data, loja e total de vendas.

- **tabela *lookup***: é a tabela que provê a informação detalhada de um atributo. Exemplo: a tabela *lookup* para o atributo Trimestre inclui a lista de todos os valores de trimestre disponíveis do *Data Warehouse*. Cada linha (cada trimestre) pode ter muitos atributos, um para a chave única que identifica o trimestre e um ou mais atributos adicionais que especificam particularmente

como aquele trimestre é representado num relatório. Por exemplo, primeiro trimestre de 2006 pode ser representado por "T1 2006".

Um modelo dimensional é composto de tabelas fato e tabelas *lookup,* conforme representado na Figura 3-2 . As tabelas fato podem se conectar a uma ou mais tabelas *lookup.* Dimensões e hierarquias são representadas por tabelas *lookup*. Atributos são todas as colunas das tabelas *lookup* que não são chave da tabela.

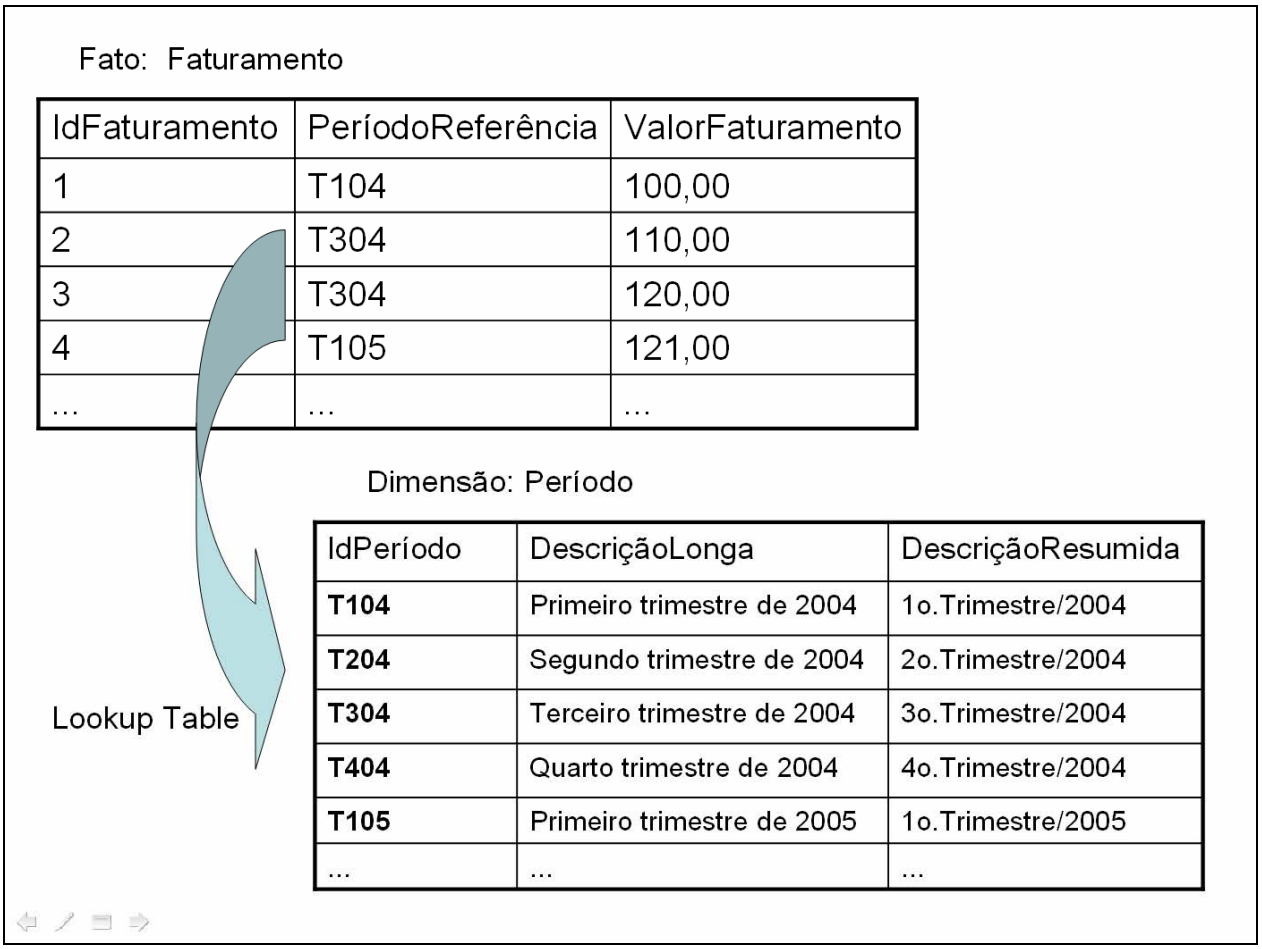


Figura 3-2: Representação da tabela *lookup* [14].

No desenho do modelo de dados de um *Data Warehouse* ou *Data Mart* os tipos de esquemas mais utilizados são: *star schema* e *snowflake schema*.

- ***star schema***: ilustrado na Figura 3-3, um único objeto (a tabela fato) fica ao centro da representação e é radialmente conectado aos outros objetos (tabelas de dimensão/*lookup*) ao redor

dele como uma estrela. Um *star schema* pode ser simples ou complexo. Uma estrela simples consiste de uma única tabela fato e uma estrela complexa pode ter mais de uma tabela fato.

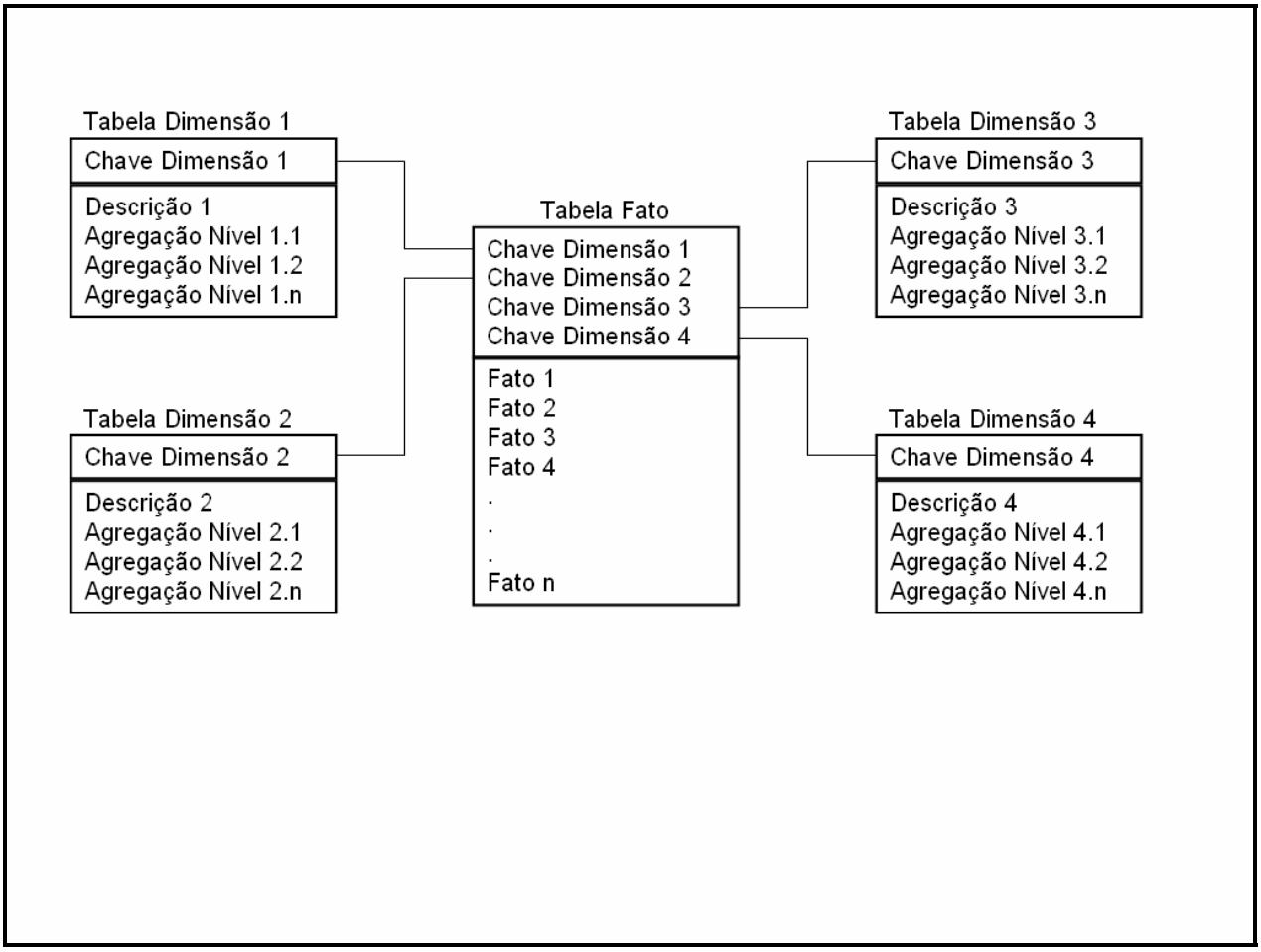


Figura 3-3: Representação de um *star schema* [17].

Neste esquema, apresentado na Figura 3-3 [17] temos a tabela fato no centro do modelo que representa quantidades de dados que podem ser agregados sem a perda de significado e necessita ser descrita por mais de uma dimensão. O propósito da tabela fato é facilitar o processamento que deve ser feito com base nas diversas ocorrências de dados encontradas na tabela fato. Os dados da tabela fato são construídos a partir de elementos de dados estruturados em níveis organizacionais que são de natureza somatória; ou seja, os valores destes elementos de dados podem ser agregados através de somas, médias ou das mais variadas formas sem prejudicar a integridade dos dados.

Ao redor da tabela fato temos todas as dimensões. As tabelas dimensão são de natureza descritiva, e não somatória; as tabelas dimensão descrevem ou categorizam os dados na respectiva tabela fato. A chave da tabela fato é uma chave composta de todas as chaves das tabelas dimensão. O resultado disto é: para cada linha na tabela fato haverá uma única combinação dos domínios das chaves das tabelas dimensão. Normalmente este tipo de modelagem garante uma melhor performance por ter um número reduzido de relacionamentos entre tabelas a ser realizado.

- ***snowflake schema***: a Figura 3-4 é uma extensão do *star schema*, onde cada ponto da estrela explode em muitos outros pontos. A grande vantagem deste esquema é a melhora no tempo de execução de uma consulta em função do mínimo uso de armazenamento em disco e um número menor de relacionamento entre tabelas *lookup*. A grande desvantagem é a manutenção em função do grande número de tabelas *lookup*.

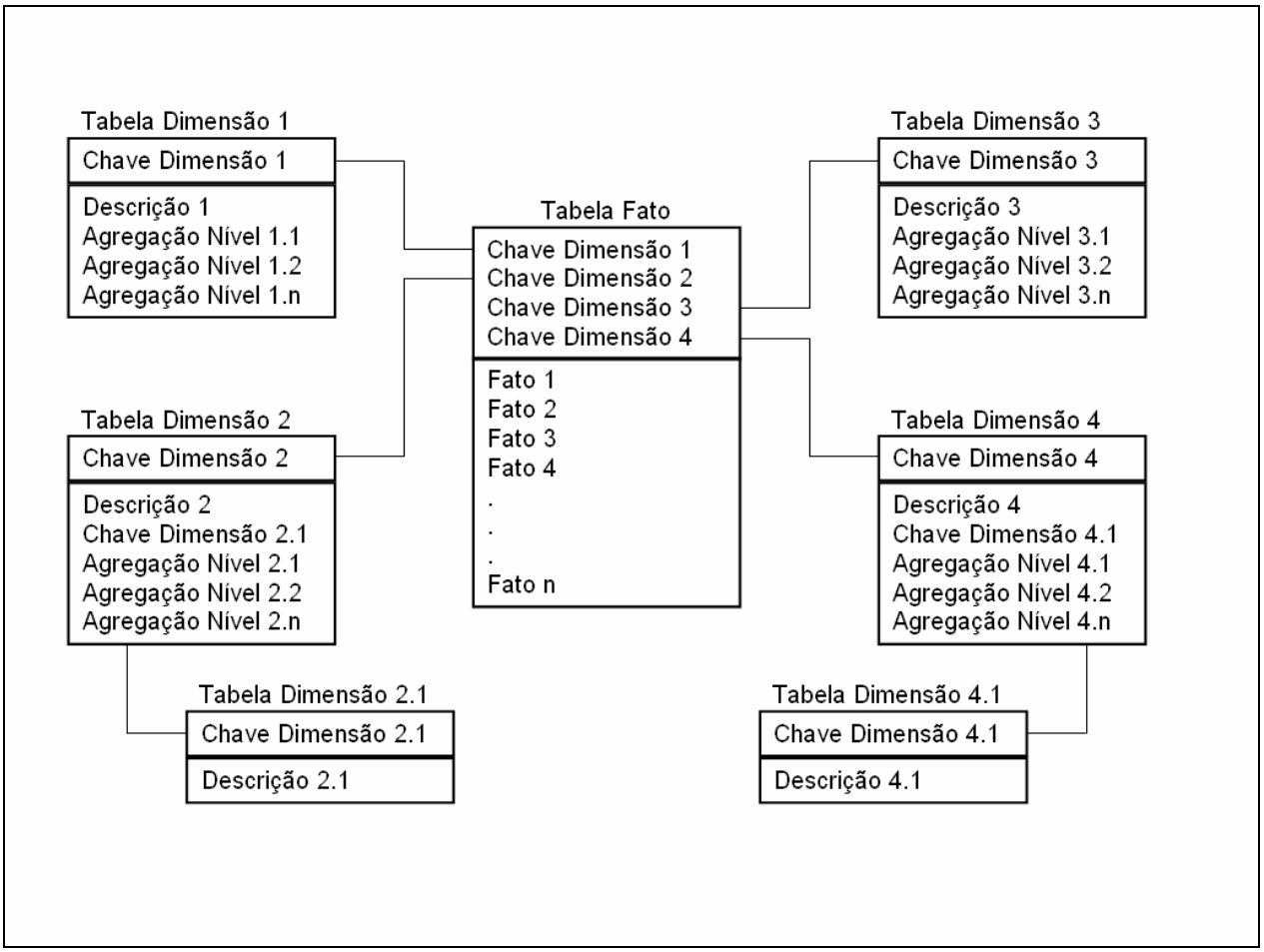


Figura 3-4: Representação de um *snowflake schema* [17].

On-Line Analytic Processing (OLAP)

O termo OLAP (*On-Line Analytic Processing*), foi criado para descrever uma abordagem dimensional para suporte à decisão [14]. Suas características não permitem uso de SQL (*Structured Query Language*) padrão para consulta de informações e sua capacidade não atende às necessidades de uma corporação no uso de um *Data Warehouse.* Por isso, algumas combinações de modelagem de dados são utilizadas no desenho de uma solução OLAP [10]: o Multidimensional OLAP (MOLAP) e o Relacional OLAP (ROLAP); ainda pode ter uma tecnologia mista que combina MOLAP e ROLAP dependendo das necessidades de negócio.

A modelagem MOLAP é a mais tradicional e mais simples de se implementar; os dados são armazenados em estruturas multidimensionais e em formatos de dados específicos de seus fornecedores e não em um banco de dados relacional. Este modelo permite facilidade no acesso e implementação dos dados por estarem disponíveis em uma única estrutura de dados. Apesar desta facilidade, o grande limitador é o volume dos dados; caso seja muito grande, implica em baixa performance no acesso a dados.

A modelagem ROLAP é implementada em um banco de dados relacional, que parece ter uma estrutura OLAP; porém, cada filtro aplicado não necessita de uma especificação para cubos OLAP, basta utilizar uma sentença SQL com filtros "WHERE" e aplicá-la à base de dados relacional. Esta estrutura permite a manipulação de grandes volumes de dados, dado que os mesmos estão armazenados em tabelas de dados distintas; em contrapartida uma série de combinações entre tabelas pode ser realizada para acessar determinado conjunto de informações resultantes de um filtro processando com baixa performance.

A modelagem HOLAP combina as facilidades das duas modelagens descritas acima: MOLAP e ROLAP, para acesso a informações agregadas; utiliza a estrutura de dados gerada pela modelagem MOLAP, pois estão calculadas e disponíveis para acesso direto. Para acesso a informações detalhadas, utiliza a estrutura de dados gerada pela modelagem ROLAP, pois o maior nível de detalhe das informações pode ser acessado diretamente da estrutura relacional.

3.2 *Data Warehouse* e sua Modelagem

Segundo Inmon [25] podemos definir *Data Warehouse* (DW) como uma coleção de dados orientada a assuntos, integrada, não volátil e variável no tempo, que é usada para apoio a decisões gerenciais. Um *Data Mart* pode ser definido como um *Data Warehouse* de pequena capacidade usado para atender a uma unidade específica de negócios [26]; sua diferenciação em relação ao *Data Warehouse* se dá pelo uso e gerenciamento e não pelo seu tamanho ou volume de dados. Os motivos que levam ao desenvolvimento de um *Data Mart* podem ser: criação de um projeto piloto, atender necessidades imediatas de uma unidade de negócio, atender a restrições de custo, tempo, entre outros.

Os objetivos de um *Data Warehouse* podem ser enumerados da seguinte forma, gerando assim os requisitos básicos para a sua construção [14]:

1. O *Data Warehouse* fornece acesso a dados corporativos ou organizacionais.
2. Os dados do *Data Warehouse* são consistentes.
3. Os dados no *Data Warehouse* podem ser separados e combinados usando-se qualquer medição possível do negócio (requisito clássico *slice and dice*)
4. O *Data Warehouse* não consiste apenas em dados, mas também em um conjunto de ferramentas para consultar, analisar e apresentar informações.
5. O *Data Warehouse* é o local em que se publica dados confiáveis.
6. A qualidade dos dados no *Data Warehouse* impulsiona a reengenharia de negócios.

Os objetivos de um *Data Warehouse* se aplicam a uma corporação enquanto os objetivos de um DBM se aplicam à área de Marketing, considerando os mesmos requisitos básicos apresentados anteriormente.

Além disso, um *Data Warehouse* garante aos usuários consulta às informações estratégicas e a grandes volumes de dados em um tempo muito curto [17].

A arquitetura típica de um *Data Warehouse* [17], apresentada na Figura 3-5, é composta dos seguintes componentes e perfis:

- **Ambiente Operacional**: é utilizado pelos usuários que devem realizar as transações e tarefas do dia a dia de uma corporação.

- **Camada de Transformações:** é utilizada pelos analistas de negócio como camada intermediária para produzir as informações necessárias para a tomada de decisão.

- **Armazenamento dos dados operacionais (ODS)**: é utilizado pelos executivos que tomam decisões táticas.

- **Dados detalhados do *Data Warehouse***: é utilizado pelos analistas de suporte à decisão que buscam padrões e tendências ainda não descobertos.

- ***Data Mart* levemente sumarizado**, como parte do *Data Warehouse*: é utilizado pelos analistas de suporte à decisão assim como pelos executivos.

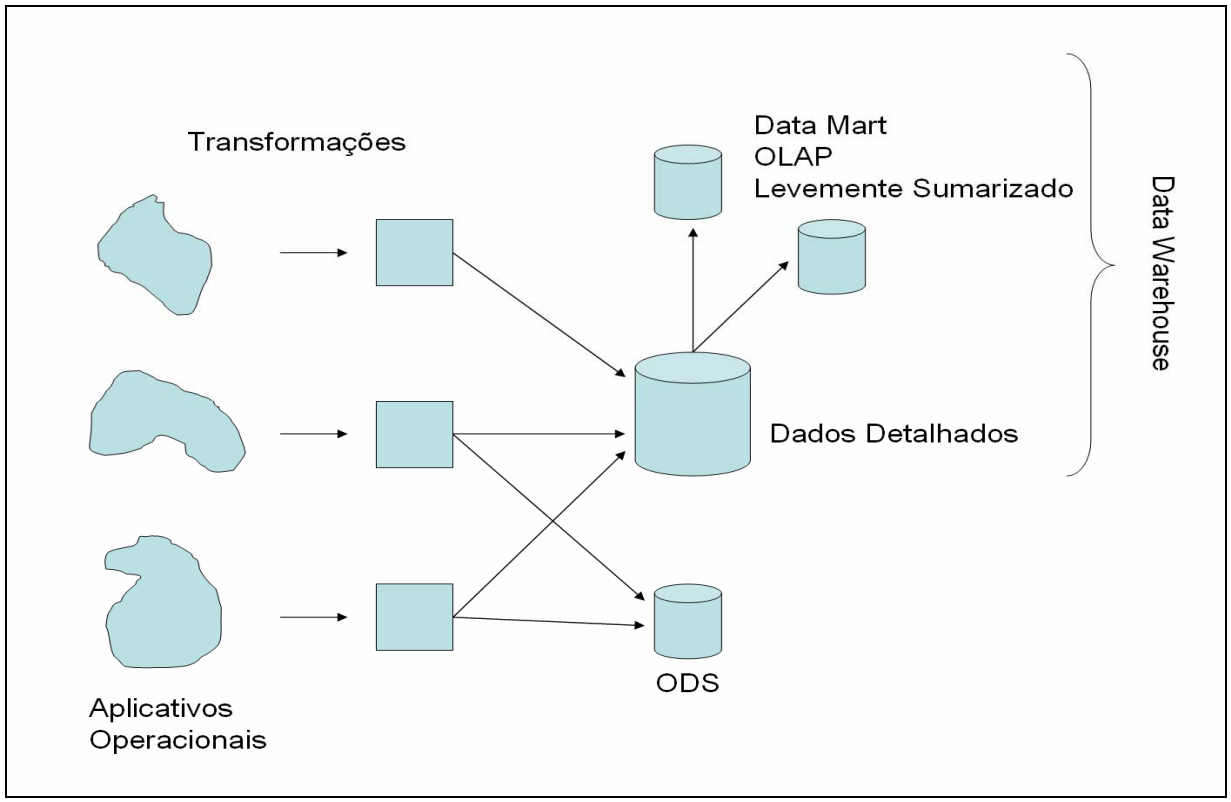


Figura 3-5: Arquitetura típica de um *Data Warehouse* [17].

Num ambiente de processamento analítico, com um *Data Warehouse* estruturado para apoiar as tomadas de decisões das áreas de negócio, alguns aspectos podem ser observados como descrito a seguir e apresentado na Figura 3-6 [17].

Os dados são transformados em informação e disponibilizados de forma que atendam às necessidades do usuário, seja por meio de estruturas corporativas com mais detalhes, mais informações históricas e mais normalizadas (um modelo relacional com o menor nível de granularidade) ou por meio de estruturas individuais que possuem um detalhe menor, possuem

informações de um período histórico menor e tipicamente não estão normalizadas (possuem estruturas otimizadas para uso nas análises realizadas pela área de negócio).

É importante salientar esta diferenciação porque cada nível hierárquico da corporação possui interesses distintos e pode ter disponíveis informações em formatos distintos e que atendam melhor às suas necessidades otimizando os recursos de *hardware* disponíveis.

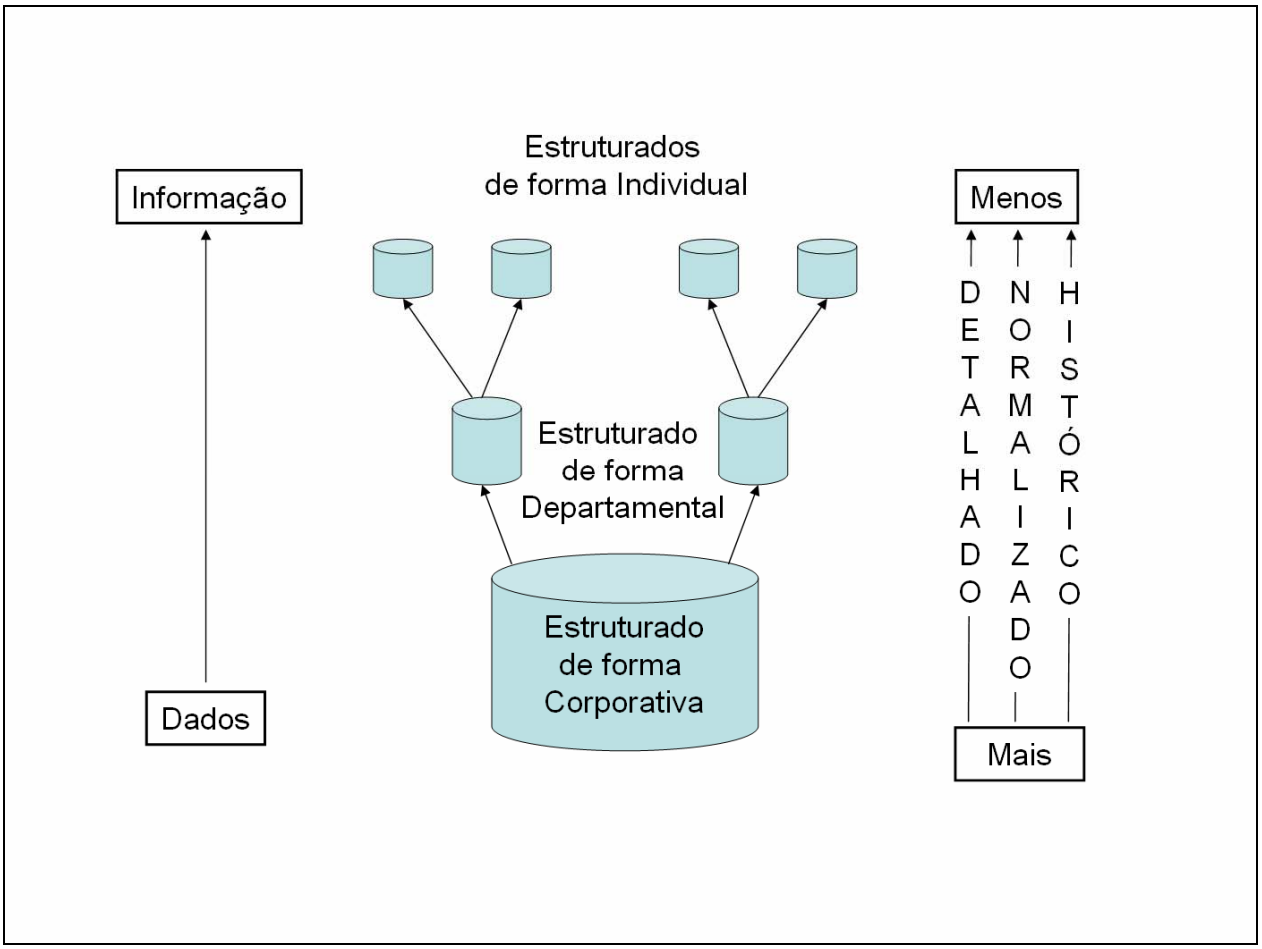


Figura 3-6: Uso da Informação em um *Data Warehouse* [17].

Um *Data Warehouse* pode ser composto de informações com diversas instâncias de dados que aparecem freqüentemente numa base de dados e com dados sumarizados, que contém um certo número de elementos de dados que podem ser utilizados como agrupamento ou categoria numa análise. Um ou mais destes agrupamentos ou categorias são hierárquicos, ou seja, possuem níveis distintos de agrupamento até o menor nível de granularidade [17].

Há cerca de 20 anos, os grandes autores e reconhecidos profissionais na área de *Data Warehouse*, Ralph Kimball e Bill Inmon, vêm desenvolvendo muitos estudos e avanços

permitindo que as organizações obtenham informações estratégicas e tomem suas decisões mais rapidamente [11][12].

Apesar dos objetivos desses autores serem os mesmos, suas visões são distintas, como veremos nos tópicos nomeados com seus nomes popularmente conhecidos, "Visão Kimball" e "Visão Inmon".

3.2.1 Visão Inmon

Bill Inmon é reconhecido como o "Pai" do *Data Warehouse* e o seu conceito de *Data Warehouse* é definido como um repositório que precisa servir às necessidades de todos os seus usuários e não somente de uma classe de usuários, "Os *Data Warehouse*s são organizados por áreas/assuntos da corporação...num modelo de dados corporativo. Usualmente, o *Data Warehouse* é construído e controlado por uma área centralizada... É realmente um esforço de toda a corporação." [11, p.2].

O paradigma de dados relacionais num *Data Warehouse* é aquele em que os dados devem estar no menor nível de granularidade e na terceira forma normal (3NF). Este tipo de modelagem de dados garante flexibilidade total às corporações caso necessitem de uma nova modelagem que atenda as necessidades do negócio.

Num ambiente corporativo existem muitas classes de usuários dos mais diferentes departamentos: contabilidade, finanças, *marketing*, produção, serviços, entre outros. Cada uma destas classes de usuários é uma comunidade distinta com a sua própria maneira de analisar os dados do *Data Warehouse*; isto exige que um *Data Warehouse* tenha um desenho de tabelas relacionadas. A vantagem de se desenhar um *Data Warehouse* baseado num formato relacional é a facilidade de remodelar e reorganizar os dados conforme a necessidade do negócio. Uma vez que o modelo relacional esteja desenhado de forma adequada e com dados existentes no menor nível de granularidade, qualquer outra configuração de dados pode ser realizada, como cubos multidimensionais, que são estruturas de dados com métricas agrupadas por categorias de dados, modelos *star schema*, *flat files*, etc.

Após ter os dados modelados, é possível "denormalizar levemente" os dados desde que sejam usados por todas as classes de usuários, ou seja, essa denormalização permite que algun dados sejam agrupados de forma a atender toda a corporação. Segundo Inmon, uma vez que a base relacional esteja estabelecida e implementada, existe a flexibilidade para apoio a *Data Marts*

multidimensionais e outras estruturas de dados como *Data Warehouse*s usados para exploração, bases de dados para mineração de dados, entre outros.

Sua metodologia aplica uma forma incremental para desenvolvimento de grandes *Data Warehouses;* assim, garante entregas parciais onde a cada iteração uma pequena parte do *Data Warehouse* é entregue até que esteja construído por completo. Desta forma, utilizando o método iterativo, ajustes podem ser aplicados a um pequeno conjunto de dados sem que seja necessária a reprogramação ou recarga de um grande volume de dados, o que poderia ocasionar sérios problemas num projeto de grandes dimensões.

Vale lembrar que os *Data Warehouse*s contém informação na maior granularidade de dados e que este nível de granularidade não é definido com base em nenhum requisito das áreas da corporação e sim com base em requisitos de dados da corporação como um todo.

O *Data Warehouse* requer uma tecnologia que dê conta de crescimento, dada a necessidade de armazenar um enorme volume de dados de toda a corporação; além disso deve prever o armazenamento dos dados históricos dos diversos sistemas legados.

Um *Data Warehouse* não é uma coleção de *Data Marts*, mas é um conjunto físico de diferentes componentes de dados, conforme ilustrado na Figura 3-7.

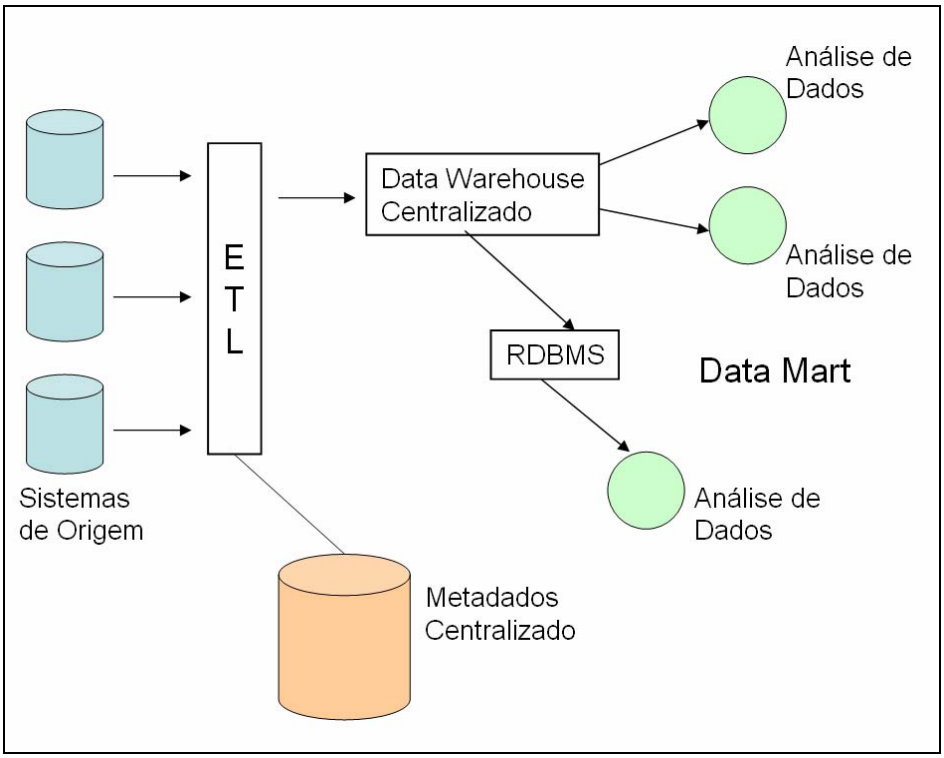


Figura 3-7: Arquitetura de um Data Warehouse Corporativo [23]

3.2.2 Visão Kimball

Seu conceito tem evoluído nos últimos 20 anos em função do ambiente da tecnologia da informação com mudanças freqüentes. Sendo assim alguns conceitos foram estabelecidos e difundidos nesta área.

A abordagem definida por Kimball como "Business Dimensional Lifecycle" é baseada nos princípios definidos a partir de sua experiência para a construção bem sucedida de um *Data Warehouse* [12].

A forma de estruturação dos dados denominada "*Data Mart*" ou um conjunto de "*Data Mart*s", é representada por um *star schema*, permitindo acesso direto às informações que cada usuário necessita. Esta abordagem é denominada "*Bottom Up*".

A arquitetura proposta por Kimball é denominada "*Data Warehouse bus architecture*" e é composta por uma *staging area*, onde os arquivos que serão usados como fonte de dados para o *Data Warehouse* são armazenados, sejam eles arquivos txt, ou qualquer outro formato de dados, e pelo próprio *Data Warehouse Bus*, que é dimensional, contém dados transacionais e sumarizados, inclui *Data Mart*s com assuntos únicos ou uma única tabela fato, e pode ter múltiplos *Data Mart*s em uma única base de dados.

O modelo dimensional contém as mesmas informações de um modelo normalizado (3NF) porém com informações agrupadas para proporcionar facilidade de uso e performance na execução de uma consulta.

É importante ressaltar que os modelos dimensionais são desenhados para atender a um processo de negócio e não a um departamento de uma organização; portanto, as informações que compõem um *Data Warehouse* corporativo estão dispostas em forma de tabelas fato que representam as métricas do processo de negócio e as dimensões que representam os atributos descritivos.

O conceito de repositório de dados definido no *Data Warehouse Bus* é virtual e não físico, ou seja, existe uma série de *Data Mart*s, cada um com seu respectivo *star schema* composto de sua tabela fato e suas dimensões associadas, e que formam o *Data Warehouse* corporativo com as informações extraídas dos sistemas de origem através de processos de ETL (*Extraction, Transformation and Load*), conforme representado na Figura 3-8. O termo RDBMS refere-se à Bancos de Dados Relacionais e o termo MDB refere-se a Bancos de Dados Multidimensionais.

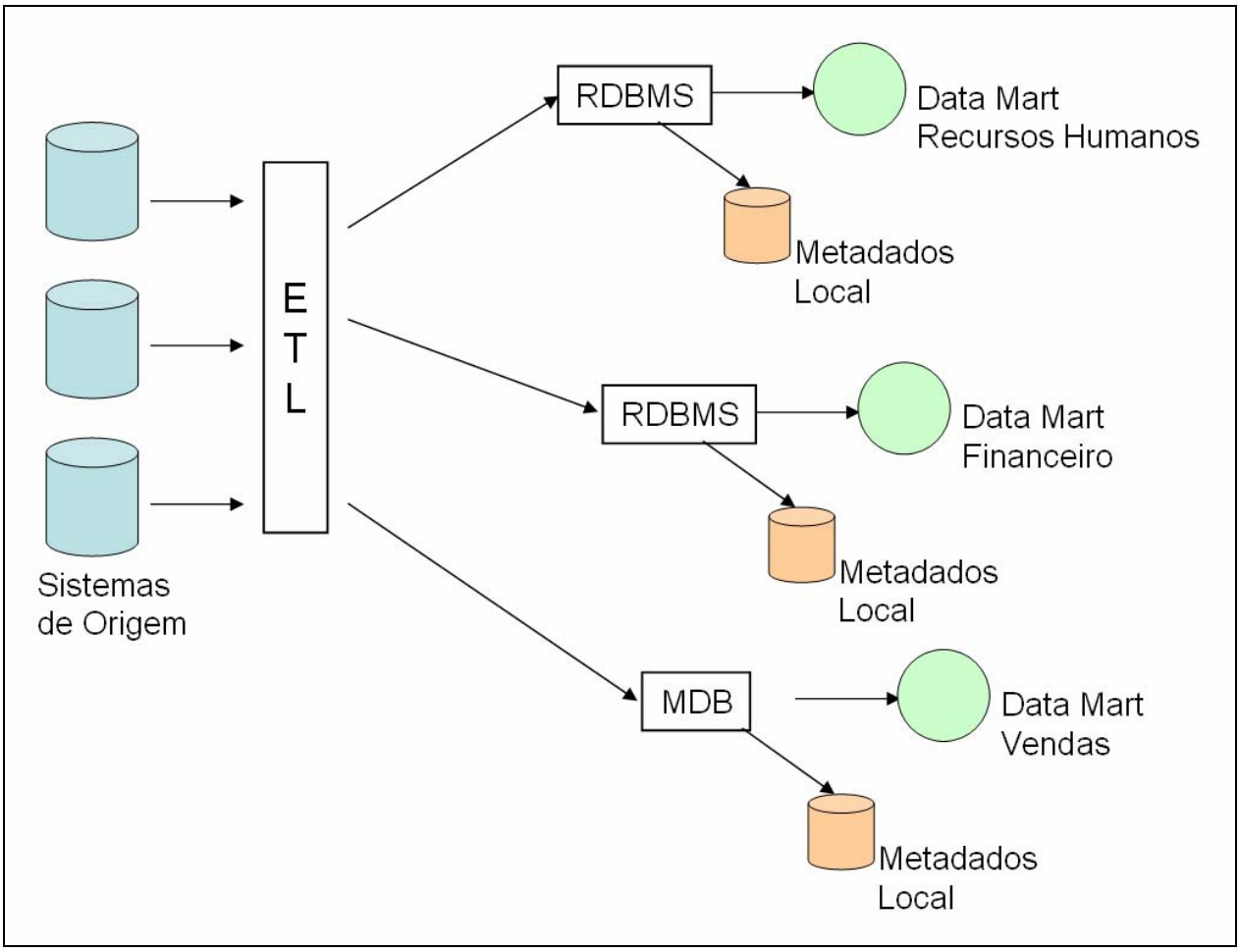


Figura 3-8: Arquitetura de um Data Warehouse baseada em Data Marts [23]

3.2.3 Diferença entre as abordagens Inmon e Kimball

Na área de *Data Warehousing*, existem muitas discussões sobre qual abordagem utilizar, a de Bill Inmon ou a de Ralph Kimball. De uma forma bastante simples segue uma explicação sobre a diferença entre estas abordagens [7].

- **Paradigma de Bill Inmon:** O *Data Warehouse* é parte de um sistema completo de inteligência do negócio. Uma corporação tem um *Data Warehouse* e *Data Marts* cuja fonte de dados é o *Data Warehouse*. No *Data Warehouse* as informações são armazenadas na 3ª. Forma Normal.

- **Paradigma de Ralph Kimball:** O *Data Warehouse* é um conglomerado de todos os *Data Mart*s da corporação. As informações são sempre armazenadas num modelo dimensional.

Entre estas abordagens não há certo ou errado e sim conceitos distintos de representação do *Data Warehouse*.

3.3 Database Marketing

O propósito do *database marketing* é combinar dados de clientes provenientes de diversas fontes de dados baseados em uma única chave, o cliente. Segundo [15], existe um tipo de transformação básica utilizado para se obter um *Data Mart* baseado na visão única de cliente e que possa ser utilizado por toda a organização além de ter suas informações relacionadas a outros *Data Mart*s através da chave cliente.

Um Database Marketing tem como objetivo principal fornecer informações para uso na mineração de dados a fim de entender comportamentos, descobrir tendências e realizar previsões. Assim como um *Data Warehouse* possui uma estrutura de dados sumarizada para publicação de informações através de relatórios, um Database Marketing contém informações sumarizadas de forma a atender as análises de mineração de dados incluindo uma série de variáveis derivadas.

3.4 Mineração de Dados *(Data Mining)*

> "Mineração de Dados" é o processo de descobrimento de novas correlações, padrões e tendências significativas filtradas a partir de grandes volumes de dados armazenados em repositórios e pelo uso de tecnologias de reconhecimento de padrões assim como técnicas estatísticas e matemáticas."
>
> - The Gartner Group

A mineração de dados tem como objetivo descobrir padrões e regras significativas nos dados [15] de forma automatizada a partir de uma quantidade muito grande de dados permitindo que as organizações melhorem seu *marketing*, suas vendas assim como operações de suporte a clientes através de um melhor conhecimento dos mesmos [16].

Para a realização da mineração de dados algumas tarefas básicas são desenvolvidas, são elas: Classificação, Estimativa, Predição, Agrupamento por Afinidade, Clusterização e Descrição

[16]. A partir de um *Database Marketing,* quaisquer destas tarefas podem ser realizadas com sucesso, pois sua estrutura de dados permite facilmente a realização de qualquer uma das análises sem que seja necessária adaptação aos dados já disponíveis, pois eles já estão adequados para este fim.

Um ciclo de mineração de dados é realizado em basicamente 4 passos representados graficamente na Figura 3-9:

1. Identificar a oportunidade de negócio.

2. Utilizar técnicas de mineração de dados para transformar os dados em informações utilizáveis.

3. Agir em função das informações descobertas.

4. Medir os resultados para saber o quão efetivo foi o contato realizado com os clientes.

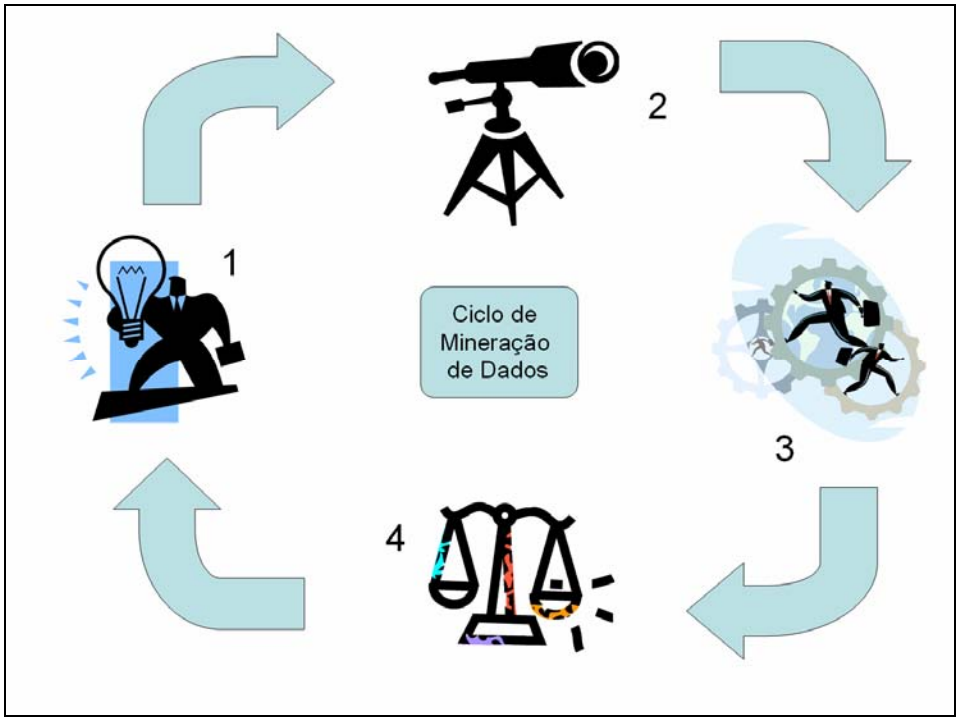


Figura 3-9: Ciclo de mineração de dados [16].

Uma definição muito importante na mineração de dados está relacionada à criação das amostras a serem utilizadas, caso a amostra selecionada não permita uma análise que represente a base completa de dados, os estudos possíveis podem resultar em informações completamente distintas da realidade não garantindo a obtenção dos resultados esperados.

3.5 Considerações Finais sobre os conceitos apresentados

Foram apresentados e descritos os diversos conceitos utilizados na construção e utilização de um DBM. Um DBM é concebido e implementado a partir de uma necessidade da área de *marketing* em coletar e analisar informações de seus clientes para tomada de decisão ou ainda para a implementação de uma estratégia de relacionamento com seus clientes.

Para que um DBM seja construído é necessário o armazenamento das informações em um repositório de dados, este repositório pode ser representado conceitualmente por um *Star Schema,* onde serão representadas todas as entidades de dados e seus atributos. Para a implementação deste modelo conceitual de dados representado pelo *Star Schema*, é necessária a definição de uma estratégia de carga dos dados e definição da forma como estes dados serão disponibilizados. Para um DBM será considerada a visão abordada por Kimbal, onde os *Data Marts* compõem um *Data Warehouse Corporativo* e neste trabalho abordaremos a construção de um *Data Mart* com foco em *marketing*, ou seja, um DBM. Para exploração e análise dos dados serão consideradas as aplicações OLAP, que permitirão a visualização das informações em relatórios de diferentes formatos e também *Data Mining,* que permitirá criar segmentos de clientes por meio de seus comportamentos e características e futuramente desenvolver modelos de previsão para definir um público-alvo para participar de um evento.

Desta forma, estes conceitos serão utilizados como base do desenvolvimento do DBM em conjunto com as práticas participativas que resultarão num modelo conceitual de dados que represente a necessidade da DBMTech.

4 Elicitação de Requisitos para um DBM: Abordagem Proposta

4.1 O Processo de Engenharia de Requisitos

A engenharia de requisitos é um processo que envolve todas as atividades necessárias para a criação e manutenção do documento de requisitos de sistemas. Conceituando o termo "Requisito", é toda condição necessária para a obtenção de um certo objetivo ou para o preenchimento de certo fim [19]. Existem 4 atividades genéricas de alto-nível no processo de engenharia de requisitos: estudo de viabilidade do sistema, elicitação e análise dos requisitos, especificação dos requisitos e sua documentação, e a validação destes requisitos.

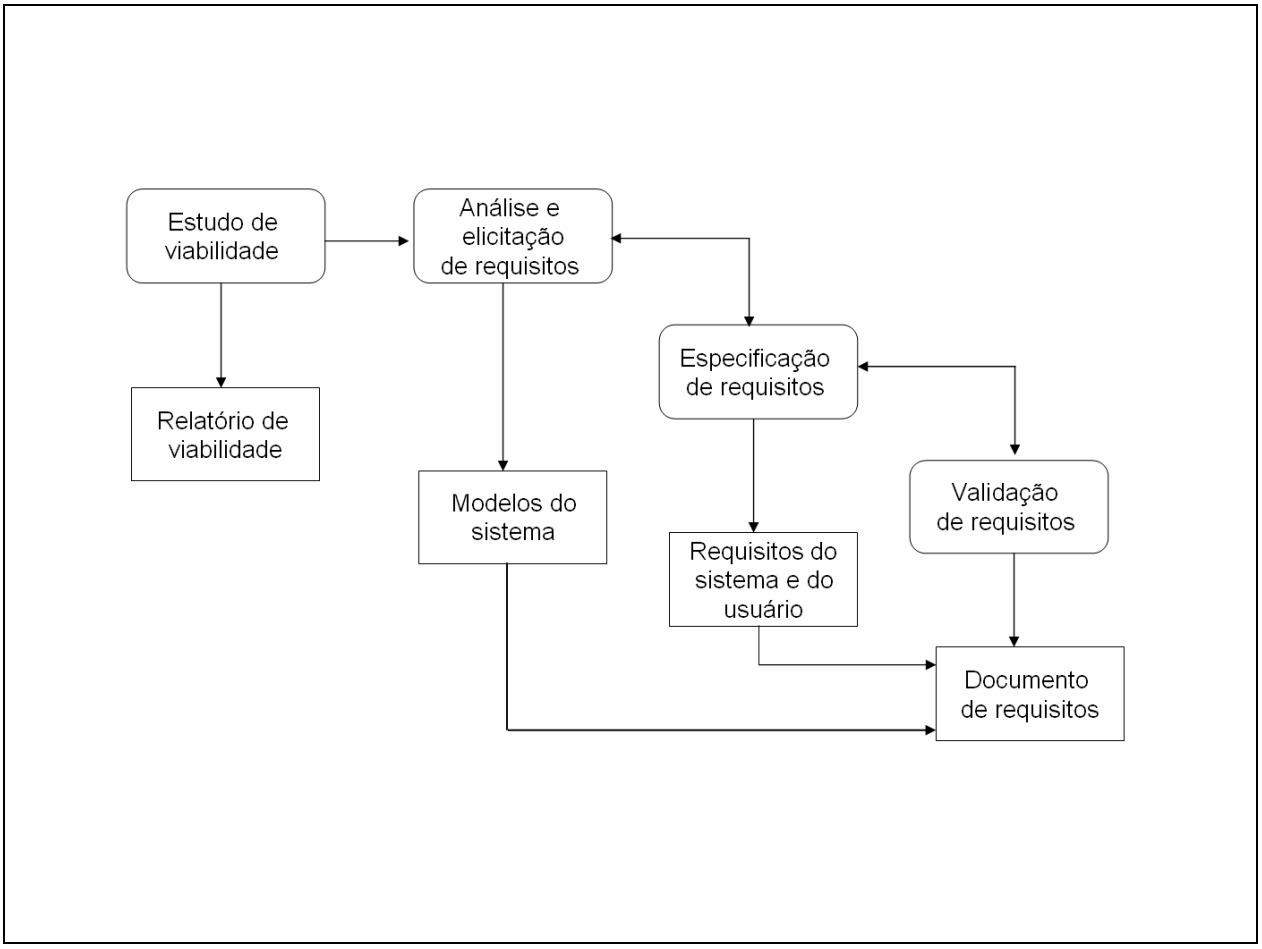


Figura 4-1: Processo de engenharia de requisitos

Segue a descrição das etapas ilustradas no modelo descrito na Figura 4-1:

- **Estudo de Viabilidade:** esta fase compreende a análise de uma descrição geral do sistema e como isto será utilizado pela organização, resultando em um relatório de recomendação do que deve ou não ser passado para o processo de engenharia de requisitos e desenvolvimento de sistemas [18]. Nesta etapa, é importante compreender o quanto este sistema está alinhado aos objetivos de negócio da organização para entender o quanto ele será valioso a organização. No detalhamento das informações levantadas e relatadas, algumas questões devem ser respondidas:

1. Este sistema contribui para os objetivos gerais da organização?
2. O sistema pode ser implementado utilizando tecnologia existente e dentro das restrições de tempo e custo propostas?
3. O sistema pode ser integrado a outros sistemas existentes?

Para esta fase os *experts* em tecnologia, usuários finais, engenheiros de sistemas e outros profissionais familiares ao sistema que será implementado são entrevistados e consultados para garantir que o documento de requisitos de alto-nível contenha o entendimento geral coletado entre os interessados no sistema.

- **Análise e Elicitação de Requisitos:** esta fase compreende atividades da equipe de desenvolvimento de sistemas com os clientes e usuários finais do sistema para determinar que serviços deverão ser entregues, qual o nível de performance esperado, quais as restrições de *hardware*, entre outros [18]. Nesta fase todos os *stakeholders* são envolvidos em entrevistas e dinâmicas para que os requisitos sejam entendidos e esclarecidos garantindo um modelo de sistema que represente a necessidade descrita pelos *stakeholders*. Não há uma abordagem perfeita e universalmente aplicável para a análise de requisitos; normalmente se utiliza de diversas abordagens para desenvolver um entendimento completo e análise dos requisitos.

- **Especificação dos Requisitos:** esta fase compreende a estruturação dos requisitos em diferentes formatos, seja em linguagem natural, linguagem formal ou semiformal, representações simbólicas ou gráficas [19], através de documentos completos e consistentes que contenham um entendimento do produto a ser desenvolvido e seus requisitos.

- **Validação dos Requisitos:** esta fase compreende a validação dos requisitos que define o sistema que os clientes desejam, é muito parecido com a fase de análise mas tem o objetivo de encontrar problemas nos requisitos [19]. Esta é uma fase de extrema importância porque os requisitos validados serão utilizados na implementação por meio do modelo espiral, caso a

validação não seja feita plenamente, durante a implementação os erros podem incorrer em custos não provisionados em função de retrabalhos que deverão ser feitos de forma a atender os requisitos do usuário final. Corrigir um erro de implementação custa 100 vezes mais do que corrigir um erro na fase de requisitos [20]. Diferentes tipos de checagem devem ser realizados [18]:

- **Validade:** garante que o sistema proverá todas as funções que atendem as necessidades do cliente.

- **Consistência**: garante que não há conflitos de requisitos, com definições contraditórias ou com descrições diferentes para a mesma função.

- **Completude:** garante que todas as funções requeridas pelo cliente estão incluídas na especificação de requisitos.

- **Realismo:** garante que os requisitos podem ser implementados com o orçamento e tecnologia disponíveis.

- **Verificabilidade**: garante que os requisitos estão representados de uma forma que tanto o usuário final quanto o desenvolvedor possam verificar os requisitos especificados.

4.2 Práticas Participativas na Engenharia de Requisitos

Segundo [27] entre outras referências, Design Participativo (DP) é um conjunto de teorias, práticas e estudos relacionados aos usuários finais como participantes em atividades que levam ao desenvolvimento de produtos de *hardware* e *software* e atividades baseadas em computação. DP teve início num explícito contexto político, como parte de um movimento democrata sobre locais de trabalho na Scandinávia, promovido pelos experimentos realizado pelos pesquisadores em conjunto com os empregadores para entendimento dos problemas da época.

O *Design* Participativo promove a decisão de forma democrática e consensual, permitindo que todas as opiniões sejam expostas e, em consenso, seja alcançado o que deve realmente ser considerado para o resultado final. Em geral, práticas participativas utilizam oficinas de trabalho, que devem ser promovidos de forma organizada para possibilitar o alcance dos objetivos propostos. Para que isto ocorra as oficinas de trabalho, segundo [28], podem ser organizadas da seguinte forma:

- dar aos usuários a voz no desenho do processo, desta forma se aumenta a probabilidade de se fazer um desenho utilizável;

- permitir que os participantes técnicos e não-técnicos participem de forma igual;

- prover uma oportunidade em que os desenvolvedores possam entender e trabalhar com os usuários finais;

- promover um encontro para identificação dos problemas;

- prover uma oportunidade de obter ou melhorar a capacidade de decisão do usuário;

- ser altamente produtivas, e

- utilizar técnicas que podem ser facilmente aprendidas e aplicadas em atividades futuras.

A escolha dos 3 métodos participativos utilizados foi realizada a partir dos 61 métodos apresenteados em [27]. Os critérios para seleção foram baseados no objetivo proposto por cada método, por exemplo, uso do *brainstorming* nas oficinas de trabalho, leitura do mapa do ambiente de trabalho para desenho e identificação das tarefas e papéis dos envolvidos, criação de um boneco através da participação direta do usuário; e nos produtos resultantes em cada método proposto, por exemplo, lista das idéias priorizadas, desenho do fluxo dos trabalhos diários, artefatos de papel representando a aparência do sistema.

A seguir apresentamos os 3 métodos participativos que utilizamos em nossa proposta para elicitação de requistos de um DBM.

4.2.1 Group Elicitation Method

Este método provê suporte à decisão sobre o que deve ser considerado durante a elicitação de requisitos, por meio do compartilhamento de idéias em seções de *brainstorming* [29].

A dinâmica dos trabalhos de elicitação por meio deste método se dá pela realização de oficinas de trabalho, onde inicialmente se estabelece o problema de forma clara e objetiva com a participação dos *experts* no assunto a ser discutido. Com o problema definido, os diferentes *stakeholders* participantes da oficina de trabalho podem ter diferentes pontos de vista sobre o mesmo problema. Para que estas idéias sejam relatadas, cada participante escreve suas idéias em uma folha de papel e, então, passam o papel ao próximo participante. Ao receber as idéias do outro participante, cada participante pode concordar, discordar, ou estabelecer um novo ponto de vista. Esta fase se encerra quando todos os participantes tiverem visto as idéias de todos os outros participantes. Após a coleta de todas as idéias sugeridas, uma reformulação do que foi escrito é

realizada a fim de reduzir o número de conceitos principais a um grupo menor de conceitos. Uma vez definidos os conceitos, os relacionamentos entre estes conceitos devem ser estabelecidos com a contribuição de cada participante, onde cada um completa uma tabela cruzada com o resultado da lista final de conceitos, estabelecendo a sua percepção da importância relativa entre todos os pares de conceitos (maior, igual, menor). Assim é possível calcular um consenso por meio da avaliação de importância atribuída a cada par de conceito definido. Ao final, uma análise crítica dos resultados é realizada a fim de validar o que foi definido e caso haja necessidade, modificações são realizadas para adequar as idéias sugeridas pelos participantes.

Ao final das atividades, são produzidos os seguintes artefatos:

1- Lista de Idéias
2- Lista de Conceitos (refinamentos e combinações de idéias).
3- Classificação por Importância dos Conceitos
4- Análise crítica da classificação de importância

4.2.2 BluePrint Mapping

Este método baseia-se no uso de um grande mapa do local de trabalho (*the blueprint*), onde os participantes adicionam fotos ou desenhos dos lugares de trabalho, das ferramentas de trabalho e dos trabalhadores [29] a este grande mapa de forma a ilustrar cada um destes componentes no fluxo das tarefas de trabalho diário. O método de trabalho utilizado durante a aplicação desta técnica envolve cada participante do local de trabalho, para que cada um deles coloque no grande mapa do local de trabalho pequenas representações, onde representações podem ser lugares de trabalho, ferramentas de trabalho ou trabalhadores. A cada representação colocada no mapa, discussões são realizadas entre os participantes para que seja explicado e compreendido por todos, o trabalho desenvolvido naquele local. Dando continuidade ao trabalho, os mesmos participantes ou participantes diferentes revisam a situação atual do mapa até aquele momento, discutindo sobre o modo como o trabalho é realizado em cada ponto representado no mapa. Para que este trabalho seja produtivo e represente a realidade da organização em termos de atividades e seus fluxos, é de extrema importância a participação das pessoas que tenham envolvimento com as tarefas diárias da organização envolvida no processo analisado, afinal estas pessoas devem descrever cada uma das tarefas a serem representadas no mapa.

Ao final das atividades são produzidos os seguintes artefatos:

1- Mapa com as ilustrações nos lugares corretos.
2- Entendimentos de como os grupos percebem como os trabalhos devem ser organizados.
3- Discussões sobre os trabalhos ordinários, fluxos de trabalho diário e exceções.

4.2.3 PICTIVE (Plastic Interface for Collaborative Technology Iniciatives through Vídeo Exploration)

Este método baseia-se na utilização de recursos de baixa tecnologia, onde os participantes prototipam a aparência e – num nível descritivo – a dinâmica de um sistema com uma interface gráfica ou textual. Esta técnica é muito útil para a fase de design, mas pode ser usada no levantamento de um protótipo feito em papel, ou para a análise [2]. Por meio da utilização de suprimentos comuns de escritório como canetas coloridas, *post-it notes* e papéis coloridos, além de materiais pré-impressos, customizados, como por exemplo, componentes de uma interface que fazem parte de um guia de estilos ou de um ambiente de desenvolvimento, ícones do domínio de trabalho são apresentados e discutidos.

Os trabalhos são iniciados pela apresentação de cada participante que abrange o expertise de cada um, o tipo de contribuição que pode oferecer para o trabalho compartilhado, para a organização e para o grupo, assim como o que é esperado de cada participante pelo grupo de trabalho. Caso seja necessário, mini-tutoriais podem ser utilizados para informação dos participantes sobre os materiais utilizados como base nos trabalhos a serem realizados. Continuando os trabalhos em conjunto com os participantes usuários é realizada a exploração do domínio das tarefas, na forma de levantamento, análise ou desenho. Ao final, se realiza uma breve revisão dos resultados atingidos com o grupo durante a sessão, preferencialmente gravada em vídeo. Nas sessões de trabalhos propostas por este método é de extrema importância a participação do maior número de *stakeholders* possível.

Ao final das atividades são produzidos os seguintes artefatos:

1- Artefatos de papel representando a aparência do sistema.
2- Artefatos de papel que possam ser utilizados na reconstrução das idéias do grupo sobre a dinâmica do sistema.
3- Revisão detalhada gravada em vídeo mostrando a aparência, a dinâmica e um resumo.

4.3 Uma Proposta Participativa para Elicitação de Requisitos de um DBM

A proposta para o uso de técnicas participativas na elicitação de requisitos prevê a utilização das técnicas *Group Elicitation Method*, *BluePrint Mapping* e *PICTIVE* para que os resultados obtidos nesta fase do projeto tenham minimizadas as falhas propagadas para as próximas fases. O diagrama da Figura 4-2 ilustra como as técnicas participativas foram utilizadas nas diferentes fases da engenharia de requisitos. O *Group Elicitation Method* foi utilizado para estudo de viabilidade e análise e elicitação de requisitos; o *BluePrint Mapping* para análise e elicitação de requisitos, o *BluePrint Mapping* e o *PICTIVE* para a especificação e validação dos requisitos.

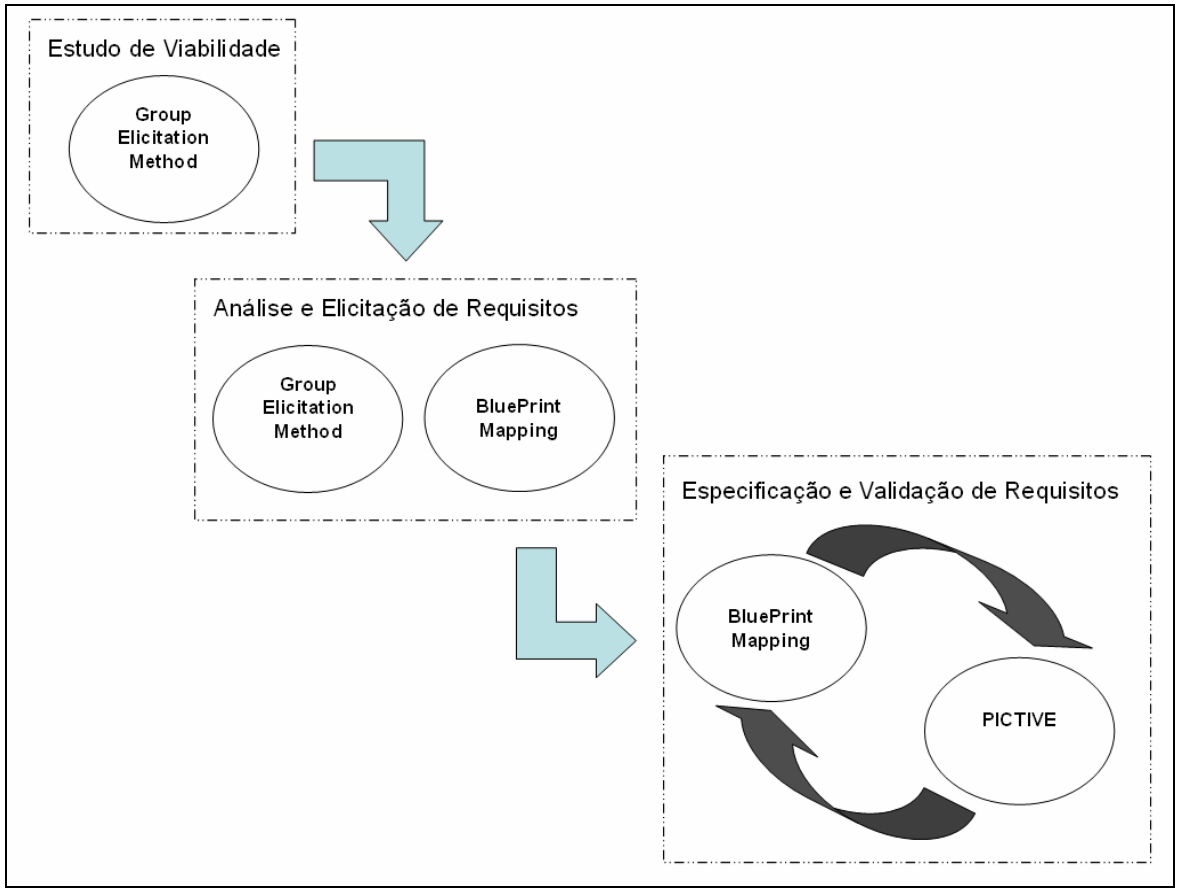


Figura 4-2: Proposta Participativa na Engenharia de Requisitos

5 Estudo de Caso: DBMTech

A DBMTech é o nome fictício de uma empresa de *software* e soluções de negócio para a área estratégica das empresas dos mais diversos ramos de atividades, seja financeiro, varejo, manufatura, entre outros.

A DBMTech tem a necessidade de criar uma base de dados que lhe permita aplicar técnicas de mineração de dados para conhecer melhor seus clientes e contatos assim como produzir relatórios gerenciais à partir das informações existentes. As informações utilizadas são aquelas provenientes de dados utilizados pela área de *marketing* e os relatórios gerenciais criados apresentam métricas e resultados baseados nestas mesmas informações que ficariam disponíveis em um DBM.

Este estudo de caso investiga a utilização de técnicas participativas na elicitação de requisitos e na modelagem de dados que atenda a estas necessidades e que possa ser implementado por uma equipe técnica de sistemas.

5.1 Cenário atual da DBMTech

Em 2003 foi implementada a estratégia de *Balanced Scorecard*, que é uma abordagem para administração estratégica baseada em indicadores desenvolvida por Robert Kaplan e David Norton em meados de 1990 [13].

Uma missão foi estabalecida para a DBMTech com uma série de objetivos estratégicos para serem atingidos, como metas financeiras, índice de satisfação dos clientes, entre outras. Para atingir os objetivos estratégicos, alguns indicadores foram criados para que fosse monitorada e mensurada a evolução da empresa e, por meio das medições, obtidos os resultados reais da empresa. Também desejava-se conhecer quanto do objetivo estabelecido estava sendo cumprido.

Um desses indicadores é o Valor do Cliente, que é derivado de informações sobre: faturamento anual do cliente, há quanto tempo ele é cliente, custos com iniciativas de *marketing* para cada cliente, entre outras informações. Para que esse indicador seja calculado, as informações precisam ser coletadas, estruturadas, exploradas e agregadas resultando no número que indica o Valor do Cliente.

No processo de cálculo do indicador de Valor do Cliente atual existem informações que não são coletadas, ou ainda, não são armazenadas de forma estruturada para a realização de estudos que permitam identificar o perfil de cada cliente e o valor de cada cliente.

Este indicador é extremamente importante para medir o nível de satisfação dos clientes em relação aos serviços prestados pela empresa, o que impacta diretamente nos resultados da empresa e, conseqüentemente, entender em que medida a empresa tem atingido os resultados em relação ao objetivo proposto pela estratégia desenhada. Para que o indicador seja atualizado com regularidade, por exemplo, mensalmente, foi criada a iniciativa de estruturar as informações permitindo que o indicador e os relatórios gerenciais sejam criados, a cada mês, através de processos mais ágeis e de forma estruturada e organizada.

Para atender a esta requisição, ficou definido que um DBM seria a melhor maneira de atender às necessidades levantadas e apresentadas pela DBMTech. Neste estudo de caso serão utilizadas as Técnicas Participativas para elicitar os requisitos necessários para a criação de um DBM que viabilize a mineração de dados e a geração de relatórios gerenciais, assim como a atualização do indicador de Valor do Cliente.

5.1.1 Equipe de Trabalho

Neste trabalho foram envolvidas as áreas de Marketing e Tecnologia, conforme organograma representado pela Figura 5-1, para que fosse especificado o modelo de dados a ser implementado.

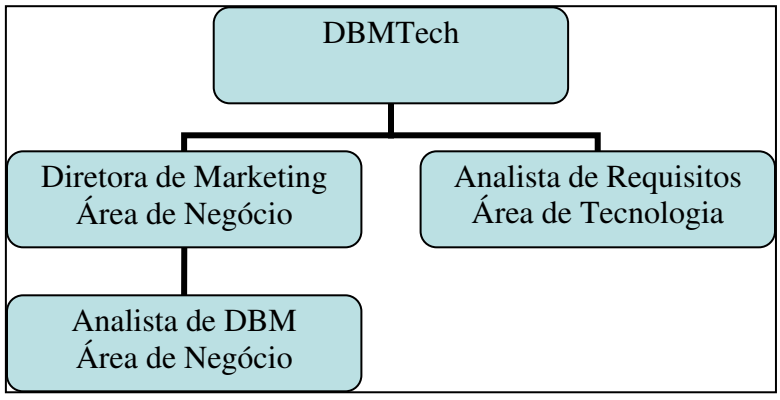


Figura 5-1: Organograma da DBMTech

A participação da área de Marketing deu-se através da Analista de DBM e da Diretora de Marketing, trazendo as suas necessidades para este estudo de caso. A área de Tecnologia foi representada pelo Analista de Requisitos que realizou as atividades de elicitação de requisitos baseada em práticas participativas.

A participação destas pessoas foi decidida em função do seu envolvimento com a utilização das informações do DBM. A Analista de DBM que será a principal usuária realizando todas as consultas de informações para geração dos relatórios, foi a pessoa chave na condução dos trabalhos pois todo o conhecimento sobre as necessidades do negócio ficam centralizadas nesta pessoa, assim sendo a participação da Diretora de Marketing aconteceu somente na validação dos resultados deste trabalho dado que todo o detalhe do processo existente era de domínio da Analista de DBM.

A DBMTech não possui uma área de tecnologia com analistas de sistemas, programadores, administradores de dados que possam implementar o modelo de dados resultado deste trabalho. Portanto, houve somente a participação do Analista de Requisitos para definir a especificação do modelo de dados que deverá ser implementado por uma equipe terceirizada pela DBMTech.

5.1.2 Necessidades da DBMTech

As necessidades da DBMTech foram relatadas num documento gerado pela Analista de DBM com todas as atividades realizadas atualmente, incluindo as informações acessadas e geradas para cada um dos relatórios e planilhas eletrônicas, além de relatórios resultantes dos processos executados através das atividades descritas. Uma transcrição do documento entregue pela Analista de DBM encontra-se no Anexo I.

Este documento (Anexo I) foi base para uma análise de viabilidade que permitiu entender, em alto nível de abstração as necessidades da DBMTech. Algumas das informações já estão disponíveis de forma estruturada, em tabelas de dados ou em planilhas e outras são capturadas através de pesquisas de opinião mas não têm um repositório próprio num banco de dados preparado para que sejam armazenadas as respostas destas pesquisas.

A

Tabela 5-1, a seguir, mostra a lista das entidades que foram identificadas previamente pela Analista de DBM e os pontos críticos em relação a cada uma delas. A expectativa é que pontos críticos devam ser melhorados com a implementação do projeto que se inicia com este estudo de caso.

Tabela 5-1: Lista de entidades levantadas inicialmente

Entidade	Descrição apresentada
Contatos	Cada Gerente de contas possui sua lista de contatos, isto dificulta a centralização das informações e por sua vez a utilização dos contatos de forma correta.
Clientes	Não existe um local único para se entender a base de clientes e retirar informações reais, como quantidade de clientes atual, principais clientes de financial, etc.
Atividades	Não existe um controle das ações, esta falta de controle dificulta o planejamento dos próximos anos e as análises de retorno.
Custos	Todos os custos das ações de *marketing* não são controlados dificultando a projeção do orçamento e a identificação dos principais gastos.

A DBMTech tem como objetivo centralizar as informações do *marketing* de forma organizada para auxiliar em todas as ações de relacionamento com o cliente realizadas pela área de *marketing*. As informações de contatos, clientes, atividades e custos foram identificadas inicialmente pela Analista de DBM, sob supervisão da Diretora de Marketing e devem ser coletadas e armazenadas de forma centralizada; isto é, em uma base de dados que suporte todas as necessidades demandadas pela área de *marketing*.

Não existe um processo automatizado para a coleta de informações e geração dos relatórios, o que dificulta o trabalho da Analista de DBM, que executa uma série de tarefas operacionais, como deduplicação de dados dos sistemas origens, análise dos dados de origem para garantir a consistência dos dados a serem apresentados nos relatórios. Este trabalho poderia ser realizado por processo automatizado colocando a Analista de DBM numa posição de analista, de fato, para que novos indicadores ou novas análises fossem criadas.

Os processos existentes atualmente dão conta das demandas da área de *marketing* e geram as informações que foram especificadas no documento criado pela Analista de DBM, utilizado como material fonte para este trabalho. Os processos tem como resultado os seguintes relatórios:

- Atualização do HTML com dados sobre a base de clientes – Por Trimestre
- Atualização dos Contatos atingidos com as últimas ações – Por Trimestre
- Atualização das últimas atividades, com nº de convidados, confirmados, participantes, pesquisa de satisfação – Por Trimestre
- Atualização dos custos de *marketing* – Por Mês
- Relatório final da base de clientes – Anual
- Relatório final dos contatos atingidos – Anual
- Relatório final das últimas atividades x nº convidados, confirmados, participantes x custo - Anual
- Relatório final de custos de *marketing* – Anual

Uma descrição mais detalhada sobre cada um dos relatórios é feita nas seções seguintes: relatórios gerados com base nas informações de clientes, relatórios gerados com base nas informações de contatos, relatórios gerados com base nas informações de atividades e relatórios gerados com base nas informações de custos.

➢ Relatórios gerados com base nas informações de clientes

Estes relatórios têm por objetivo produzir métricas baseadas na receita trazida pelos clientes e pelos produtos vendidos a estes clientes. A Analista de DBM utiliza a aplicação SAS *Enterprise Guide*, ferramenta que permite o uso de técnicas estatísticas e funcionalidades de bancos de dados para manipular informações de diferentes fontes de dados, para gerar os relatórios relacionados a seguir no formato HTML (*Hyper Text Markup Language*).

- Análise de Cliente/Receitas.

- Evolução da Receita 2001 a 2004
- Clientes X Segmento 2001 a 2004
- Clientes X Produto/Solução 2001 a 2004
- *Revenue Weathervane* 2001 a 2004

- Faturamento Acumulado X Cliente 2001 a 2004
- *Top* 10 Clientes 2001 a 2004

- Atualização por Trimestre.

- Clientes X Segmento 2005
- Clientes X Produto/Solução 2005
- *Revenue Weathervane* 2005
- Lista de Clientes Atual x Trimestre 2005
- Receita X Trimestre 2005
- Evolução de Receita 2001 a 2005

➢ Relatórios gerados com base nas informações de contatos

Estes relatórios têm por objetivo produzir listas de nomes dos contatos atuais nas empresas onde a DBMTech tem seus sistemas instalados. A aplicação SFA (*Sales Force Automation*) é utilizada atualmente como origem para extração dos dados e seus resultados são gerados em planilhas eletrônicas. Esta aplicação controla e mantém um banco de dados com todos os contatos relacionados a cada empresa cliente ou potencial cliente da DBMTech.

- Controle de Contatos

- Lista de contatos para Brasil

1º Limpeza da Base: excluir aqueles contatos que aparecem mais de uma vez na planilha que atualmente é considerada a base de contatos da empresa.

2º Alterações no Aplicativo: após a limpeza da base de contatos, todas as modificações são realizadas no SFA.

3º Cruzamento da Base Limpa com Base SFA: as bases geradas no 1º e 2º passos são juntadas para obter uma única base de dados dos contatos.

4º Carrega dados no SFA: os dados adicionais de cada contato, como empresa, cargo, entre outros é incluído no SFA.

5º Carga das últimas listas de eventos no SFA: todos os eventos realizados, possuem uma lista de participantes que são registrados no SFA.

6º Criação de planilha sobre Perfil dos Contatos: os dados finais são gerados na planilha que contém o perfil de cada contato das empresas clientes da DBMTech.

- Lista de contatos para Argentina e Chile

1º Recebimento da Base limpa de Argentina e Chile: cada país possui sua própria lista de contatos e envia esta para a Analista de DBM que realiza todo o trabalho de consolidação dos contatos.

2º Criação do Aplicativo de Contatos para Argentina e Chile: todos os contatos são registrados no sistema SFA.

3º Cruzamento da Base limpa com Base SFA: as bases geradas no 1º e 2º passos são juntadas para obter uma única base de dados dos contatos.

4º Carga dos dados no SFA: os dados adicionais de cada contato, como empresa, cargo, entre outros é incluído no SFA.

5º Carga das últimas listas de eventos no SFA: todos os eventos realizados possuem uma lista de participantes que são registrados no SFA.

6º Criação de planilha sobre Perfil dos Contatos: os dados finais são gerados na planilha que contém o perfil de cada contato das empresas clientes da DBMTech.

➢ Relatórios gerados com base nas informações de atividades

Estes relatórios têm por objetivo analisar todas as atividades planejadas e realizadas pela área de *marketing* para manter o relacionamento contínuo com os clientes. A aplicação SIM (aplicação de controle de eventos e atividades do *marketing*) é utilizada atualmente como origem de dados para gerar os relatórios relacionados a seguir e seus resultados são gerados em planilhas eletrônicas.

- Controle de Atividades

- Para abrangência Brasil

1º Listar todas as atividades do Ano X Trimestre

2º A cada trimestre preencher o SIM com as atividades do Trimestre

3º Mensalmente enviar um e-mail atualizado sobre o andamento das atividades do trimestre para a Diretora de Marketing e Gerente de Marketing

- Para abrangência Argentina e Chile

1º Responsável pelos eventos informa todas as atividades do Ano

2º Listar todas as atividades do Ano x Trimestre

3º A cada trimestre preencher o SIM com as atividades do Trimestre

4º Mensalmente enviar um e-mail atualizado sobre o andamento das atividades do trimestre para a Diretora de Marketing.

➢ Relatórios gerados com base nas informações de custos

Estes relatórios têm por objetivo demonstrar todo o custeio das atividades realizadas pela área de *marketing*. A aplicação de controle financeiro é utilizada atualmente como origem de dados para gerar os relatórios relacionados a seguir e seus resultados são gerados em planilhas eletrônicas.

- Controle de Custos

1º Fazer novo comunicado a todos sobre processo de controle de custos do *marketing*

2º Cobrança dos Custos Antecipados das Ações

3º Recebimento – Cadastramento – Assinatura - Entrega

4º Comparativo mensal com os relatórios do Analista Financeiro

5.2 Abordagem Proposta Aplicada ao Estudo de Caso

O material apresentado pela Analista de DBM possui um relato muito superficial e com muitas falhas no que se refere ao detalhe das informações necessárias para a produção dos relatórios, como:

- ausência da definição das fontes de informações (sistemas e dados detalhados).
- ausência da definição das informações geradas nos relatórios.
- ausência do detalhe das atividades no processo de geração dos relatórios.

Com estas falhas ressaltadas não é possível a uma equipe de implementação de sistemas utilizá-lo para análise e desenho de um DBM.

Dada esta necessidade, o objetivo deste estudo de caso é utilizar práticas participativas para a elicitação de requisitos do DBM para que seja desenhado um modelo de dados que atenda às necessidades da DBMTech para análise e produção dos relatórios solicitados.

5.2.1 O uso de Práticas Participativas

Algumas práticas participativas foram consultadas em [2] e avaliadas para que as mais adequadas a este processo de elicitação de requisitos para uma base de dados fossem utilizadas. Ao final da análise deste material, 3 práticas participativas foram escolhidas para utilização: *Group Elicitation Method*, *Blueprint Mapping* e PICTIVE. A escolha foi baseada em alguns critérios como: permitir a análise sistemática dos processos atuais e desenhá-los com a participação do usuário, permitir o detalhamento de cada processo de forma organizada a fim de identificar os *stakeholders*, as fontes de dados e os relatórios gerados, e permitir o uso de artefatos visuais para a criação de um diagrama que pudesse ser modificado durante as seções de trabalho, além de identificar aquelas técnicas com maior aderência à fase de Análise e Requisitos.

- ***Group Elicitation Method*** – esta prática foi utilizada para que fossem discutidos e entendidos, com mais detalhes as necessidades apresentadas pela Analista de DBM no documento que encontra-se no Anexo I.

- ***Blueprint Mapping*** – esta prática foi utilizada para que fossem definidos os processos de geração de relatórios, e identificadas suas fontes de dados e os *stakeholders* nos processos;

- ***PICTIVE*** – esta prática foi utilizada para que fossem detalhadas as entidades de dados identificadas através do resultado do processo desenhado no *Blueprint Mapping*, permitindo um detalhamento do modelo de dados e fornecendo subsídios necessários para a construção de um *Star Schema*.

O diagrama ilustrado na Figura 5-2, ilustra a maneira como as técnicas foram combinadas e utilizadas. Inicialmente uma análise preliminar foi realizada para clarificar o entendimento das necessidades da DBMTech por meio do *Group Elicitation Method* utilizando o documento com detalhes dos processos atuais de geração de relatórios. O resultado obtido foi a validação das necessidades, priorização e tabelas com a seleção dos processos a serem detalhados na próxima etapa. No próximo passo, o *BluePrint Mapping* foi utilizado para o desenho dos processos de criação dos relatórios, identificação dos *stakeholders*, o fluxo dos dados, os relatórios gerados como resultados dos processos, assim como as origens das informações. Ao final desta etapa os dados de origem e os dados utilizados nos relatórios foram identificados em cada um dos processos. A partir dos processos com os dados identificados e também dos leiautes de relatórios apresentados foi possível detalhar cada uma destas entidades de dados através do *PICTIVE*,

resultando em um diagrama de entidades e atributos que definiu o modelo de dados conceitual do DBM. As etapas que envolveram o *BluePrint Mapping* e o *PICTIVE* foram realizadas através de iterações porque os dados identificados na etapa do *PICTIVE* algumas vezes exigiam a reavaliação e atualização do processo desenhado na etapa do *BluePrint Mapping*.

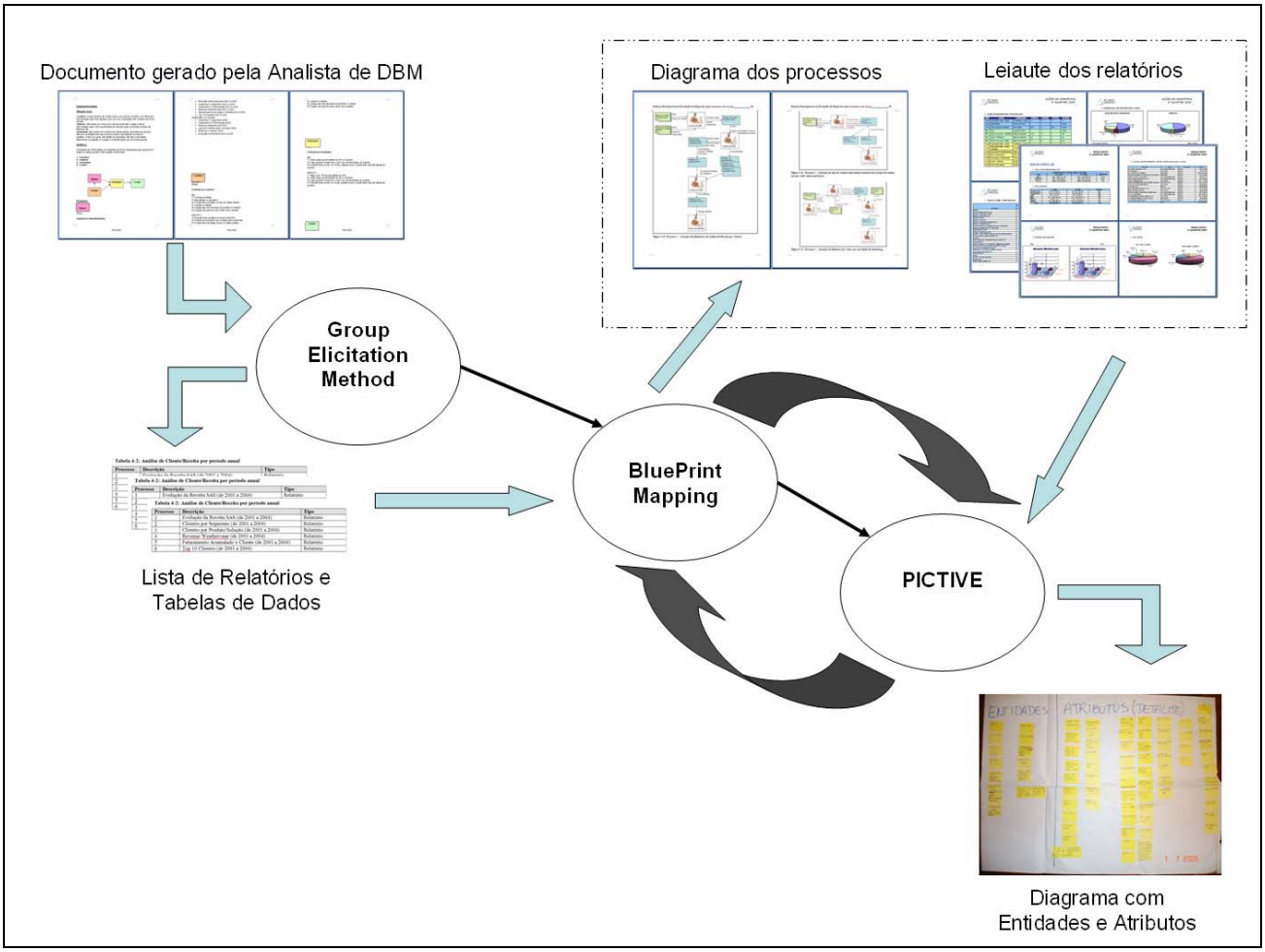


Figura 5-2: Ilustração do método criado para elicitação de requisitos com uso das técnicas participativas.

➢ *Group Elicitation Method* no Contexto do Estudo de Caso

Quando a discussão do problema foi iniciada, um breve relato das necessidades e das atividades realizadas atualmente foi apresentada por meio de um documento criado pela Analista de DBM. Utilizando como base o documento que foi entregue, foi realizada uma sessão de *brainstorming* e esclarecimentos, conforme a sugestão deste método no suporte a sessões de *brainstorming*, o que permitiu a exploração de diversas idéias para saber o que seria necessário elicitar neste momento e o que poderia ser deixado para o futuro.

Desta forma foi possível clarificar a problemática e identificar os pontos importantes que deveriam ser considerados em termos de processos, relatórios e dados resultando em algumas tabelas com os relatórios e tabelas de dados a serem gerados, estas tabelas estão descritas no capítulo 5.3.1.

➢ *BluePrint Mapping* no Contexto do Estudo de Caso

Este método foi utilizado para definir os *stakeholders*, o fluxo dos dados, os relatórios gerados como resultados dos processos, assim como as origens das informações que fazem parte do cenário da DBMTech e que devem ser consideradas no processo de elicitação de requisitos para o DBM e utilização nos relatórios e análises produzidas. Por exemplo: o analista financeiro gera a planilha de custos referente a cada ação/atividade realizada; neste caso identificamos o *stakeholder*, que é o analista financeiro, os dados disponibilizados (dados de custos e os relatórios gerados) que permitem analisar os custos por ação/atividade.

Foi possível envolver as pessoas do processo de criação de relatórios e determinar as suas atividades diárias através de um mapa que representa as tarefas realizadas para a obtenção dos dados ou para a produção das tabelas e relatórios, conforme a representação da Figura 5-3.

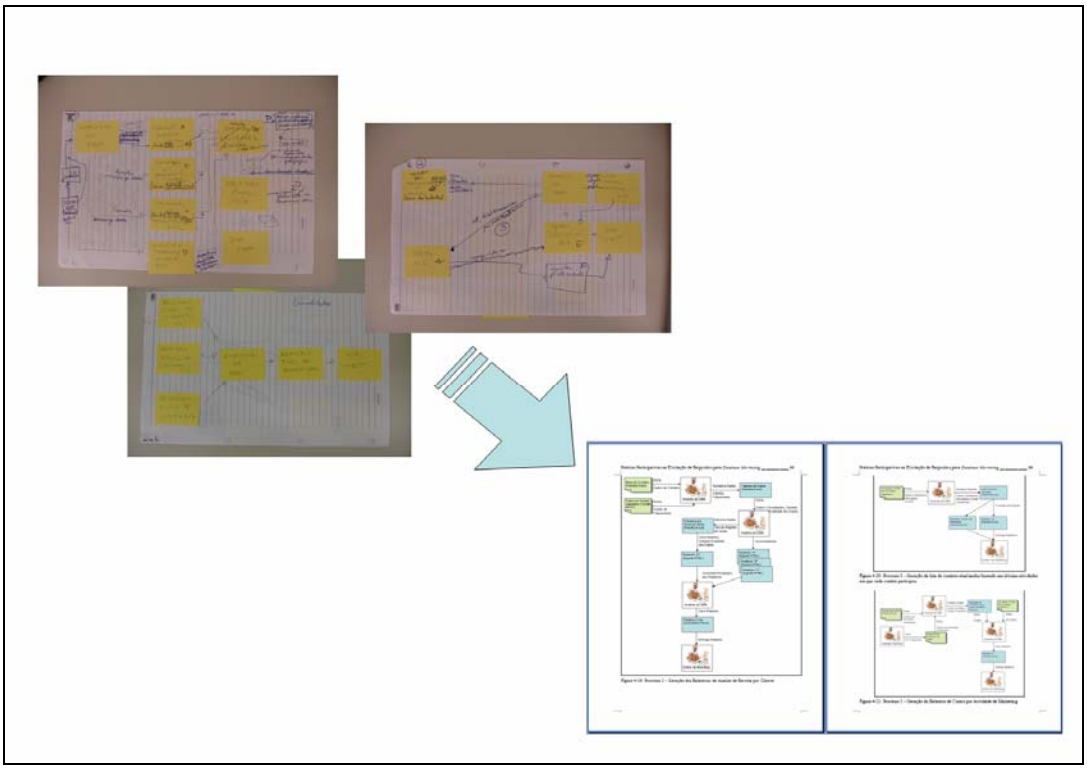


Figura 5-3: Diagrama gerado durante as sessões onde o *BluePrint Mapping* foi utilizado e resultado após o término das sessões.

Durante a revisão do processo desenhado, o diagrama gerado facilitou a identificação dos dados existentes e suas fontes, das informações que serão geradas em relatórios ou tabelas de dados, assim como a forma como estes dados são utilizados e gerados.

É importante salientar que este trabalho é desenvolvido durante uma oficina de trabalho onde os participantes fazem a diagramação das atividades com base em seu próprio conhecimento da organização, não existe um roteiro definido a ser seguido durante a oficina de trabalho para definir os papéis e as tarefas que farão parte do cenário e que devem ser envolvidos no processo de levantamento e utilização.

➢ PICTIVE no Contexto do Estudo de Caso

Este método permitiu aos participantes realizarem a prototipagem da aparência e a dinâmica de um sistema com interface gráfica utilizando recursos de baixa tecnologia. Esta técnica é muito útil para a fase de desenho, mas pode ser usada no levantamento de um protótipo feito em papel, ou para a análise [2].

Baseando-se nesta dinâmica, uma adaptação foi realizada para que os mesmos recursos pudessem ser utilizados, papel, *post-it notes*, canetas; e pudessemos detalhar as entidades de dados identificadas a partir do processo desenhado através do método participativo *Blueprint Mapping*.

Para esta adequação o quadro ilustrado na Figura 5-4 foi utilizado como ponto de partida para que o preenchimento fosse realizado ao longo dos trabalhos deste estudo de caso, com todas as entidades e os atributos que eram identificadas e documentados por meio de *post-its*.

Exemplo:

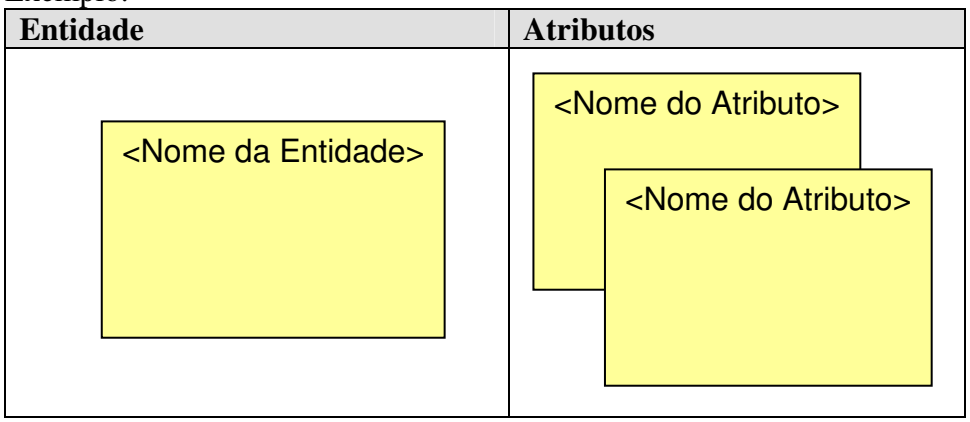


Figura 5-4: Quadro de Mapeamento das Entidades e seus atributos.

Através do processo desenhado no *BluePrint Mapping*, conseguimos identificar as entidades e seus atributos, mas em alguns casos surgiu a necessidade da alteração do processo previamente desenhado no *BluePrint Mapping*. A revisão do processo permitiu identificar, nesta fase inicial do projeto, falhas que poderiam trazer grandes transtornos e custos durante a implementação, podemos citar como exemplo, a ausência da informação na granularidade adequada utilizando a variável tempo como exemplo, se somente o ano for armazenado não é possível fazer qualquer relatório agregado por trimestre ou mês, e esta falha vai gerar uma nova requisição para que este dado seja implementado, gerando trabalho adicional a equipe de implementação.

5.3 Visão Geral do Processo

Para a realização dos trabalhos foram organizadas 4 oficinas de trabalho conforme o diagrama da Figura 5-5, onde cada um deles viabilizou o uso de diferentes técnicas participativas. Na primeira oficina de trabalho foi utilizada a técnica *Group Elicitation Method* para estudo de viabilidade e análise e elicitação de requisitos; na segunda oficina de trabalho foi utilizada a técnica *BluePrint Mapping* para análise e elicitação de requisitos, e na terceira e quarta oficina de trabalho foram utilizadas as técnicas *BluePrint Mapping* e *PICTIVE* para a especificação e validação dos requisitos.

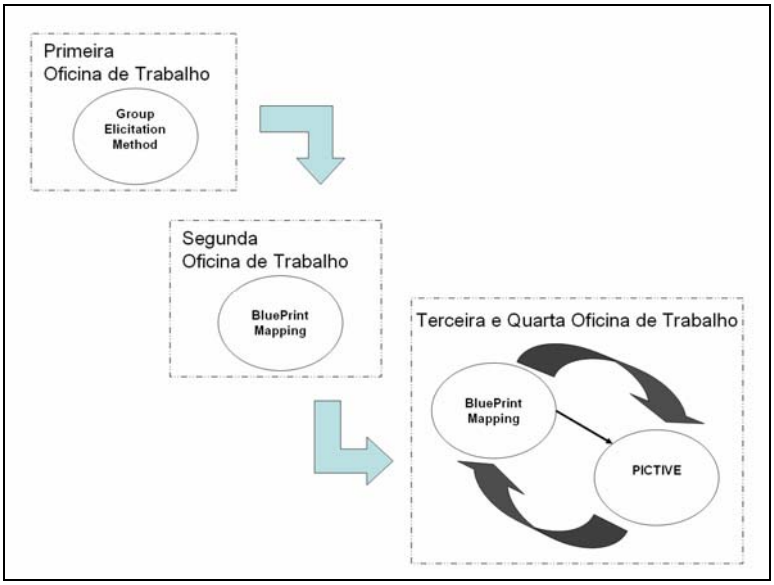


Figura 5-5: Organização das Oficinas de Trabalho

5.3.1 Primeira Oficina de Trabalho

Em uma sessão de 2 horas o Analista de DBM e o Analista de Requisitos se reuniram para a validação e detalhamento das necessidades por meio da leitura de um documento elaborado com as necessidades da DBMTech e os processos atuais, descrito no capítulo 5.1.2. Muitas dúvidas foram esclarecidas pela Analista de DBM, alguns pontos foram descartados e outros priorizados.

Ao realizar a leitura do documento foram identificadas primeiramente, uma série de processos que produziam diferentes relatórios para diferentes assuntos e que precisavam ser

organizados para que os requisitos fossem especificados de formas comuns aos grupos de relatórios que tivessem as mesmas informações apresentadas, evitando inconsistências.

A seguir, a Tabela 5-2, a Tabela 5-3, a Tabela 5-4 e a Tabela 5-5 onde demonstram os relatórios identificados e seus respectivos processos e formato dos resultados:

Tabela 5-2: Análise de Cliente/Receita por período anual

Processo	Descrição	Tipo
1	Evolução da Receita (de 2001 a 2004)	Relatório
2	Clientes por Segmento (de 2001 a 2004)	Relatório
3	Clientes por Produto/Solução (de 2001 a 2004)	Relatório
4	Revenue Weathervane (de 2001 a 2004)	Relatório
5	Faturamento Acumulado x Cliente (de 2001 a 2004)	Relatório
6	Top 10 Clientes (de 2001 a 2004)	Relatório

Tabela 5-3: Assunto: Análise de Cliente/Receita por trimestre

Processo	Descrição	Tipo
1	Evolução da Receita (de 2001 a 2005)	Relatório
2	Clientes por Segmento (2005)	Relatório
3	Clientes por Produto/Solução (2005)	Relatório
4	Revenue Weathervane (2005)	Relatório
5	Receita Anual da Empresa x Trimestre (2005)	Relatório
6	Lista de Clientes Atual x Quarter (2005)	Relatório
7	Base de contatos extraída do SFA e aplicativo Web.	Tabela de Dados

Tabela 5-4: Custos (Trimestral, Mensal e Anual)

Processo	Descrição	Tipo
1	HTML atualizado com dados sobre a base de clientes	Relatório
2	Lista de contatos atualizados baseado nas últimas atividades realizadas	Relatório
3	Atividades x (qtd. de convidados, confirmados, participantes,custos)	Relatório
4	Custos por atividade de Marketing	Relatório

Tabela 5-5: Atividades

Processo	Descrição	Tipo
1	Contatos por Atividade (Anual, com quebra trimestral)	Relatório

Ao final desta oficina de trabalho, foi concluído, através dos dados apresentados e tabulados a partir do documento de necessidades, que existiam 18 processos para a produção de 18 relatórios ou tabelas de dados. Ficou entendido que dentre estes processos muitos deles eram similares e deveriam ser analisados de forma a simplificar e reduzir o número de processos, tornando a realização das atividades da área de *marketing* mais simples. Este resultado foi utilizado como entrada para a segunda oficina de trabalho onde foi trabalhado o aspecto de redução do número de processos e o desenho dos processos definidos por meio dos artefatos apresentados na Figura 5-6.

5.3.2 Segunda Oficina de Trabalho

Em uma sessão de 2 horas, o Analista de DBM e o Analista de Requisitos se reuniram para reduzir o número de processos e definir o desenho destes processos, que foram identificados na primeira oficina de trabalho. Para realizar estas atividades foi utilizada a técnica participativa *Blueprint Mapping* para elicitar as tarefas que realmente faziam parte dos processos identificados e para identificar os *stakeholders* assim como o fluxo dos dados utilizados na geração dos relatórios. Somente a Analista de DBM participou desta oficina de trabalho, pelo lado dos usuários, juntamente com a pesquisadora por ser a única envolvida no processo de manipulação das informações, apesar de outras áreas e outras pessoas da área de *marketing* estarem envolvidas, de alguma forma, no processo de geração da informação.

Ao iniciar os trabalhos, foram criados *post-its* com o nome de cada um dos 18 relatórios nomeados na primeira oficina de trabalho. Foram criados também *post-its* com o nome dos *stakeholders* identificados inicialmente:

- Analista de DBM, que em todos os processos manipula as informações para produzir os relatórios.

- Diretora de Marketing, que em todos os processos recebe os relatórios criados.

Os processos criados ao longo desta oficina de trabalho foram gerados da forma apresentada na Figura 5-6.

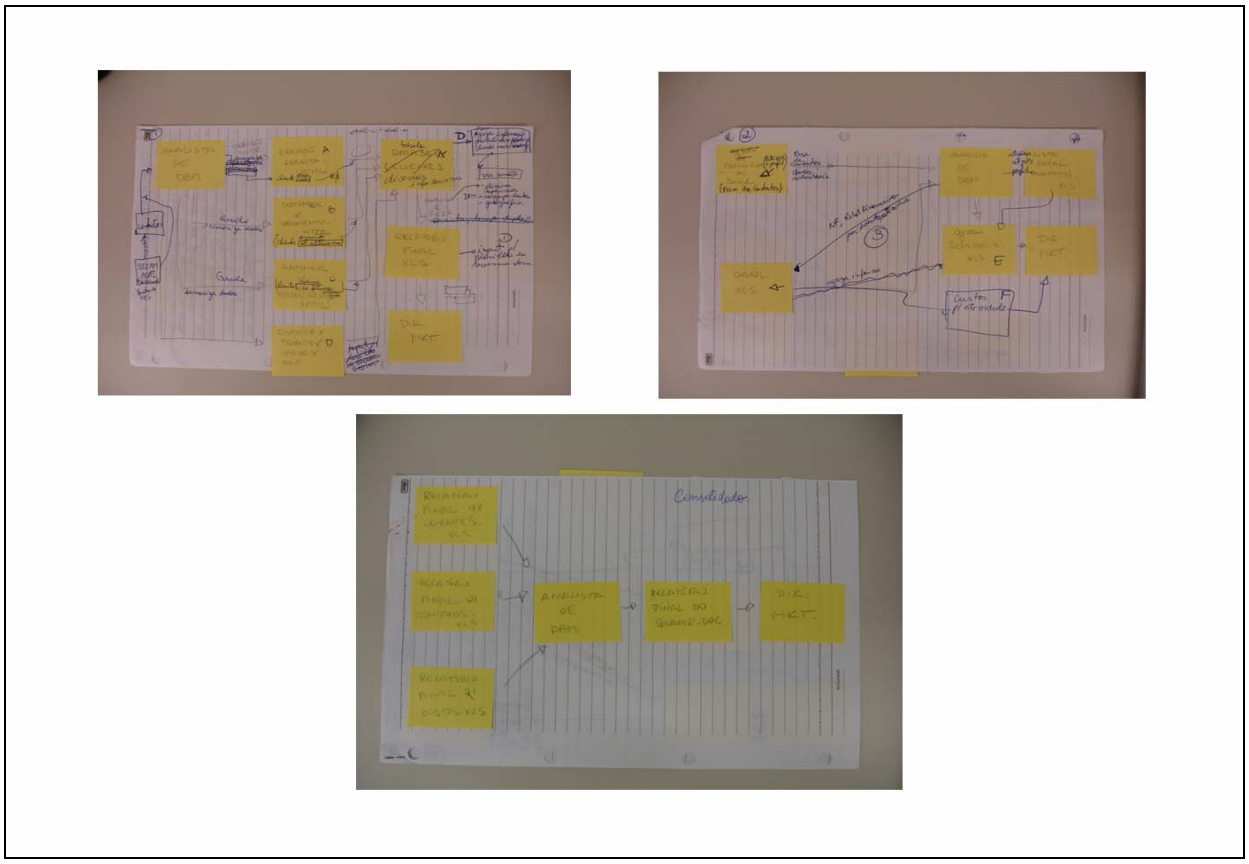


Figura 5-6: Artefato produzido na segunda oficina de trabalho por meio da técnica *BluePrint Mapping*.

Dando continuidade ao trabalho de redução do número de processos, os *post-its* referentes aos relatórios que possuíam o mesmo tipo de informação e que eram visualizados por períodos diferentes, por exemplo, anual, trimestral, foram agrupados. Após esta análise foi concluído que 8 dos relatórios identificados eram produzidos pelo mesmo processo com apenas 4 variações, ou seja, 4 relatórios com os resultados, e outros 4 relatórios eram produzidos por este mesmo processo porém com uma pequena variação. Por meio do *BluePrint Mapping* foi possível dispor de forma visual os relatórios e identificar 12 dos 18 relatórios como parte de um único processo. Este processo foi denominado "Processo 1" e tem como objetivo principal a análise de informações cadastrais dos clientes e faturamento destes clientes. A Figura 5-7 ilustra o Processo 1.

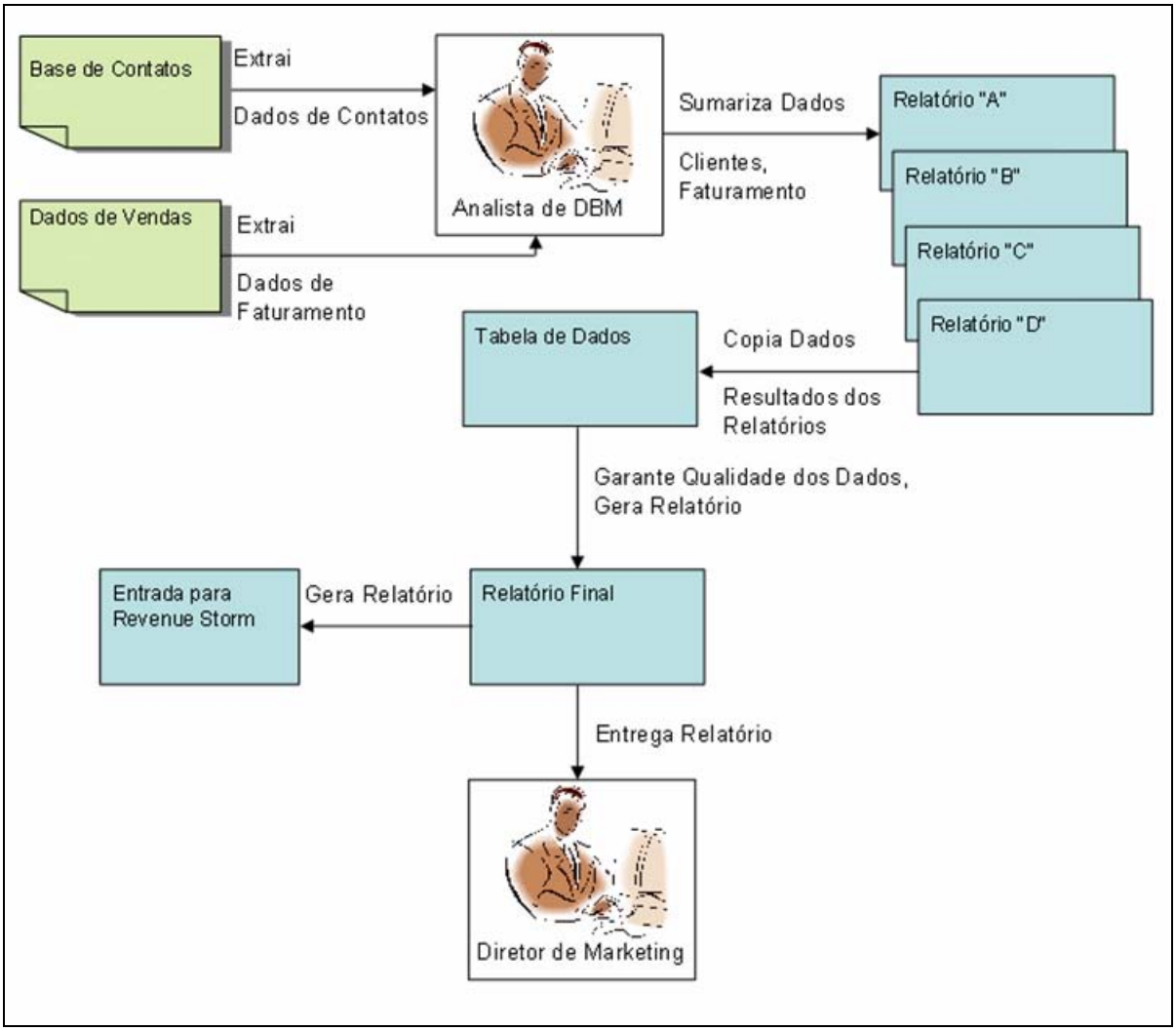


Figura 5-7: Processo 1 - Geração dos Relatórios de Análise de Receita por Cliente.

Tabela 5-6: Relatórios analisados no Processo 1.

Relatórios e suas descrições	
A	- Evolução da Receita SAS (de 2001 a 2004) anual - Evolução da Receita SAS (de 2001 a 2004) trimestral
B	- Clientes por Segmento (de 2001 a 2004) anual - Clientes por Segmento (de 2001 a 2004) e trimestral
C	- Clientes por Produto/Solução (de 2001 a 2004) anual - Clientes por Produto/Solução (de 2001 a 2004) trimestral
D	- Revenue Weathervane (de 2001 a 2004) anual - Revenue Weathervane (de 2001 a 2004) trimestral
Relatório Final	- Faturamento Acumulado x Cliente (de 2001 a 2004) - Top 10 Clientes (de 2001 a 2004) - Receita Anual da Empresa x Trimestre (2005)

Na Figura 5-7, a origem de dados "Base de Contatos" e os relatórios "Tabela de Dados", "Relatório Final" e "Entrada para Revenue Storm" são disponibilizados em Planilha Excel. Os relatórios "A", "B", "C" e "D" são disponibilizados em formato HTML. Os "Dados de Vendas" são disponibilizados pelo aplicativo *Steam Boat*.

	- Lista de Clientes Atual x Trimestre (2005)

Restaram 6 processos para serem analisados. O mesmo princípio foi aplicado para agrupamento utilizado na definição do Processo 1; foi definido um grupo de relatórios através da comparação dos dados fontes acessados para a geração dos resultados; 4 dos 6 relatórios apresentados foram definidos como parte deste processo. Este processo foi denominado "Processo 2" e tem como objetivo principal a análise de informações dos contatos dos clientes e informações das atividades de *marketing* das quais estes contatos participaram. A Figura 5-8 ilustra o Processo 2.

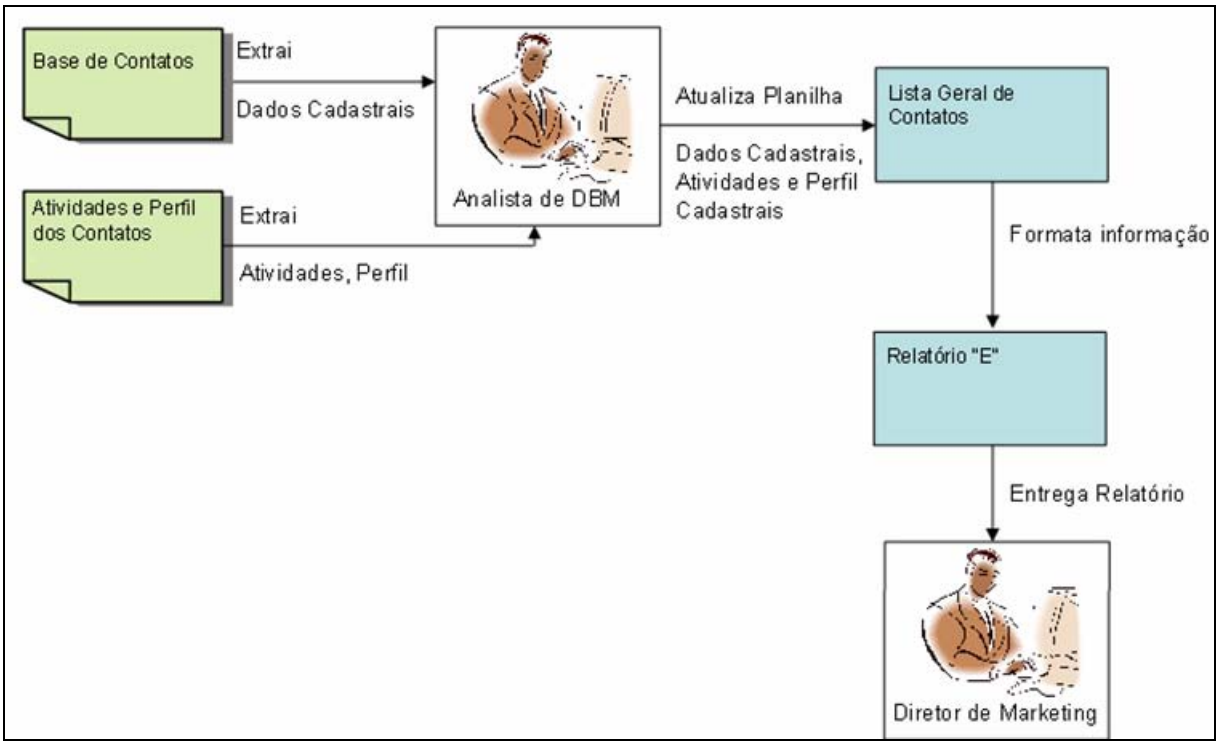


Figura 5-8: Processo 2 - Geração da lista de contatos atualizados baseado nas últimas atividades em que cada contato participou.

Tabela 5-7: Relatórios analisados no Processo 2.

Relatórios e suas descrições	
E	- Lista de contatos atualizada baseada nas últimas atividades realizadas - Contatos por Atividade (Anual, com quebra trimestral) - HTML atualizado com dados sobre a base de clientes
Base de Contatos	- Base de contatos extraída do SFA e aplicativo Web

Na Figura 5-8, a origem de dados "Base de Contatos" e os relatórios "Lista Geral de Contatos" e "Relatório E" são disponibilizados em Planilha Excel. A origem de dados "Atividades e Perfil dos Contatos" é disponibilizada pelo aplicativo SIM.

Finalizando esta etapa, foi desenhado um último processo que resulta na produção dos 2 relatórios restantes. Ao desenhar este processo utilizando os *post-it* foi possível identificar um *stakeholder* que não havia sido mencionado: o Analista Financeiro, que produz os dados utilizados nas análises de custos. As informações geradas por ele são de extrema importância e, em nenhum momento anterior, ele tinha sido mencionado no processo de desenvolvimento das análises de informações. Este processo foi denominado "Processo 3" e tem como objetivo principal a análise de informações de custos das atividades de *marketing*. A

Figura 5-9 ilustra o Processo 3.

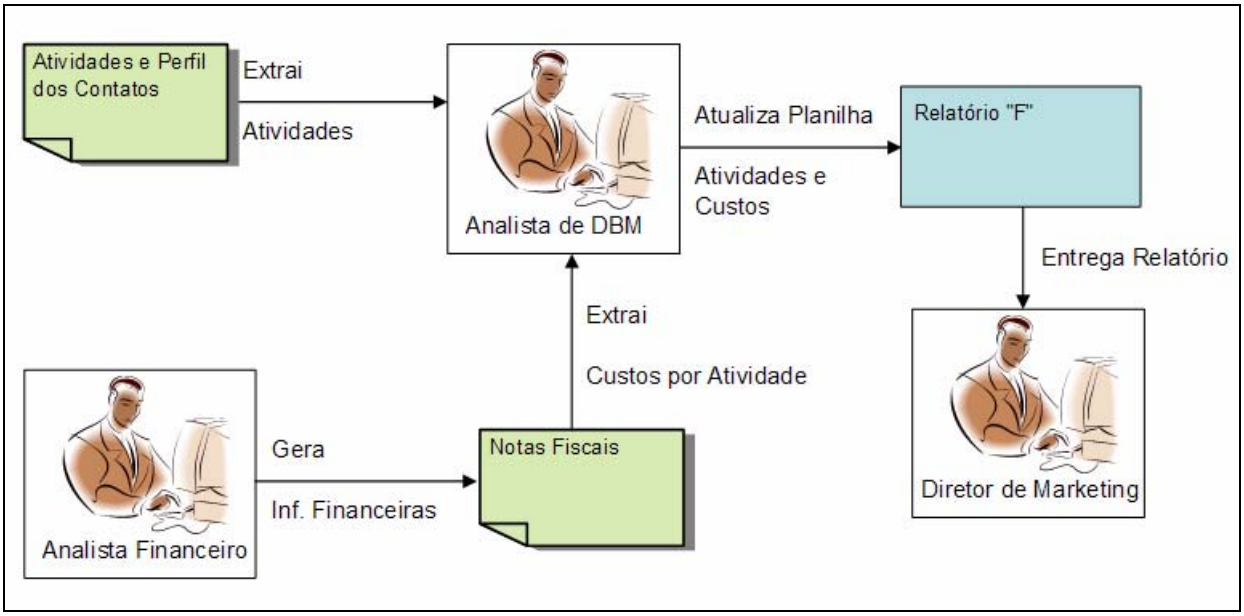


Figura 5-9: Processo 3 - Geração do Relatório de Custos por Atividade de Marketing.

Tabela 5-8: Relatórios analisados no Processo 3.

Relatórios e suas descrições	
F	- Custos por atividade de Marketing - Qtd. de Convidados, Confirmados, Participantes e Custos por Atividade

Ao final desta oficina de trabalho, a quantidade de processos foi reduzida de 18 para 3, consolidando nestes processos as informações que são geradas a partir das mesmas bases de dados de origem ou ainda que resultam num mesmo tipo de informação. Ao longo desta oficina de trabalho foram mapeados todos os *stakeholders* envolvidos diretamente no processo além de

Na

Figura 5-9, a origem de dados "Notas Fiscais" e o relatório "Relatório F" são disponibilizados em Planilha Excel. A origem de dados "Atividades e Perfil dos Contatos" é disponibilizada pelo aplicativo SIM.

traçar todo o fluxo de dados desde a sua origem, que são as fontes de dados, até a sua utilização; foram mapeados também os tipos de dados usados em cada relatório resultante.

A dinâmica utilizada facilitou ao usuário, a Analista de DBM, explicar e detalhar cada um dos processos. Além disso, a facilidade de realizar mudanças no processo ou acrescentar etapas, deixou o usuário à vontade para definir o processo com detalhes. Como já existia um processo para criar os relatórios atuais, alguns passos foram explicitados com a ajuda da metodologia, permitindo identificar passos implícitos no processo inicialmente relatado pelos *stakeholders*. Os artefatos produzidos até este momento foram claramente posicionados no diagrama permitindo identificar os dados existentes e quais dados são necessários para produzir estes relatórios.

Este resultado foi utilizado como entrada para a terceira e quarta oficina de trabalho onde trabalhamos na elicitação dos dados relevantes para a estruturação do DBM.

5.3.3 Terceira e Quarta Oficinas de Trabalho

Foram realizadas duas sessões que totalizaram 4 horas de trabalho com o objetivo de revisar cada um dos 3 processos desenhados nas oficinas de trabalho realizadas anteriormente e extrair as informações necessárias para o desenho do modelo de dados *Star Schema* através da técnica participativa PICTIVE adequada à necessidade de banco de dados (adequação demonstrada no capítulo 5.2.1). Participou desta oficina de trabalho somente a Analista de DBM.

Os trabalhos foram iniciados pela análise de cada um dos processos, 1, 2 e 3, e foram identificados inicialmente os dados utilizados como entrada em cada processo e os mesmos foram nomeados como entidades. A fim de facilitar a identificação destas entidades nos desenhos dos processos, as mesmas foram circundadas por elipses como mostrado nos diagramas da Figura 5-10, Figura 5-11 e Figura 5-12.

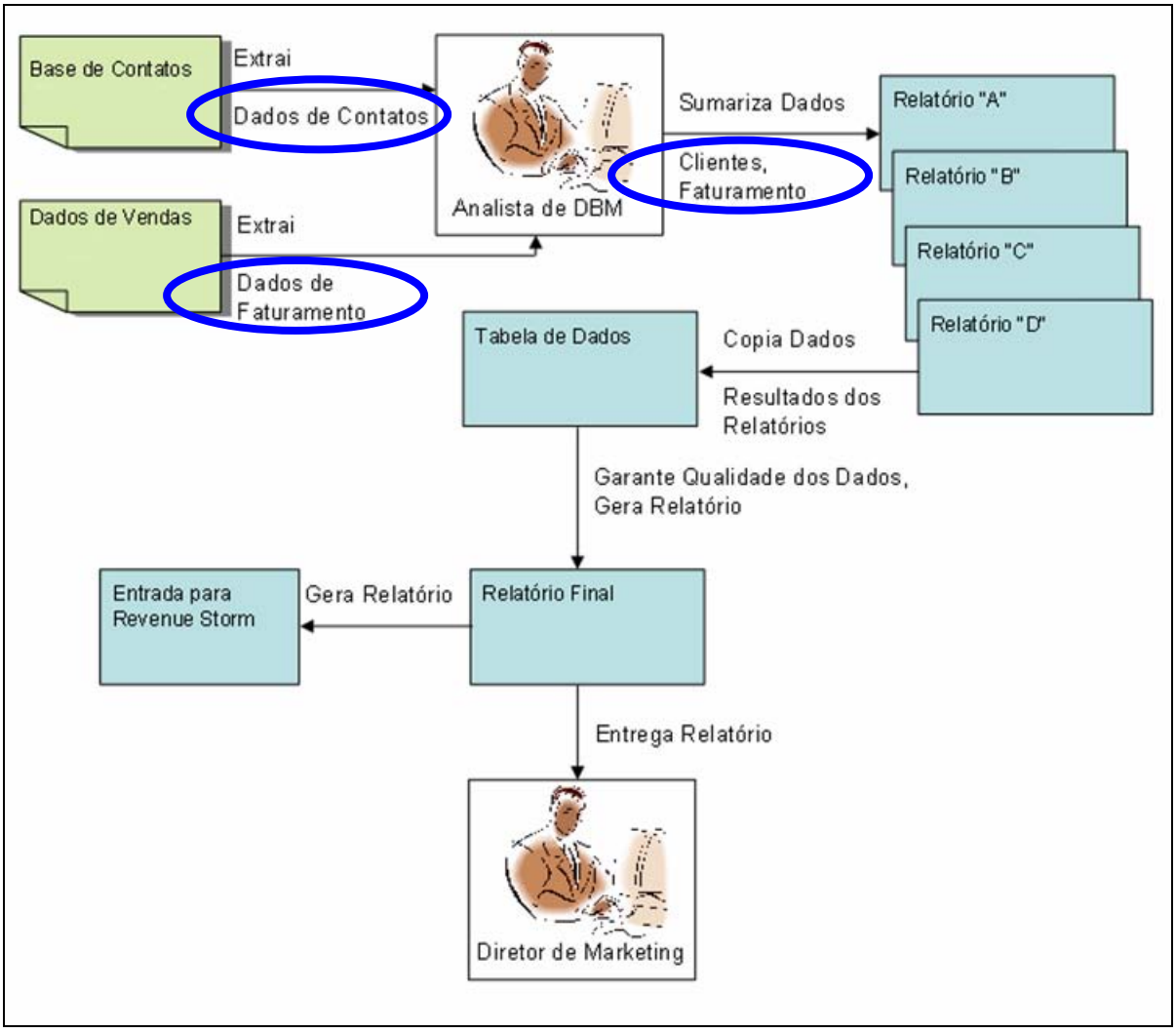


Figura 5-10: Processo 1 - Geração dos Relatórios de Análise de Receita por Cliente com identificação das entidades.

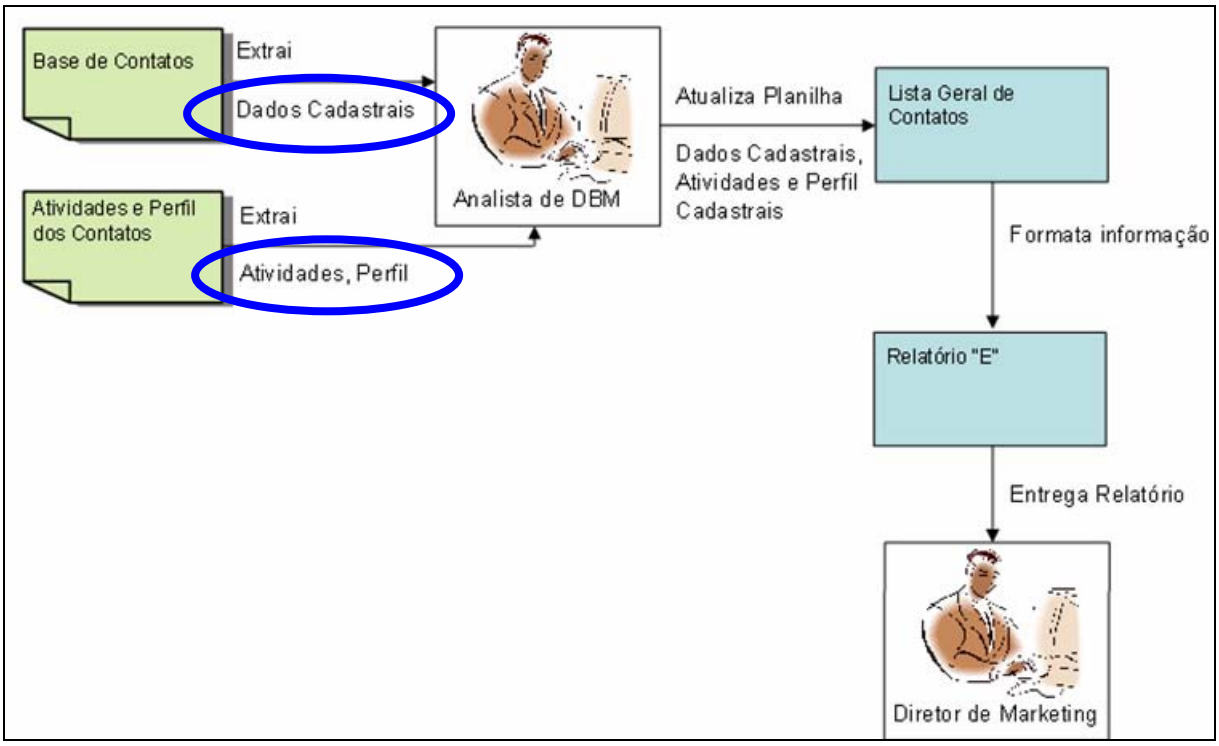


Figura 5-11: Processo 2 - Geração da lista de contatos atualizados baseado nas últimas atividades em que cada contato participou, com identificação das entidades.

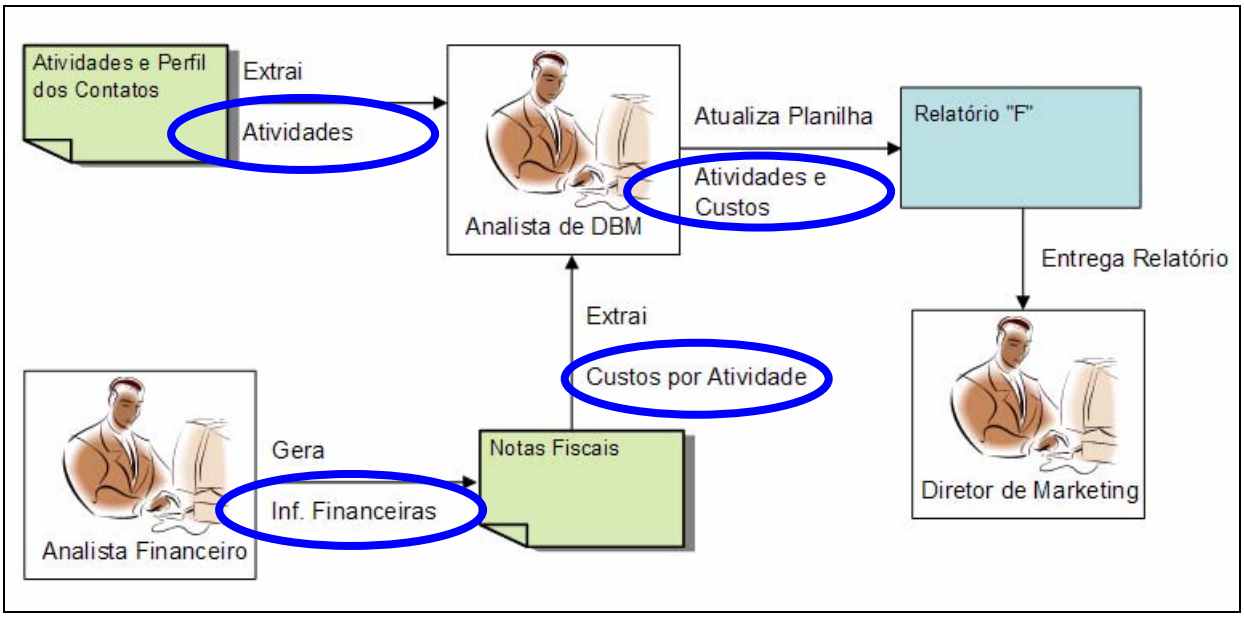


Figura 5-12: Processo 3 - Geração do Relatório de Custos por Atividade de Marketing, com identificação das entidades.

Visualizando os processos 1, 2 e 3 ilustrados nas respectivas figuras; 9 entidades foram identificadas através desta análise preliminar:

- Dados de Contato.
- Dados de Faturamento.
- Clientes.
- Dados Cadastrais.
- Perfil.
- Atividades.
- Custos (Mktg).
- Custos por Atividade.
- Custos Financeiro (Inf. Financeiras).

Para cada uma destas entidades foram criados *post-its* e anexados ao quadro descrito no capítulo 5.2.1 na coluna "ENTIDADE".

Até o início da terceira oficina de trabalho nenhum formato de relatório havia sido apresentado, para que os detalhes ou campos de dados utilizados fossem identificados; a única informação obtida havia sido o tipo de dados utilizado, por exemplo, dados de custos, informação dos contatos, entre outras, que permitiram a identicação das entidades. Para identificar os atributos de cada entidade uma exploração detalhada em cada relatório ou tabela foi necessária em cada um dos 3 processos desenhados.

O processo de análise foi iniciado pelo processo 1, que tem como resultado os relatórios A, B, C e D.

- **Análise do Processo 1**

Relatório A: Evolução da Receita (2001 a 2004)

Atributos identificados:

- Nome do Cliente.
- Valor do Faturamento Anual do Cliente (Histórico de todos os anos).

➢ **EVOLUÇÃO DA RECEITA (25 CLIENTES MAIS RENTÁVEIS)**

	Invoice Year				TOTAL ACUMULADO	%PARTICIPAÇÃO NO FAT. ACUMULADO
	2001	2002	2003	2004		
	Amount in Local	Amount in Local	Amount in Local	Amount in Local		
Cliente 1	R$ 100.00	R$ 105.00	R$ 110.00	R$ 115.00	R$ 430.00	17%
Cliente 2	R$ 90.00	R$ 95.00	R$ 100.00	R$ 105.00	R$ 390.00	16%
Cliente 3	R$ 80.00	R$ 85.00	R$ 90.00	R$ 95.00	R$ 350.00	14%
Cliente 4	R$ 70.00	R$ 75.00	R$ 80.00	R$ 85.00	R$ 310.00	12%
Cliente 5	R$ 60.00	R$ 65.00	R$ 70.00	R$ 75.00	R$ 270.00	11%
Cliente 6	R$ 50.00	R$ 55.00	R$ 60.00	R$ 65.00	R$ 230.00	9%
Cliente 7	R$ 40.00	R$ 45.00	R$ 50.00	R$ 55.00	R$ 190.00	8%
Cliente 8	R$ 30.00	R$ 35.00	R$ 40.00	R$ 45.00	R$ 150.00	6%
Cliente 9	R$ 20.00	R$ 25.00	R$ 30.00	R$ 35.00	R$ 110.00	4%
Cliente ...	R$ 10.00	R$ 15.00	R$ 20.00	R$ 25.00	R$ 70.00	3%

Figura 5-13: Relatório A - Evolução da Receita (de 2001 a 2004).

As informações de percentual serão criadas em tempo de execução do relatório para representar o percentual de crescimento e o percentual de participação no total de faturamento. Não há necessidade de termos estas variáveis pré-definidas nas bases de dados por ser um dado variável em relação ao faturamento anual e à quantidade de clientes existente.

Relatório B: Clientes por Segmento (de 2001 a 2004)

Atributos identificados:

- Valor do Faturamento = Valor de Novas Vendas + Valor de Renovação.
- Segmento de Indústria do Cliente.
- Nome do Cliente.

➢ **CLIENTES POR SEGMENTO (DE 2001 a 2004)**

EMPRESA	SETOR	VALOR
Cliente 1	SETOR A	R$ 100.00
Cliente 2	SETOR B	R$ 90.00
Cliente 3	SETOR A	R$ 80.00
Cliente 4	SETOR C	R$ 70.00
Cliente 5	SETOR B	R$ 60.00
Cliente 6	SETOR B	R$ 50.00
Cliente 7	SETOR C	R$ 40.00
Cliente 8	SETOR A	R$ 30.00
Cliente 9	SETOR A	R$ 20.00
Cliente ...	SETOR B	R$ 10.00
	TOTAL	**R$ 550.00**

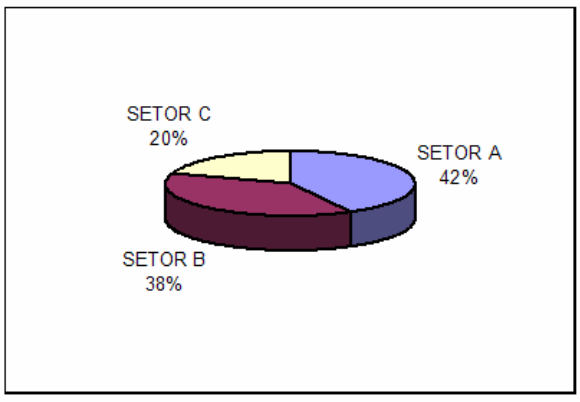


Figura 5-14: Relatório B - Clientes por Segmento (de 2001 a 2004).

Relatório C: Clientes por Produto/Solução

Atributos identificados:

- Nome da Empresa.
- Produto ou Solução Comprado.
- Valor da Compra.

➢ **LISTA DE CLIENTES (EMPRESA X PRODUTO X VALOR)**

EMPRESA	PRODUTO	VALOR
Cliente 1	TOOLS	R$ 10,00
Cliente 2	SOLUTION 1	R$ 10,00
Cliente 3	TOOLS	R$ 10,00

Figura 5-15: Relatório C - Clientes por Produto/Solução.

Relatório D: Tabela de Clientes e Produtos usadas no Revenue Weathervane

Atributos identificados:

- Tipo de Indústria.
- Produto.
- Nome do Cliente.
- Tipo do Registro de Venda (Nova/Renovação).
- Valor do Faturamento.
- Existing Places (Cliente já existente).
- New Places (Cliente Novo).
- Existing Stuffs (Produto que já foi vendido no Brasil).
- New Stuffs (Produto vendido no Brasil pela 1ª. vez).

Tipo de Indústria Categoria	Código do Produto	Nome do Cliente	Tipo de Registro de Venda	Total em Moeda Local	Existing Places	New Places	Existing Stuffs	New Stuff
SETOR A	PROD 1	CLIENTE 1	R	R$ 0.00	X		X	
SETOR A	PROD 2	CLIENTE 1	R	R$ 0.00	X		X	
SETOR A	PROD 3	CLIENTE 1	R	R$ 0.00	X		X	
SETOR A	PROD 4	CLIENTE 1	R	R$ 0.00	X		X	
SETOR B	PROD 1	CLIENTE 2	N	R$ 0.00		X	X	
SETOR B	PROD 2	CLIENTE 2	N	R$ 0.00		X	X	

Figura 5-16: Fonte de Dados para o Relatório D - Tabela de Clientes e Produtos.

Na Figura 5-16, os termos “Existing Places”, “New Places”, “Existing Stuffs”, “New Stuff” foram mantidos em inglês por opção do próprio usuário em adotar um vocabulário utilizado pela DBMTech.

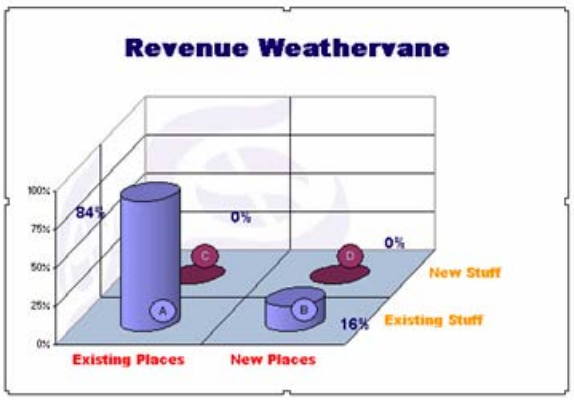


Figura 5-17: Revenue Weathervane – Relatório gerado a partir dos dados apresentados na Figura 5-16.

Ao final desta oficina de trabalho, foi possível determinar todos os atributos necessários para as entidades identificadas nos relatórios do Processo 1.

A dinâmica permitiu ao usuário, a Analista de DBM, identificar um passo que foi omitido na segunda oficina de trabalho durante o desenho do Processo 1.

Neste processo uma das fontes de dados mapeada é utilizada para a obtenção de dados de contato: a "Base de Contatos – Planilha Excel". Porém, uma nova fonte será disponibilizada após a implementação da aplicação que controlará todo o gerenciamento dos contatos de cada empresa. Esta fonte de dados não está disponível atualmente portanto não foi mapeada por não ter disponível o formato dos dados que serão gerados. Além disso, um novo passo foi identificado no processo de criação dos relatórios A, B, C e D: existe um passo de "Garantia da Qualidade dos Dados", que foi inserido antes da criação dos relatórios. Este passo é realizado para eliminar os dados duplicados de vendas extraídos do STEAM BOAT e também para acrescentar as colunas "Existing Places, New Places, Existing Stuffs e New Stuffs", utilizadas na criação do relatório D.

Estas correções no processo 1 causaram modificação no diagrama definido inicialmente para este processo e estão representadas na Figura 5-18.

Na Figura 5-17, os termos "Existing Places", "New Places", "Existing Stuffs", "New Stuff" e "Revenue Weathervane" foram mantidos em inglês por opção do próprio usuário em adotar um vocabulário utilizado pela DBMTech.

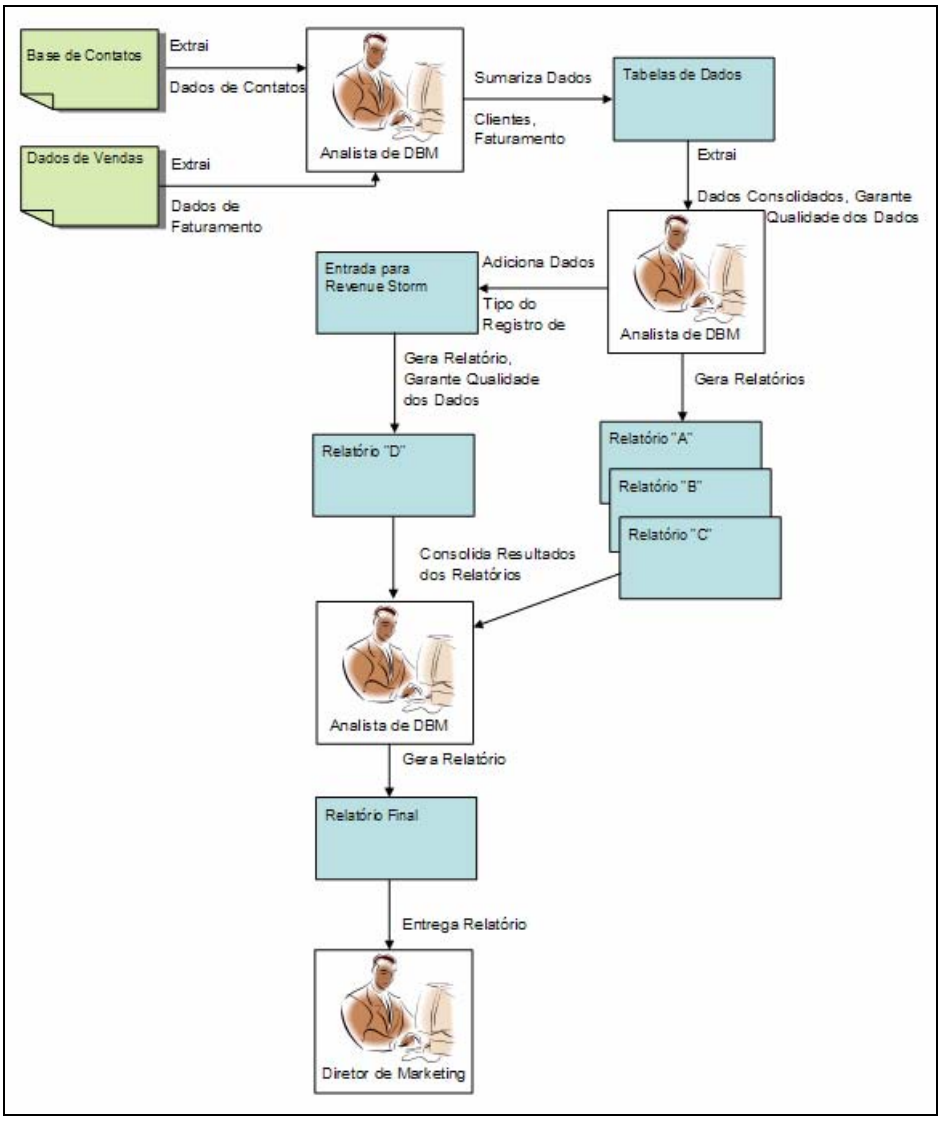


Figura 5-18: Processo 1 – redesenhado com base nas alterações identificadas na terceira oficina de trabalho.

A quarta oficina de trabalho foi iniciada com o término da análise do Processo 1, para definir as informações necessárias na produção do "Relatório Final – Documento Word".

Relatório Final: Resumo dos Resultados do Trimestre

Atributos identificados:

- Faturamento por setor
- Valor da Venda por Trimestre
- Valor de Novas Vendas por Trimestre

- Faturamento Anual por Cliente

- Valor da Renovação por Trimestre

BASE DE CLIENTES - REGIÃO

➢ **TOTAL DE RECEITA DE SOFTWARE (TSR)**

TSR	TRIMESTRE 1 - 2004	TRIMESTRE 1 - 2005	TAXA CRESCIMENTO
NOVO	R$ 100.00	R$ 200.00	50%
RENOVAÇÃO	R$ 150.00	R$ 250.00	60%
TOTAL	**R$ 250.00**	**R$ 450.00**	**56%**

➢ **TOTAL DE RECEITA**

TIPO DE VENDA	TRIMESTRE 1 - 2004	TRIMESTRE 1 - 2005	TAXA CRESCIMENTO
SERVIÇOS	R$ 100.00	R$ 150.00	67%
TREINAMENTO	R$ 110.00	R$ 160.00	69%
LIVROS	R$ 120.00	R$ 170.00	71%
NOVO	R$ 130.00	R$ 180.00	72%
RENOVAÇÃO	R$ 140.00	R$ 190.00	74%
TOTAL	**R$ 600.00**	**R$ 850.00**	**71%**

➢ **LISTA DE CLIENTES (EMPRESA – SETOR – PRODUTO/SOLUÇÃO – VALOR)**

EMPRESA	SETOR	SOLUÇÃO	VALOR
CLIENTE 1	SETOR A	TOOLS	R$ 10.00
CLIENTE 2	SETOR B	TOOLS	R$ 20.00
CLIENTE 3	SETOR C	TOOLS	R$ 30.00
		TOTAL	**R$ 60.00**

Figura 5-19: Relatório Final – baseado em informações dos relatórios B,C e D.

➢ **Análise do Processo 2**

Relatório E: Lista de Contatos atualizada

Atributos identificados:

- Código do Contato

- Grupo de interesses pessoais

- Detalhe do interesse pessoal

- Classificação de Jornais

- Classificação de Revistas

- Classificação de Hobbies

- Última data de atualização
- Tipo do Contato (Decisor/User)
- Nome do Contato
- Estado Civil
- Sexo
- Posição DBMTech (Analista/Assessor/Consultor/Assistente/Diretor/Gerente/Presidente/Proprietário)
- Área DBMTech
- Data de Aniversário
- Nível de Contato (High/Medium/Level)
- Cidade
- Estado
- Tem Filhos (Sim/Não)
- Estudante/Professor (Estudante/Professor/Ambos)
- Participa de Entidade de Classe (Sim/Não)
- Conhece a DBMTech (Sim/Não)
- Receptivo a e-mail (Sim/Não)
- Receptivo a mala direta (Sim/Não)
- Assinante da DBMTech.COM (Sim/Não)
- Que jornal lê (lista de jornais que lê)
- Que revista lê (lista de revistas que lê)
- Status de Atualização (Correspondência devolvida/E-mail devolvido/E-mail validado/Sem e-mail/Inválido/Provavelmente inválido/Válido)

A Tabela 5-9 é uma amostra dos dados de contatos existentes atualmente e a Tabela 5-10 são os eventos dos quais cada contato participou.

Tabela 5-9: Lista Geral de Contatos

AÇÃO TELEMARKETING	CORE/SME	CLUB VIP	CÓD. SFA	STATUS	REF4	SEGMENTO	RAZÃO SOCIAL	COD. CONTATO	TIPO
	CORE		2040	CLIENTE	200	SETOR A	EMPRESA 1		H
	CORE		2040	CLIENTE	200	SETOR A	EMPRESA 1		H
	CORE		2040	CLIENTE	200	SETOR A	EMPRESA 1		L
NÃO CONTATAR	CORE		39657	PROSPECT		SETOR A	EMPRESA 2		H
	SME	NÃO	407	CLIENTE	100	SETOR A	EMPRESA 3		L
	SME	NÃO	407	CLIENTE	100	SETOR A	EMPRESA 3		L
	SME	NÃO	407	CLIENTE	100	SETOR A	EMPRESA 3		L
LIGAR NOVAMENTE	SME		23098	PROSPECT		SETOR A	EMPRESA 4		H
	SME		23098	PROSPECT		SETOR A	EMPRESA 4		H

Tabela 5-10: Relatório "Ações de Marketing" – Lista

SEGMENTO	TIPO CONTA	CONTATO	CARGO	TIPO AÇÃO	AÇÃO	DATA
SETOR A	CORE	CONTATO 1	CLevel	EVENTO	ENCONTRO SETOR A	Q1/2004
SETOR A	CORE	CONTATO 2	Superindente	EVENTO	ENCONTRO SETOR A	Q1/2005
SETOR A	CORE	CONTATO 3	Diretor	EVENTO	ENCONTRO SETOR A	Q1/2006
SETOR A	CORE	CONTATO 4	Gerente	EVENTO	ENCONTRO SETOR A	Q1/2007
SETOR A	SME	CONTATO 5	Analista	EVENTO	ENCONTRO SETOR A	Q1/2008
SETOR A	SME	CONTATO 6	Assistente	EVENTO	ENCONTRO SETOR A	Q1/2009
SETOR A	SME	CONTATO 7	Diretor	EVENTO	ENCONTRO SETOR A	Q1/2010
SETOR A	SME	CONTATO 8	Gerente	EVENTO	ENCONTRO SETOR A	Q1/2011
SETOR A	SME	CONTATO 9	Analista	EVENTO	ENCONTRO SETOR A	Q1/2012

Tabela 5-11: Relatório E - Contatos por Atividade

EMPRESA	AÇÃO						
	EVENTO 1	EVENTO 2	EVENTO 3	EVENTO 4	EVENTO 5	EVENTO 6	EVENTO 7
Empresa 1	.	.	10	.	.	.	1
Empresa 2	.	.	19	.	.	.	.
Empresa 3	1	24	193	7	.	.	.
Empresa 4	.	17	174	9	2	.	.
Empresa 5	.	.	24	5	.	.	.

Ao avaliar o Processo 2 foi notada a ausência do relatório representado na Tabela 5-9, fato que gerou atualização do Processo 2 com a inserção deste novo relatório. A avaliação de cada um dos resultados produzidos neste processo levou ao entendimento e

As informações disponíveis nas tabelas são fictícias para que não fossem divulgadas informações confidenciais da DBMTech neste trabalho. Apesar disso, utilizamos as tabelas com dados reais durante as oficinas de trabalho.

definição de mudanças num dos sistemas de origem, o "Controle de Contato". Foi notada a ausência de algumas informações de perfil dos contatos, como "quais os Hobbies preferidos", "quais os jornais lidos", "quais as revistas lidas" e não somente se o cliente tem hobby, lê jornal, lê revista.

- **Análise do Processo 3**

Relatório F: Custos por Atividade de Marketing e Análise de Participação dos Convidados

Atributos identificados:

- Área que gerou o custo
- Sub-Áreas geradoras do custo
- Código da ação
- Detalhe do custo
- Descrição do custo
- Mês do custo
- Ano do custo
- Nome do fornecedor
- Valor do custo
- Quitado (Sim/Não)
- Código do fornecedor
- Tipo do serviço prestado
- Canal da Ação (E-mail *marketing*/Evento/Mala Direta/TeleMarketing)
- Tipo da Ação
- Código da Ação
- Data de Início da Ação
- Data Fim da Ação
- Nome da Ação de Marketing

Tabela 5-12: Relatório F - Custos de Marketing por Ação

Evento/Atividade	Quarter	Mês	Descrição	Fornecedor	Valor
Evento 1	Q1	2	Reserva de Sala	Fornecedor 1	R$ 10.00
Evento 2	Q1	2	Transporte	Fornecedor 2	R$ 20.00
Evento 3	Q1	2	Coordenação do Evento	Fornecedor 3	R$ 30.00
Atividade 1	Q2	8	Celular	Fornecedor 4	R$ 40.00
Atividade 2	Q2	8	Cartão de Crédito	Fornecedor 5	R$ 50.00
Evento 6	Q2	8	Patrocínio	Fornecedor 6	R$ 60.00

➢ **AÇÕES DE MARKETING X PARTICIPAÇÃO**

MÊS	AÇÃO	CAMPANHA	CONVID.	CONFIR.	PARTIC.	% PART.
JAN	Evento 1	Campanha 1	---	---	---	---
JAN	Material 1	Campanha 1	1	1	1	100%
JAN	Material 2	Campanha 2	300	---	300	100%
FEV	Evento 2	Campanha 2	1	1	1	100%
FEV	Evento 3	Campanha 3	183	13	4	2%
MAR	Evento 4	Campanha 3	---	---	---	---
MAR	Evento 5	Campanha 3	---	---	---	---

Figura 5-20: Relatório F - Participação em Ações de Marketing.

Após a avaliação do Processo 3, alguns atributos foram identificados e foi solicitada a sua inclusão nos dados de origem para que fossem utilizados nos relatórios de análise.

Ao final da quarta oficina de trabalho foi realizada uma revisão geral dos resultados obtidos durante os trabalhos em todas as oficinas de trabalho e algumas revisões foram realizadas no diagrama resultante dos trabalhos, ilustra a Figura 5-21.

Figura 5-21: Diagrama resultado dos trabalhos com PICTIVE, fonte para o StarSchema.

Ao avaliar o nome dado para a entidade "Atividade", foi notado que não era o nome adequado; esta modificação foi realizada e foi representada no modelo de dados estrela como entidade "Ação". Esta mudança causou a necessidade de reavaliar todos os atributos existentes para cada entidade que estava relacionada à "Ação", a fim de manter o documento consistente. O artefato produzido como resultado da aplicação do PICTIVE, permitiu que, de forma visual e simples, estes atributos fossem identificados e modificados.

Um conjunto de 9 entidades foram definidas: cliente, dados cadastrais, perfil do contato, ação, custos por atividade, custos (financeiro), faturamento, fornecedor, lista de participante, representadas no diagrama da Figura 5-21. Este diagrama ainda não possui um formato formal de modelo de dados, mas possibilitou a discussão até que o resultado final deste trabalho fosse desenvolvido, o *StarSchema*.

Estas 9 entidades foram reavaliadas e algumas modificações foram aplicadas porque foram notados os seguintes aspectos:

- a entidade dados cadastrais possuía informações de clientes, que são as empresas clientes da DBMTech assim como informações dos seus funcionários que possuem algum tipo de relacionamento com a DBMTech. Foi necessária a divisão desta entidade em 2 distintas: “Clientes” e “Contatos”, o que permite uma análise das informações de clientes de forma única, pois da forma como estava definido em uma única entidade o mesmo atributo de cliente poderia ter valores distintos para cada contato existente, gerando inconsistência nos dados.

- a entidade custos (financeiro) não era necessária nas análises que serão realizadas com os dados gerados no DBM. Os dados provenientes desta entidade serão utilizados para validação dos dados no processo de carga dos dados da entidade custos, porém não precisam estar disponíveis para consulta.

- a entidade custos por atividade possuía informações cadastrais como nome, data, das atividades realizadas assim como valores dos seus custos. Os atributos com estas características já haviam sido definidos em outras 2 entidades. As informações cadastrais passaram a ser atributos da entidade “Ação” e as informações de custos passaram a ser atributos da entidade “Custos”.

Na entidade “perfil do contato”, a reavaliação permitiu que a falta de atributos referente à informações de interesse pessoal fosse percebida e inserida como novo atributo nesta entidade. Além disso, foi gerada uma solicitação aos desenvolvedores do sistema origem dos dados para que informações novas identificadas durante as oficinas de trabalho sejam incorporadas ao sistema transacional que fornece as informações.

A entidade lista de participantes, que representava todos os contatos que já participaram de alguma ação, terá seus atributos incorporados na entidade “Ação”. Esta definição teve base no fato de uma ação sempre ter um participante.

A entidade custos sempre possui um fornecedor associado que é o gerador do custo. Para que análises de volume de custos gerado por fornecedor possam ser feitas e entender o perfil dos fornecedores, foi criada a entidade “Fornecedor”.

Com esta revisão realizada encerramos a versão final do modelo conceitual com apenas 7 entidades: Cliente, Faturamento, Ação, Custos, Fornecedor, Contato e Perfil do Contato.

Ao finalizar estas revisões, todas as modificações necessárias estavam representadas no diagrama do modelo de dados, o que agilizou o trabalho de desenho do modelo conceitual por meio de um sistema computacional.

5.4 Processos e Modelo de Dados proposto (*StarSchema*) para a DBMTech

Ao final dos trabalhos realizados por meio das práticas participativas, 3 processos foram definidos para que os relatórios e análises criados a partir dos dados do DBM fossem mapeados também foram identificadas as pessoas que atuam no processo, seja fornecendo os dados de origem ou ainda utilizando os dados resultantes para obter os indicadores necessários para a área de marketing da DBMTech. As figuras Figura **5-22**, Figura **5-23** e Figura **5-24**, representam os processos desenhados e ajustados ao final das oficinas de trabalho realizadas.

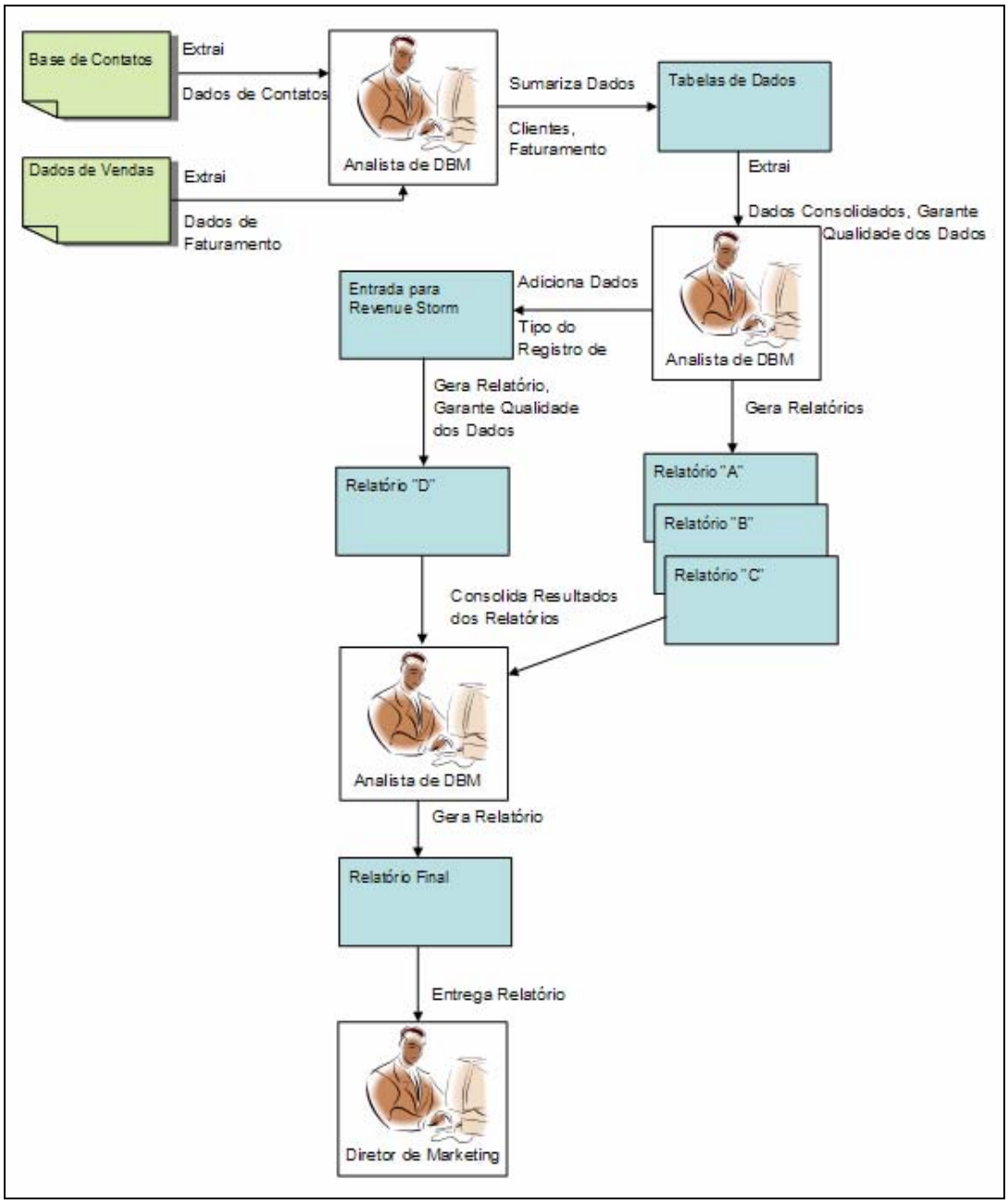


Figura 5-22: Processo 1 - Geração dos Relatórios de Análise de Receita por Cliente.

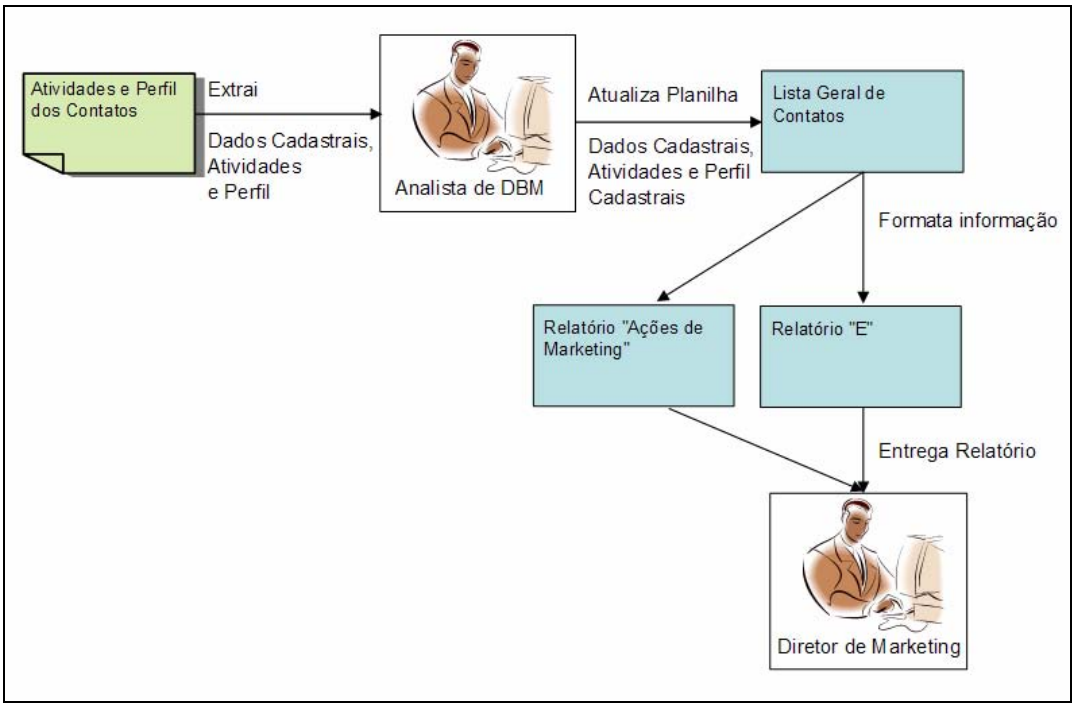


Figura 5-23: Processo 2 - Geração da lista de contatos atualizados baseado nas últimas atividades em que cada contato participou.

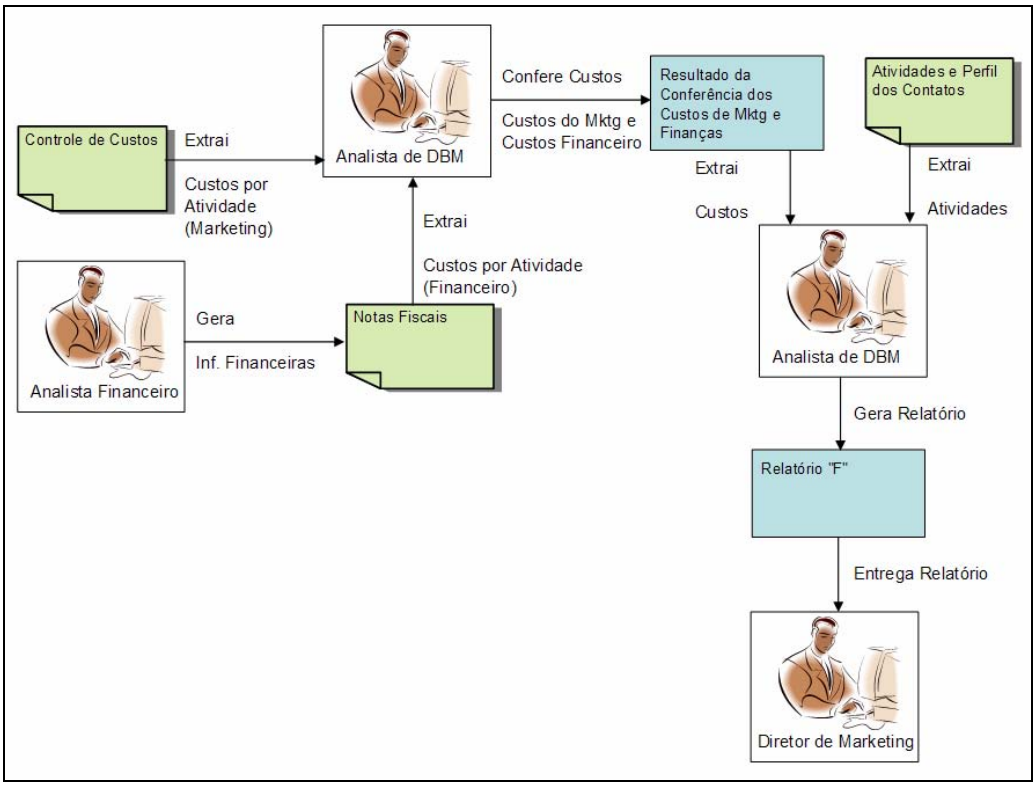


Figura 5-24: Processo 3 - Geração do Relatório de Custos por Atividade de Marketing.

Como resultado do trabalho de análise dos processos 1, 2 e 3 realizado com a DBMTech foi definido o modelo de dados conceitual. Este modelo de dados foi criado ao longo das oficinas de trabalho realizadas utilizando as práticas participativas PICTIVE e *Blueprint Mapping*. O modelo está representado por 2 diagramas, um *Star Schem*a que representa a tabela fato de custos e suas dimensões, na Figura 5-25, e outro que representa a tabela fato de faturamento e sua dimensão, na Figura 5-26.

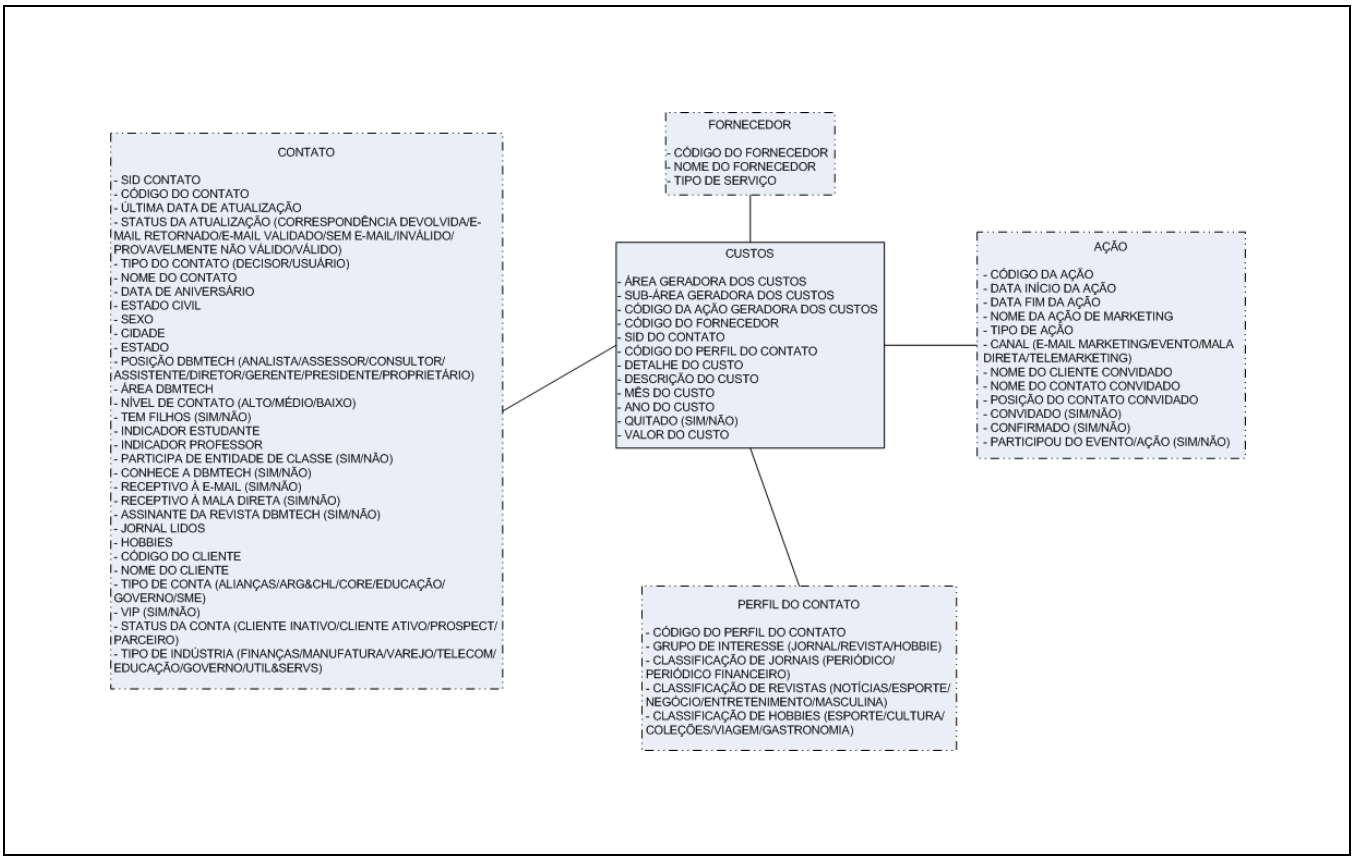


Figura 5-25: Modelo de Dados Conceitual (*StarSchema*) de Custos.

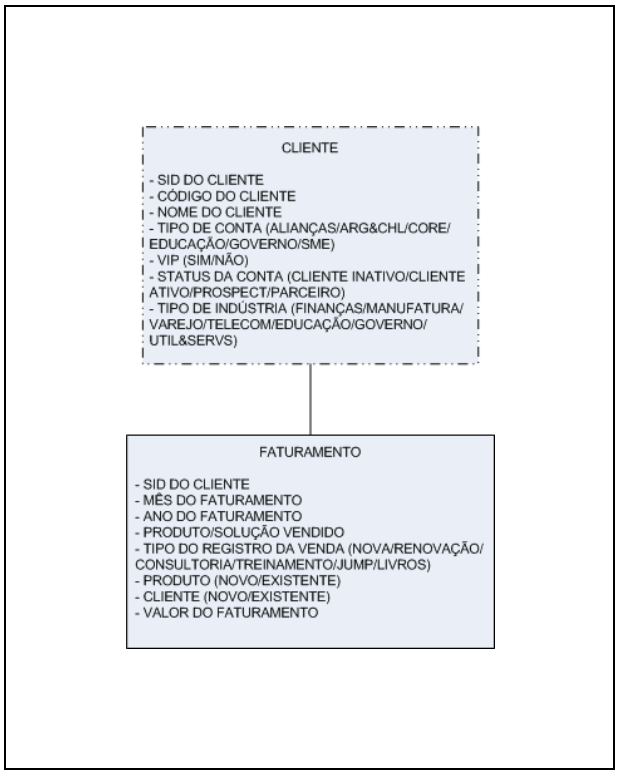


Figura 5-26: Modelo de Dados Conceitual (*StarSchema*) de Faturamento.

6 Discussão, Conclusão e Trabalhos Futuros

Os resultados obtidos após a realização dos trabalhos propostos com o uso de práticas participativas na elicitação de requisitos serão comentados neste capítulo sob os aspectos positivos e negativos, ressaltando quais foram os pontos fracos e fortes de cada prática participativa utilizada.

De maneira geral, foi possível obter uma melhora no processo tradicional de elicitação de requisitos onde são utilizados questionários e entrevistas. Por meio das práticas participativas foi melhorada a participação e envolvimento dos usuários na definição dos requisitos necessários para a implementação do DBM, consequentemente minimizando os erros que pudessem ser propagados. As práticas participativas aplicadas com o uso de artefatos distintos estimulam os usuários a contribuir com seus conhecimentos; apesar disso se o grupo de usuários envolvido for muito pequeno, o aproveitamento das técnicas participativas pode não ser pleno.

No caso da elicitação de requisitos para DBM, que é um banco de dados, adaptações foram realizadas nas técnicas participativas originalmente utilizadas para a definição de interfaces, para que o resultado esperado deste trabalho fosse atingido. Neste caso, além da elicitação de requisitos, o uso das práticas participativas foi expandido nas fases de análise, especificação e validação dos requisitos, o que permitiu criar um modelo conceitual de dados.

6.1 Discussão dos Resultados Obtidos

A construção de um DBM depende da decisão de uma organização em coletar e organizar dados de seus clientes. Esta coleta pode se dar pela realização de pesquisas junto aos clientes ou pela extração de informações de sistemas transacionais existentes numa organização. Cada uma destas formas de coleta exige uma força tarefa muito grande de todos os envolvidos no processo de criação do DBM para que os resultados apresentados sejam aqueles esperados pelos usuários em relação aos dados coletados e armazenados.

No contexto de desenvolvimento de DBMs em geral, durante a fase de elicitação de requisitos dos dados que devem ser armazenados e disponibilizados, um questionário que aborda

aspectos do negócio e aspectos de tecnologia é preenchido antes da realização de qualquer entrevista com os *stakeholders*. As entrevistas junto aos usuários são realizadas individualmente com cada *stakeholder* com base no questionário respondido, para validar e obter um entendimento claro dos dados que devem ser coletados e armazenados.

Ao utilizar esta abordagem muitas falhas na fase de elicitação de requisitos podem ocorrer, como por exemplo, ausência de informações necessárias para as análises sobre os clientes, ausência de conhecimento das análises desejadas e dificuldade de obtenção dos dados sobre os clientes. Tais falhas, que podem inviabilizar a construção do DBM, podem ser minimizadas por meio do uso de técnicas participativas onde os *stakeholders* são incentivados a interagir, discutir e concluír suas necessidades baseadas no consenso das diferentes visões sobre a ótica do DBM.

No estudo de caso utilizado para a aplicação das técnicas participativas, descrito neste trabalho, não houve uma mensuração dos resultados de forma quantitativa, comparando indicadores resultantes de métricas estabelecidas para obtenção destes resultados, mas houve a análise dos pontos fortes e fracos percebidos durante as oficinas de trabalho realizadas em relação aos objetivos propostos para cada um deles e pode-se verificar a influência destes pontos no resultado obtido.

6.1.1 Pontos Fortes no Uso das Práticas Participativas

Ao utilizar as práticas participativas na elicitação, análise e validação dos requisitos foram observados os seguintes pontos fortes:

- O uso do *Blueprint Mapping* permitiu organizar de forma rápida, clara e objetiva os artefatos disponíveis para o início dos trabalhos.
- Como os resultados são visuais, por exemplo, um diagrama, um quadro com figuras, imediatamente ao final de uma oficina de trabalho os resultados estão prontos e definidos pelos próprios usuários, garantindo o entendimento das suas necessidades.
- O uso de artefatos, por exemplo, um processo de criação de relatórios, produzidos pelos próprios usuários em conjunto com o analista de requisitos, facilitou a avaliação dos detalhes de cada passo durante a avaliação do processo atual e identificação dos dados e dos *stakeholders* envolvidos em cada processo.

- Ao avaliar cada processo com o objetivo de identificar as entidades e atributos de dados, o próprio usuário percebeu a ausência de certas informações e a ausência de certos *stakeholders*, o que contribuiu para reduzir falhas no desenho do modelo de dados apresentado como resultado final do trabalho.
- O desenho dos processos criados facilitará a identificação das fontes dos dados que serão carregados nas tabelas fato e dimensões por parte do analista de sistemas que fará a implementação do modelo de dados proposto.
- O desafio de utilizar as práticas participativas que são normalmente aplicadas ao desenvolvimento de interfaces de sistemas, foi completamente atendido com as adaptações criadas durante o desenvolvimento deste trabalho e demonstrou aderência às necessidades de um projeto de construção de um DBM.
- O fato do analista de requisitos conhecer a empresa DBMTech, facilitou no esclarecimento e definição dos processos e dados fonte para o DBM.

6.1.2 Pontos Fracos no Uso das Práticas Participativas

Ao utilizar as práticas participativas na elicitação, análise e validação dos requisitos foram observados os seguintes pontos fracos:

- As técnicas participativas podem contribuir de forma mais efetiva quando as oficinas de trabalho têm participação de um grupo maior de *stakeholders*, no caso deste trabalho tivemos somente um *stakeholder* participando ativamente de todas as oficinas de trabalho, porque a empresa é muito pequena e o DBM fica restrito ao uso desta pessoa envolvida.
- O analista de requisitos deve conhecer do negócio e das necessidades comuns para uma empresa que deseja construir um DBM, caso contrário não conseguirá mediar bem as sessões de análise e validação.
- A necessidade de trazer todos os envolvidos no projeto para as sessões que acontecem num único local, pode tornar o projeto inviável, por não conseguir reunir todos os interessados.
- Os *stakeholders* participantes das sessões devem conhecer a fundo suas necessidades e estarem dispostos a discutir para a definição de uma opinião que seja consenso do grupo; caso não exista este perfil, não é aconselhável o uso de técnicas participativas.

- Não existe um padrão para a documentação gerada como resultados dos trabalhos, portanto documentações de outras metodologias devem ser utilizadas na composição dos documentos gerados.

6.1.3 Trabalhos Futuros

Entre alguns aspectos que podem ser explorados em trabalhos futuros como continuidade ou aprimoramento deste trabalho, podem ser considerados os seguintes:

- Implementação do modelo de dados conceitual sugerido, por uma equipe de tecnologia.
- Aplicar a combinação de técnicas participativas propostas, em sessões de elicitação de requisitos de outros tipos de repositórios de dados, por exemplo, *Data Warehouse* e *Data Mart*.
- Expandir o uso das técnicas participativas em outras fases do projeto de implementação.
- Utilizar outros métodos participativos para compor uma metodologia melhorada em relação a que foi apresentada neste trabalho.
- Melhoria nas adaptações realizadas às técnicas participativas já propostas neste trabalho, por exemplo, criando um conjunto de materiais de apoio pré-formatados que possam ser utilizados prontamente em um novo projeto.

Este trabalho permitiu que a necessidade da DBMTech pudesse ser compreendida e definida para um equipe de implementação, além de demonstrar como uma abordagem diferenciada pode melhorar um processo existente e melhorar os resultados esperados desta fase de elicitação, que gera tanto desentendimento e falta de compreensão por parte dos envolvidos. Esta experiência contribuiu na melhora da abordagem utilizada atualmente na empresa e poderá ser expandida para outros projetos onde seja necessária a elicitação de requisitos para um banco de dados.

7 REFERÊNCIAS

[1] D'ALCÂNTARA, MÁRIO. **"Como começar pelo começo . Dicas muito importantes para se fazer DBM".** Disponível em: <http://www.globaldbm.com.br/empresa/dicas.htm>. Acesso em: 18/12/2003.

[2] HELANDER, MARTIN G., LANDAUER, THOMAS K., PRABHU, PRASAD V.. **Handbook of Human-Computer Interaction.** Second, Completely Revised Edition. Elsevier Science B.V., 1997.

[3] HOLTZ, HERMAN. ***Databased Marketing***. Coleção Eficácia Empresarial. Tradução Lenke Peres Alves de Araújo. São Paulo: Makron Books, 1994.

[4] LOPES, PAULO S. N. D. **Engenharia de Requisitos: Uma Visão Geral.** Seminário de Engenharia de Software, IME-USP, 1999. Disponível em: <http://www.naddeo.com.br/pdfFiles/Generico/SeminarioEngDeRequisitos.pdf>. Acesso em: 18/12/2003.

[5] NASH, EDWARD L. ***Database Marketing: Uma ferramenta atual e decisive do marketing***. Tradução Ana Terzi Giova. São Paulo: Makron Books, 1994.

[6] SHEPPARD ASSOCIATES,D. ***Database Marketing***: O novo *marketing* direto. Tradução Kátia Aparecida Roque. São Paulo: Makron Books, 1993. 347 p.

[7] 1KeyData.com. ***"Bill Inmon vs. Ralph Kimball"*** Disponível em: <http://www.1keydata.com/datawarehousing/inmon-kimball.html>. Acesso em: 16/06/2005.

[8] 1KeyData.com. ***"Conceptual, Logical, and Physical Data Models"*** Disponível em: <http://www.1keydata.com/datawarehousing/data-modeling-levels.html>. Acesso em: 16/06/2005.

[9] 1KeyData.com. ***"Dimensional Data Model".*** Disponível em: <http://www.1keydata.com/datawarehousing/dimensional.html>. Acesso em: 16/06/2005.

[10] 1KeyData.com. ***"MOLAP, ROLAP, and HOLAP".*** Disponível em: <http://www.1keydata.com/datawarehousing/ molap-rolap.html >. Acesso em: 16/06/2005.

[11] DREWEK, KATHERINE. ***"Data Warehouse: Bill Inmon's Vision".*** Disponível em: < http://www.b-eye-network.com/view/727 >. Publicado em: 23/03/2005. Acesso em: 16/06/2005.

[12] DREWEK, KATHERINE. ***"Data Warehouse: Ralph Kimball´s".*** Disponível em: < http://www.b-eye-network.com/view/713?filter_role=0>. Publicado em: 16/03/2005. Acesso em: 16/06/2005.

[13] ARVESON, PAUL. ***"What is the Balanced Scorecard?"***. Balanced Scorecard Institute, 1998. Disponível em: < http://www.balancedscorecard.org/basics/bsc1.html >. Acesso em: 24/06/2005.

[14] KIMBALL, RALPH. ***Data Warehouse Toolkit***. Tradução Monica Rosemberg. São Paulo: Makron Books, 1998.

[15] KIMBALL, RALPH. ***Preparing for Data Mining.*** DBMS On Line, *Data Warehouse* Architect. Publicado em: Novembro, 2005. Disponível em: <http://www.dbmsmag.com/9711d05.html#Figure1>. Acesso em: 16/06/2005.

[16] BERRY, MICHAEL, LINOFF, GORDON. ***Data Mining techniques: for marketing, sales, and customer support***. USA: Wiley Computer Publishing, 1997.

[17] INMON, W.H., WELCH, J.D., GLASSEY, K.L.. ***Managing the Data Warehouse***. USA: Wiley Computer Publishing, 1997.

[18] SOMERVILLE, I. ***Software Engineering.*** USA: Addison-Wesley, 1996.

[19] CARVALHO, A.M.B.R., CHIOSSI, T.C.S. ***Introdução à engenharia de software.*** Campinas, SP: Editora da Unicamp, 2001.

[20] ARAÚJO, JOÃO. Universidade Nova de Lisboa. ***"Processo de Engenharia de Requisitos"***. Notas de aula da disciplina de engenharia de software. Disponível em: <ctp.di.fct.unl.pt/~ja/mei-es/engreq.ppt>. Acesso em: 16/06/2005.

[21] GRAY, P., WATSON H.J. ***Present and future directions in data warehousing***. ACM SIGMIS Database, Volume 29 Issue 3. ACM Press, Junho 1998. Disponível em: The ACM Digital Library. Acesso em: 11/05/2006.

[22] LYON, J.K. ***What is a database***. ACM SIGMOD Record, Volume 5 Issue 1 Pages 44-62. ACM Press, Junho 1973. Disponível em: The ACM Digital Library. Acesso em: 11/05/2006.

[23] SEN, A., SINHA A.P. ***Reconsidering Multi-Dimensional schemas.*** Communications of the ACM, Volume 48, Issue 3, Pages 79-84. ACM Press, Março 2005. Disponível em: The ACM Digital Library. Acesso em: 11/05/2006.

[24] MARTYN, T. ***A comparison of data warehousing methodologies.*** ACM SIGMOD Record, Volume 33, Issue 1, Pages 83-88. ACM Press, New York, USA, Março 2004. Disponível em: The ACM Digital Library. Acesso em: 16/05/2006.

[25] INMON, W.H. ***Building the Data Warehouse***. John Wiley, 1992.

[26] VIEIRA, M.T.P., UFSCAR. ***"Data Warehouse"***. Apostila sobre Data Warehouse. Publicada em Novembro, 1999, revisada em Dezembro, 2001.

[27] MULLER, M.J. ***"Participatory Design: The Third Space in HCI"***. IBM Watson Research Center in Cambridge Julho, 2001. Disponível em: <http://domino.research.ibm.com/ cambridge/research.nsf/pages/index.html> Acesso em: 31/03/2005.

[28] GAFFNEY, G.. ***"Usability Techniques Series: Participatory Design Workshops"***. InfoDesign, Austrália, 1999. Disponível em: < http://www.infodesign.com.au/ ftp/ParticipatoryDesign.pdf > Acesso em: 01/03/2005.

[29] MULLER, M.J. ***"Catalogue of Scenario-Based Methods and Methodologies"***. IBM Watson Research Center in Cambridge, 1999. Disponível em: <http://domino.research.ibm.com/ cambridge/research.nsf/pages/index.html> Acesso em: 24/05/2006.

[30] OLIVEIRA, E.S.R.B. ***"Portaria nº 5, de 27 de agosto de 2002"*** . Ministério da Justiça, 2002. Disponível em: < http://www.mj.gov.br/dpdc/servicos/legislacao/pdf/Portaria5DPDC-Cl%C3%A1usulas%20Abusivas.pdf >. Acesso em: 31/07/2006.

8 Anexo I

- Documento apresentado pela Analista de DBM

Database Marketing

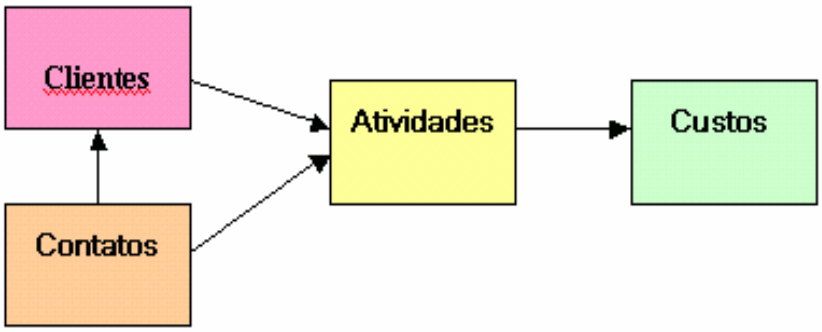
Situação Atual:

Contatos: Cada Gerente de contas possui sua lista de contatos, isto dificulta a centralização das informações e por sua vez a utilização dos contatos de forma correta.
Clientes: Não existe um local único para se entender a base e retirar informações reais, como quantidade de clientes atual, principais clientes de financial, etc.
Atividades: Não existe um controle de nossas ações, esta falta de controle dificulta o planejamento dos próximos anos e as análises de retorno.
Custos: Todos os custos das ações de marketing não são controlados dificultando a projeção do budget e a identificação dos principais gastos.

Objetivos:

Centralizar as informações do marketing de forma organizada para auxiliar em todas as nossas ações. Informações a centralizar:

1 - Contatos
2 - Clientes



3 - Atividades

4 - Custos

Processos:

Clientes

Guide

HTML

Análise de Cliente/Receitas

- Evolução da Receita SAS 2001 a 2004
- Customers X Segmento 2001 a 2004
- Customers X Prod/Solução 2001 a 2004
- Revenue Weathervane 2001 a 2004
- Faturamento Acumulado X Cliente 2001 a 2004
- Top 10 Clientes 2001 a 2004

Atualização por Quarter...

- Customers X Segmento 2005
- Customers X Prod/Solução 2005
- Revenue Weathervane 2005
- Lista de Clientes Atual x Quarter 2005
- Revenue X Quarter 2005
- Evolução do Revenue 2001 a 2005

Contatos

Aplicativo
SFA

Controle de Contatos

BZL

1º Limpeza da Base

2º Alterações no Aplicativo

3º Cruzamento da Base Limpa com Base SFA

4º Upload no SFA

5º Upload das últimas listas de eventos no SFA

6º Criação de planilha sobre Perfil dos Contatos

ARG/CHL

1º Recebimento da Base limpa de Arg/Chile

2º Criação do Aplicativo de Contatos para Arg/Chile

3º Cruzamento da Base limpa com Base SFA

4º Upload no SFA

5º Upload das últimas listas de eventos no SFA

6º Criação de planilha sobre perfil dos contatos

Atividades

SIM

Controle de Atividades

BZL

1º Listar todas as atividades do Ano X Quarter

2º Cada Quarter Preencher o SIM com as atividades do Quarter

3º Mensalmente enviar um e-mail update sobre o andamento das atividades do Quarter

ARG/CHL

1º Pegar com Tati as atividades do Ano

2º Listar todas as atividades do Ano X Quarter

3º Cada Quarter Preencher o SIM com as atividades do Quarter

4º Mensalmente enviar um e-mail update sobre o andamento das atividades do Quarter

Custos

Excel

Controle de Custos

1º Fazer novo comunicado a todos sobre processo de controle de custos do marketing
2º Cobrança dos Custos Antecipados das Ações
3º Recebimento – Cadastramento – Assinatura - Entrega
4º Comparativo mensal com os relatórios do Garcia

Resultados dos Processos:

- Update do HTML com dados sobre a base de clientes – Por Quarter
- Update dos Contatos atingidos com nossas últimas ações – Por Quarter
- Update das nossas últimas atividades, com nº de convidados, confirmados, participantes, pesquisa de satisfação – Por Quarter
- Update dos nossos custos de marketing – Por Mês
- Relatório final nossa base de clientes – Anual
- Relatório final dos contatos atingidos – Anual
- Relatório final das nossas últimas atividadades x nº convidados, confirmados, participantes x custo - Anual
- Relatório final de custos de marketing – Anual