

**Artefatos da Semiótica Organizacional na
Elicitação de Requisitos para Soluções de Data
Warehouse**

João Marcos Bonadio de Faria

Trabalho Final de Mestrado Profissional

Artefatos da Semiótica Organizacional na Elicitação de Requisitos para Soluções de Data Warehouse

João Marcos Bonadio de Faria

Fevereiro de 2006

Banca Examinadora:

- **Profa. Dra. Maria Cecília Calani Baranauskas (Orientadora)**
Instituto de Computação, UNICAMP
- **Profa. Dra. Eliane Martins**
Instituto de Computação, UNICAMP
- **Prof. Dr. Rodrigo Bonacin**
Centro de Pesquisas Renato Archer, CENPRA.

F225a Faria, João Marcos Bonadio de
Artefatos da semiótica organizacional na elicitação de
requisitos para soluções de data warehouse / -- Campinas, [S.P. :s.n.],
2006.

Orientador : Maria Cecília Calani Baranauskas
Trabalho final (mestrado profissional) - Universidade Estadual de
Campinas, Instituto de Computação.

1. Semiótica. 2. Engenharia de software. Data warehousing. I.
Baranauskas, Maria Cecília. II. Universidade Estadual de Campinas.
Instituto de Computação. III. Título.

Artefatos da Semiótica Organizacional na Elicitação de Requisitos para Soluções de Data Warehouse

Este exemplar corresponde à redação final do Trabalho Final devidamente corrigida e defendida por *João Marcos Bonadio de Faria* e aprovada pela Banca Examinadora.

Campinas, Fevereiro 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 22 de fevereiro de 2006, pela Banca Examinadora composta pelos Professores Doutores:



Profa. Dra. Eliane Martins
IC - UNICAMP



Prof. Dr. Rodrigo Bonacin
CenPRA



Profa. Dra. Maria Cecília Calani Baranauskas
IC - UNICAMP

© João Marcos Bonadio de Faria, 2006
Todos os direitos reservados

À Jacy

“Contos de Fadas são mais que verdade: não porque nos dizem que dragões existem, mas porque eles nos ensinam que dragões podem ser vencidos”.

G.K. Chesterton.

Resumo

Este trabalho tem como objetivo propor o uso das ferramentas e métodos da Semiótica Organizacional para a elicitação de requisitos de uma solução de Data Warehouse. As principais motivações para o trabalho advêm da falta de uma metodologia de elicitação de requisitos que apoiem de maneira efetiva todas as necessidades particulares para a elicitação de requisitos de Data Warehouse. Os problemas apontados pela literatura são explorados, bem como algumas propostas de metodologia apresentadas recentemente. Em particular, discutimos alguns dos aspectos que demonstram a necessidade de uma técnica de elicitação de requisitos desenvolvida especialmente para esse tipo de aplicação.

Explicamos nesse trabalho o que é um Data Warehouse e suas diferentes tecnologias, introduzimos o conceito de Semiótica Organizacional e apresentamos as ferramentas que são utilizadas durante o estudo de caso. Esse estudo de caso é descrito e os resultados apresentados, fornecendo a base para a proposta de uma maneira de uso das ferramentas da Semiótica Organizacional para a elicitação de requisitos. Ao final do trabalho pudemos ver que as ferramentas foram eficazes em seu propósito, inclusive com resultados além dos esperados, e propostas de trabalhos futuros são feitas.

Abstract

This work has the main goal of proposing the use of the tools and methods of Organizational Semiotics to elicit user requirements to a Data Warehouse solution. The main motivations for this work come from the lack of a proper methodology to elicit requirements that fully address the particular needs for a Data Warehouse solution. The issues discussed by previous works are analyzed; as well as some newly presented methodologies are discussed. Particularly we present some aspects that show the need of a new technique to elicit requirements tailored for Data Warehouse solutions.

We explain what is a Data Warehouse and its technologies, introduce the concept of Organizational Semiotics and present the tools used during the case study. The Case Study is described and the results shown giving the base to propose a method to use the Semiotic techniques to elicit the requirements for a Data Warehouse solution. With the results of the work we are able to understand that the Semiotic tools are quite efficient and the results were above expectation and finally some considerations for future works are made.

Agradecimentos

À Professora Maria Cecília Calani Baranauskas pela orientação e apoio na realização deste trabalho.

Ao Professor Pedro Rezende pelo apoio e motivação durante a graduação e mestrado.

Ao colega Carlos Alberto pela ajuda e apoio.

Ao Governo do Estado de São Paulo pela possibilidade do estudo de caso.

À minha família, em especial minha mãe e minha esposa, pela paciência e compreensão durante a confecção deste trabalho.

À todos aqueles que torceram por mim durante essa jornada.

Sumário

RESUMO.....	XIII
ABSTRACT.....	XV
AGRADECIMENTOS.....	XVII
SUMÁRIO.....	XIX
LISTA DE FIGURAS.....	XXI
LISTA DE TABELAS.....	XXIII
CAPÍTULO 1.....	1
INTRODUÇÃO.....	1
CAPÍTULO 2.....	7
A PROBLEMÁTICA DA ELICITAÇÃO DE REQUISITOS EM SOLUÇÕES DE DATA WAREHOUSE.....	7
2.1 CONCEITOS DE DATA WAREHOUSE.....	7
2.2 DESAFIOS DA ELICITAÇÃO DE REQUISITOS PARA DATA WAREHOUSE.....	17
2.3 REFERÊNCIAS DA LITERATURA.....	24
2.4 CONSIDERAÇÕES FINAIS.....	29
CAPÍTULO 3.....	33
OBJETIVO E METODOLOGIA.....	33
CAPÍTULO 4.....	39
SEMIÓTICA ORGANIZACIONAL COMO REFERENCIAL PARA A ELICITAÇÃO DE REQUISITOS.....	39
4.1 A SEMIÓTICA.....	39
4.2 A SEMIÓTICA ORGANIZACIONAL.....	43
4.3 O MÉTODO DE ARTICULAÇÃO DE PROBLEMAS (PAM).....	50
4.3.1 <i>Análise de organização e contexto</i>	50
4.3.2 <i>Análise da Morfologia Funcional</i>	53
4.3.3 <i>Análise Colateral</i>	54
4.4 CONSIDERAÇÕES FINAIS.....	56
CAPÍTULO 5.....	57
BUSCANDO UMA SOLUÇÃO PRELIMINAR DE DATA WAREHOUSE: UM ESTUDO DE CASO.....	57
5.1 PLANEJAMENTO DO USO DAS FERRAMENTAS DA SEMIÓTICA ORGANIZACIONAL.....	57
5.2 DESCREVENDO A ORGANIZAÇÃO E O CONTEXTO.....	68
5.2.1 <i>A definição da organização a ser estudada</i>	68
5.2.2 <i>A DIR I e o seu papel na área de Saúde</i>	71
5.3 O ESTUDO DE CASO.....	74
5.4 ANÁLISE DOS RESULTADOS E DESENHO DA SOLUÇÃO.....	96
5.5 CONSIDERAÇÕES SOBRE O RESULTADO.....	106

CAPÍTULO 6	111
CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS	111
REFERÊNCIAS BIBLIOGRÁFICAS	115
ANEXOS.....	121

Lista de Figuras

Figura 1	Cubos de dados de um Data Warehouse.....	9
Figura 2	Esquema simplificado de uma solução Data Warehouse (adaptada de Gorla (2003)).....	11
Figura 3	Dif. de implementação de tecnologias MOLAP e ROLAP (adaptado de Moreno & Mancuso(2003)).	13
Figura 4	Exemplo do uso de Data Marts (adaptado de (Paz, Cielo (1999-2000)))	15
Figura 5	Exemplo de um Star Schema (adaptado de Kimball (1998), p. 10).....	17
Figura 6	Diagrama de Etapas de Projeto de uma Solução de Data Warehouse (Shanks e O'Donnell (ND)).....	25
Figura 7	A interpretação de representamem por interpretantes diferentes	42
Figura 8	O Framework Semiótico (adapatado de Liu (2000, p. 27))	47
Figura 9	Camadas da Análise de Stakeholders	52
Figura 10	Representação da Análise de Morfologia Funcional	54
Figura 11	Ciclo da Análise Colateral. (Simoni, 2003, adaptado de Liu, 2000).....	56
Figura 12	Seqüência de Técnicas sugerida	59
Figura 13	Representação esquemática da solução final	60
Figura 14	Diagrama de uso dos resultados das técnicas empregadas	61
Figura 15	Aplicações representadas por <i>tabelas-fato</i>	66
Figura 16	Representação de uma aplicação e suas dimensões	67
Figura 17	Exemplos de Telas do Sistema SIAP	70
Figura 18	Exemplos de Telas do Sistema SIVISA	70
Figura 19	Resultado da Análise de Stakeholders	81
Figura 20	Sistemas fonte e Camada de ETL.....	98
Figura 21	A Aplicação elicitada	99
Figura 22	A aplicação e suas dimensões.....	102
Figura 23	As dimensões com chaves primárias e informações necessárias	104

Lista de Tabelas

Tabela 1 Diferenças entre os paradigmas Objetivista e Subjetivista. (adaptada de Liu (2000, p. 25))	45
Tabela 2 Resultados do Quadro de Avaliação - Camada de Contribuição	127
Tabela 3 Resultados do Quadro de Avaliação - Camada de Fonte	130
Tabela 4 Resultados do Quadro de Avaliação - Camada de Mercado	132
Tabela 5 Resultados do Quadro de Avaliação - Camada de Comunidade	133
Tabela 6 Resultados do Framework Semiótico	134
Tabela 7 Resultados da Análise Colateral	135

Capítulo 1

Introdução

O objetivo desse trabalho é descrever e caracterizar os problemas para a elicitação de requisitos na construção de um Data Warehouse e estudar a viabilidade do uso de técnicas de Semiótica Organizacional para apoiar a fase de elicitação de requisitos de uma solução de Data Warehouse.

Nos últimos anos houve um claro aumento da necessidade de informação da sociedade atual. Por meio de televisão, livros, revistas, Internet ou quaisquer outros meios de comunicação, as pessoas são apresentadas a milhares de informações ao mesmo tempo, seja ela requisitada ou não. A maneira como essa informação é passada muitas vezes é evasiva, e nem sempre é a informação desejada. Por causa desse acúmulo de informação, muitos dados inúteis podem se sobrepor a dados úteis e dificultar que eles sejam utilizados.

O valor da informação mudou. Ela é necessária em tempo real, no momento que ela é requisitada. O ciclo de informação tem se modificado e ficado menor. Notícias correm o mundo em segundos e tornam-se obsoletas em questão de horas.

Essa volatilidade da informação também afeta as empresas. O mercado cada vez mais competitivo não dá margens para manobras mal feitas. As decisões têm que ser tomadas imediatamente baseadas em dados confiáveis e atualizadas. Qualquer erro estratégico pode custar milhões em divisas e materiais. Estratégias de curto, médio e longo prazo tem de ser elaboradas e verificadas constantemente. Mudanças de cenário devem ser esperadas, verificadas e contornadas da maneira mais eficaz e transparente possível.

Com base nesse cenário, o conceito de Data Warehouse começou a ser desenvolvido. Sua principal função é criar um repositório de informações confiáveis, que sirvam de apoio para a tomada de decisão estratégicas e corporativas nas empresas. A qualidade dessa

informação, a maneira como ela é armazenada e disponibilizada, as diversas fontes de dados que carregam os bancos de dados, todos fazem parte do escopo de uma solução de Data Warehouse. As implementações de Data Warehouse, apesar de estarem cada vez mais comuns, ainda não se apoiam de forma consistente em ferramentas e técnicas criadas especificamente para esse tipo de solução. Esse panorama cria uma situação de risco, ainda mais levando-se em conta a quantidade de recursos e tempo necessários para finalizar a implementação total de um Data Warehouse. A etapa de elicitação de requisitos de um Data Warehouse ainda é muito pouco explorada pela literatura da área. Isso cria um grande risco para o sucesso do projeto, pois a percepção de qualidade e a confiança do usuário final vão determinar como a ferramenta será utilizada e como o Retorno de Investimento dos fundos gastos com a implementação será conseguido (Giannoccaro, Shanks, Darke (1999)). Se a ferramenta não puder dispor das informações requeridas pelos usuários de forma clara, de fácil acesso e com a qualidade esperada, a implementação do Data Warehouse não terá atingido seu objetivo.

Durante alguns anos de minha vida profissional, trabalhei em uma consultoria de informática de uma grande multinacional alemã, em seu escritório em São Paulo, SP. Essa consultoria provia serviços de implementação, integração de sistemas e consultoria de processos de uma grande gama de ferramentas, entre elas sistemas *Enterprise Resource Planning*, *Customer Relationship Management*, sistemas de Gerenciamento Eletrônico de Documentos , uma série de pequenos sistemas legados e outras soluções e processos de informática.

O grupo em que eu trabalhava era conhecido como Suporte ao Desenvolvimento e tinha como missão ser um primeiro nível de apoio para os consultores e desenvolvedores da empresa. Era nossa responsabilidade conhecer novas ferramentas, adquirir experiência com elas e repassar aos analistas novas tecnologias, ferramentas e processos e ajudá-los no uso desses novos conhecimentos.

Entre os anos de 2001 e 2004, estive bastante envolvido em implementações de ferramentas da empresa SAP, em especial o SAP R/3, conhecida ferramenta de ERP (*Enterprise Resource Planning*) e de toda a suíte conhecida como MySap, que continha toda a gama de ferramentas dessa empresa disponíveis. Embora tenha ficado focado nas soluções de CRM (*Customer Relationship Management*), sabia da existência de uma equipe que estava implementando a ferramenta da SAP conhecida como *Business Warehouse*. Essa ferramenta, em conjunto com uma outra conhecida como SAP SEM, era uma solução de *Balance Scorecard*, para apoio de decisões estratégicas. O BW, como era conhecido, era a versão da SAP para a tecnologia conhecida como *Data Warehouse*. Uma solução de *Balance Scorecard* é um tipo de ferramenta de apoio a decisões estratégicas baseada em metas e objetivos. Em geral, utiliza uma implementação de Data Warehouse como apoio.

No começo do projeto a equipe responsável pelo BW ficou um pouco isolada, sendo que o meu primeiro contato com a ferramenta foi uma apresentação ministrada pela líder da equipe. Algum tempo depois, como era normal em todas as implementações da consultoria, alguns problemas começaram a aparecer e eu fui escalado para entender o contexto, a forma de uso da ferramenta e ajudar em eventuais problemas da equipe. Nessa época comecei a tomar conhecimento dos principais problemas na área de Data Warehouse.

Esse período da minha vida foi de intensa atividade e aprendizado. Apesar de na época estar focado mais nos aspectos técnicos das ferramentas, as atividades de elicitação de requisitos e planejamento sempre me chamaram a atenção. Por causa dessas atividades eu notei a existência de alguns problemas muito específicos de projetos de desenvolvimento de soluções de Data Warehouse. Entre esses problemas estavam a falta de métodos e ferramentas específicos para apoiar o desenvolvimento das soluções de Data Warehouse.

Cada empresa de consultoria ou fabricante de software que apresente soluções para Data Warehouses geralmente cai no mesmo problema: elas tentam adequar os requisitos dos usuários em suas ferramentas e desse modo não garantem que a realidade do usuário será espelhada. Do mesmo modo a literatura sobre Data Warehouses não apresenta

metodologias novas, apesar de reconhecer a deficiência do uso de métodos clássicos de desenvolvimento no âmbito de Data Warehouses.

Especificamente no caso da análise de requisitos há o maior problema encontrado. Como o Data Warehouse é uma ferramenta para o apoio da análise estratégica de uma empresa, vários fatores como, por exemplo, os valores, missão e modos de operação dessa empresa têm de estar espelhados no Data Warehouse, mas em geral alguns desses fatores são de difícil explicação por parte dos usuários e muitas vezes há um certo receio dos usuários de contar a maneira como a análise estratégica, ponto de extrema importância nos negócios, é feita e avaliada. Somado a isso, temos a dificuldade de acesso às pessoas de alto escalão das empresas, geralmente quem mais aproveitaria da solução de Data Warehouse.

Dessa forma, a idéia principal deste trabalho é verificar se as ferramentas da Semiótica Organizacional e em especial os métodos apresentados na MEASUR podem servir como metodologia e ferramentas de apoio para projetos de soluções de Data Warehouse. O uso específico da Semiótica Organizacional foi sugerido devido a uma série de características dessas ferramentas que coincidiam com o que eu esperava que uma metodologia de apoio para a construção de Data Warehouse tivesse. Num primeiro momento o que mais me chamou a atenção foi o *Framework Semiótico*, em especial o sugerido por Liu (2000), pois seus seis degraus refletiam as etapas de desenvolvimento que eram consideradas importantes com base na minha experiência pessoal. Além disso, um conceito da Semiótica Organizacional chamava muito a minha atenção: o de subjetividade. Como podemos falar de apoio às estratégias se não entendemos as reais intenções de uma empresa, como podemos representar a linha de raciocínio de um executivo se não sabemos como ele encara a realidade e de que maneira ela a manipula para chegar aos resultados esperados? Além disso, como garantir que as pessoas responsáveis por interagir com os desenvolvedores realmente conheçam o mais importante, entendam as reais estratégias e objetivos da corporação?

A principal motivação desse trabalho saiu diretamente dessas dificuldades. Nosso principal objetivo era estudar as ferramentas da Semiótica e determinar se seria possível a partir delas

constituir uma metodologia adequada para a elicitação de requisitos de uma solução de Data Warehouse. A fase de elicitação de requisitos é de extrema importância em qualquer projeto, mas no desenvolvimento de um Data Warehouse ela se torna crítica. Os projetos de Data Warehouse são longos, caros e geram muita expectativa do usuário. Além disso, os requisitos não são claros nem mesmo aos usuários, pois o uso do Data Warehouse nem sempre segue regras específicas de acesso.

Para atingir o objetivo foi proposto um estudo de caso para o uso das técnicas da Semiótica Organizacional e verificação do comportamento dos usuários frente a essas técnicas. Com esse estudo de caso fomos capazes de definir uma metodologia de uso das ferramentas da Semiótica Organizacional, além de verificar se a empresa tinha condições para implementar uma solução do porte de um Data Warehouse.

Este trabalho está organizado da seguinte forma:

- No Capítulo 2 são exploradas as tecnologias de Data Warehouse, descritos problemas conhecidos para a elicitação de requisitos e o que a literatura da área apresenta como soluções e problemas;
- No Capítulo 3 o objetivo desse trabalho é apresentado, bem como os fatores de motivação e a metodologia empregada;
- No Capítulo 4 é introduzida a teoria da Semiótica Organizacional;
- No Capítulo 5 é desenvolvido o estudo de caso e a maneira como as técnicas da Semiótica Organizacional foram empregadas, bem como o desenvolvimento do modelo de uso dessas ferramentas;
- No Capítulo 6 estão as considerações finais e trabalhos futuros.

Capítulo 2

A Problemática da Elicitação de Requisitos em Soluções de Data Warehouse

2.1 Conceitos de Data Warehouse

Não é simples caracterizar o que é uma solução de Data Warehouse ou de sua real utilidade em uma empresa. Uma simples busca pela literatura fará que o leitor mais atento verifique que há uma grande diferença entre os conceitos e atribuições dessas soluções.

O conceito de Data Warehouse vigente é indicado na seguinte definição básica:

“Um Data Warehouse é um repositório de informações integradas, disponíveis para análises e consultas. Dados e informações são extraídos de fontes heterogêneas de onde elas são geradas. A principal vantagem é que se torna muito mais fácil e eficiente rodar consultas sobre esses dados que sobre os sistemas originais.”. Inmon (ND f).

Levando-se em consideração a definição anterior, notamos que a principal finalidade do Data Warehouse é prover um acesso mais direto às informações. Mas entende-se como acesso mais direto não apenas a maneira de se chegar aos dados, mas também a facilidade de se reconhecer inter-relações entre os dados de diversas fontes. A maneira como uma aplicação de Data Warehouse é desenhada pode facilitar em muito a tarefa de analistas de negócio, gerentes e técnicos que dependam de informação segura e precisa para as suas tomadas de decisão.

A qualidade dessas informações também é essencial. Tomadas de decisão baseadas em dados errados ou mal interpretados podem levar a resultados catastróficos para o objetivo da empresa. A qualidade das informações é relativa a cada organização e de sua interpretação dependem as diretrizes dos projetos de Data Warehouse. E porque a qualidade

de informação é tão importante? Porque, antes de tudo, o Data Warehouse é um sistema considerado *Data-Driven*, ou seja, sua função principal e motivo de existência é prover dados consistentes, de boa qualidade e de forma amigável a um usuário.

A subjetividade do conceito de qualidade de informação em um Data Warehouse pode ser exemplificada de forma simples. Tomemos como exemplo dois diferentes setores de uma empresa, que tenham diferentes funções, mas necessitem do mesmo tipo de dado, mas que façam uso diferenciado desses dados. Digamos que para o Setor A os dados sejam críticos, e que quando se deseja verificá-los tenha-se pouco tempo para se tomar uma decisão. O Setor B necessita dos mesmos dados, mas digamos que seja apenas um setor de catalogação e que se os dados demorarem algumas horas para ficarem prontos não causará maior impacto no seu processo produtivo. O tempo de acesso aos dados pode ser considerado como um aspecto de qualidade, mas se um analista apenas satisfizer os requisitos do setor B, para o setor A esses dados, e o seu método de acesso, faltarão com um aspecto muito importante de qualidade.

Apesar de haver uma grande gama de métodos de implementação de sistemas de Data Warehouse, a mais conhecida é utilizar-se uma ferramenta do tipo OLAP Server. OLAP significa “*OnLine Analytical Processing*”. A principal característica dessa tecnologia é o armazenamento de dados pré-processados, metaforicamente numa forma de cubo, onde o cruzamento de parâmetros de consulta proporciona o acesso à informação de forma rápida e consistente. Cada face do cubo representa uma dimensão. Essa dimensão pode ser, a grosso modo, comparada a uma tabela num banco de dados relacional. Obviamente, a figura geométrica de um cubo é apenas representativa, podendo o banco conter uma infinidade de dimensões. Os dados armazenados no banco estão posicionados nos pontos de interface entre as dimensões, ou seja, ao fazer uma consulta, o banco faz uma interseção perpendicular ao valor desejado de uma dimensão com a interseção de outra dimensão. A linha ou ponto gerado por essa consulta representa os valores desejados. Caso a mesma consulta fosse gerada em um banco de dados relacional, uma grande quantidade de tabelas seria necessária e a consulta deveria fazer acesso a várias delas, aumentando o tempo de processamento. Além disso, para preservar os relacionamentos entre essas dimensões, uma

grande quantidade de chaves estrangeiras seria necessária, fazendo com que a redundância de dados fosse enorme, o que provocaria um maior custo de processamento e de armazenagem. Levando-se em consideração que muitas vezes, como já dissemos, um usuário de um Data Warehouse não sabe exatamente o que está procurando e que a necessidade de extrapolação de dados desejada poderia não ter sido imaginada durante a confecção do banco de dados, a consulta desejada pode ser impossível, por uma simples falta de relacionamento entre as chaves estrangeiras das tabelas. Sendo assim, uma nova instrução para a consulta do banco de dados deveria ser gerada, tarefa que necessita de níveis de acesso ao servidor e conhecimentos técnicos acima dos encontrados na média dos usuários das empresas.

A Figura 1 representa um cubo de dados de um Data Warehouse.

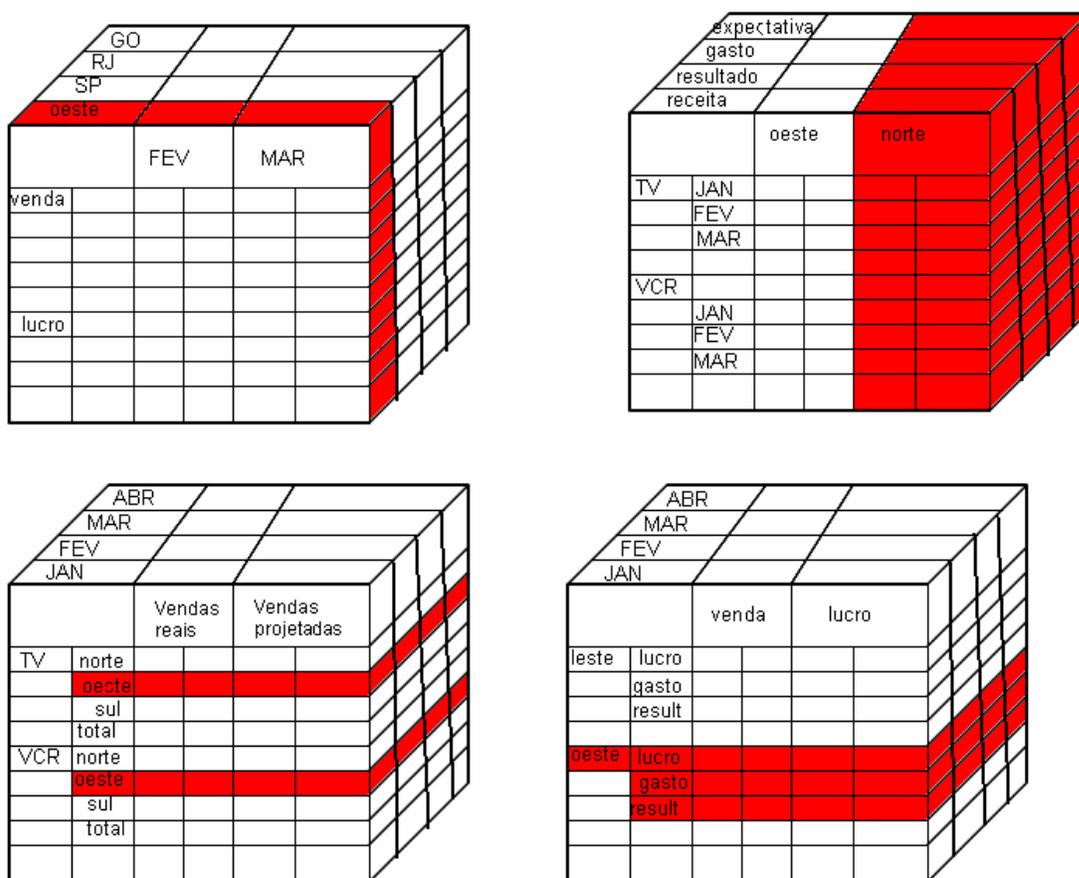


Figura 1 : Faces de um Cubo de dados de um Data Warehouse

A Figura 1 representa um “cubo” com cinco faces, cada uma delas uma dimensão demonstrando uma série de resultados de uma empresa fictícia. As dimensões apresentadas são tempo, resultado, produto, unidade de venda e orçamento. A área sombreada indica a interface entre um determinado valor de uma dimensão e a sua correlação com as outras dimensões. Caso desejemos saber qual foi a margem de lucro da unidade Oeste em janeiro, basta “hachurar” esses valores no cubo e os resultados estarão dentro da área de interseção. Além da facilidade em se fazer consultas nos cubos, uma outra vantagem do modelo OLAP é a facilidade com a qual se faz o chamado *Drill-Down*. O *Drill-Down*¹ é a busca de informações mais detalhadas de um determinado cruzamento, fazendo o caminho inverso das agregações. Os dados que estão na interseção de dimensões geralmente são valores agregados, ou seja, é a somatória de outros valores que compõem o valor final desejado. O Data Warehouse tem a capacidade de guardar vários níveis de detalhes. Geralmente só a agregação final já contém os dados necessários para a tomada de decisão, mas caso o usuário que esteja verificando os dados deseje, ele pode navegar pelos valores formadores, e verificar os dados num nível de detalhe maior. Por exemplo, um analista de produção sabe que a unidade Oeste obteve um resultado que mostra uma queda de vendas de televisores nos últimos meses. Para esse analista, pode ser necessário descobrir se a queda foi geral para todos os modelos ou se foi pontual em um. Navegando pela ferramenta, ele deve ser capaz de enxergar esses valores formadores e verificar, por exemplo, que um determinado modelo não tem mais a procura esperada e determinar a retirada do produto do mercado e substituí-lo por outro, economizando assim recursos de produção e podendo prever o tempo médio de vida de seus produtos.

A evolução das tecnologias OLAP resultou numa série de outros produtos que passam pelo mesmo conceito. Podemos exemplificar com sistemas que se utilizam de tecnologias MOLAP (*Multidimensional OLAP*) e ROLAP (*Relational OLAP*).

As ferramentas MOLAP são os OLAPs multidimensionais. Nelas, o próprio banco de dados, geralmente proprietário, armazena os dados em estruturas pré-definidas como os

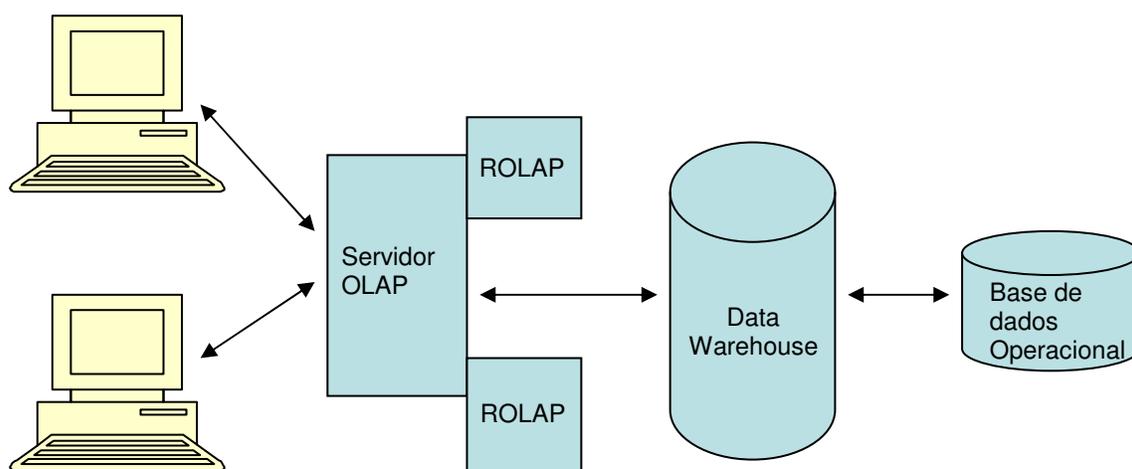
¹ O termo Drill-Down é utilizado pelos analistas de Data Warehouse e não tem uma tradução direta. O seu significado remete à ação de desdobrar o valor avaliado em seus valores componentes, navegando “para baixo” na pirâmide de agregação.

próprios cubos. Os dados são tratados e carregados periodicamente para os cubos, de modo a poderem ser acessados pelos usuários. O acesso a esses dados também é proprietário e a modelagem do cubo acontece de forma direta na ferramenta.

As ferramentas ROLAP utilizam recursos de bancos de dados convencionais para simular a criação dos cubos. Cada dimensão dos cubos é convertida num modelo de banco de dados adequado ao tipo de consulta necessária para os usuários.

Além das tecnologias MOLAP e ROLAP, há ainda o conceito de HOLAP (*Hybrid OLAP*), que resulta em ferramentas híbridas que tentam juntar as vantagens de um MOLAP com as de um ROLAP.

A Figura 2 exemplifica uma aplicação Data Warehouse simples.



Interface de usuários

Figura 2: Esquema simplificado de uma solução Data Warehouse (adaptada de Gorla (2003)).

Nesse exemplo, o usuário tem acesso direto ao OLAP Server, que se utiliza da tecnologia ROLAP para acesso ao Data Warehouse. O Data Warehouse, por sua vez, é carregado por uma base de dados operacional. Pelo fato do Data Warehouse utilizar-se de ROLAP, os dados estarão armazenados em tabelas relacionais dentro de um banco de dados de

mercado. A tecnologia a ser utilizada, seja ela a MOLAP ou a ROLAP, depende da maneira como o a aplicação será construída e também do que é esperado como resultado.

Podemos notar, por meio desse exemplo simples, que um erro comum é considerar-se como o Data Warehouse a ferramenta OLAP, por causa de sua representação de cubo, que é na verdade parte da interface da solução. O Data Warehouse real é o conjunto de dados agregados que podem ser utilizados para a carga das ferramentas de OLAP. Esses dados podem estar guardados em bancos relacionais, arquivos ou outra forma de armazenamento adequada. A importância nessa questão é que as informações contidas sejam as apropriadas para as cargas das dimensões e as mais úteis para os usuários. Bill Inmon, um dos criadores do conceito de Data Warehouse, explica o conceito fazendo analogia com um “balde de Legos”, em alusão ao popular brinquedo(Inmon, Tenderman, Imhoff (2003)). Os dados no Data Warehouse são os “bloquinhos de Lego” ou seja, unidades formadoras que podem ser utilizados para a construção de virtualmente qualquer coisa. Esses dados, recebidos das diversas fontes de dados, podem conter informações históricas, atualizações de produção e qualquer outra coisa. Nesse momento, as informações podem conter qualquer nível de agregação, mantendo-se o nível de detalhes desejado e as informações em estado mais “bruto” ou já com um tratamento prévio. Depende apenas das necessidades dos usuários e da solução como um todo para determinar qual a melhor característica de armazenamento e uso dessas informações.

A partir de uma solução OLAP, os dados do Data Warehouse são carregados para a mesma, podendo então ser manipuladas pelos usuários ou por uma grande gama de ferramentas de apoio à decisão.

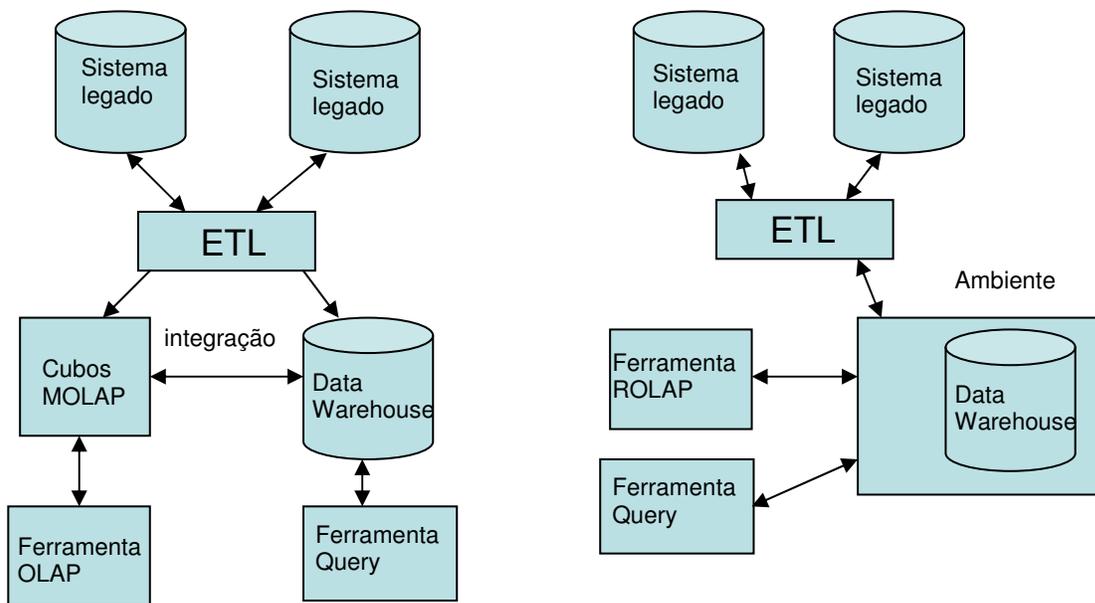


Figura 3 : Diferenças de implementação de tecnologias MOLAP e ROLAP (adaptado de Moreno & Mancuso(2003)).

A Figura 3 representa duas implementações típicas de Data Warehouse, uma utilizando ferramentas MOLAP e outra utilizando um ROLAP. Podemos verificar, nas duas implementações, as informações vindo de outros sistemas, como *Enterprise Resource Planning's* , ferramentas de *Customer Relationship Management* e outros sistemas legados. Para a busca de dados desses sistemas de carga no Data Warehouse, utiliza-se de ferramentas diversas na fase conhecida como Extração Transformação e Carga (*Load*) (ETL). Essa fase é uma das mais críticas para o Data Warehouse, pois a conexão do mesmo com os sistemas fonte implica em quais dados serão trazidos, quais os principais refinamentos que serão feitos nessas informações e como essa informação será disponibilizada.

A fase de extração implica no conhecimento de protocolos capazes de comunicar com todos os sistemas que servirão como fonte. Como podemos imaginar, a quantidade de tecnologias envolvidas é muito grande. Podemos ter sistemas que utilizam de banco de dados para guardar as informações, podemos ter sistemas baseados em arquivos texto, e uma infinidade de outros formatos. Além disso, geralmente as informações estão em servidores diferentes, com plataformas de tecnologias diferentes, como servidores com

sistemas operacionais baseados em UNIX ou Windows, podemos ter velhos sistemas baseados em Main Frames e, o que costuma ser mais problemático, podemos ter pequenos sistemas departamentais desenvolvidos especificamente para um pequeno grupo de usuários, pequenos programas que rodam em base local e não raramente baseados em planilhas eletrônicas. Como podemos ver, é uma tarefa árdua do ponto de vista de integração de sistemas. As tecnologias para integração de sistemas têm evoluído muito, como por exemplo, o uso do XML, e com isso essa tarefa têm sido facilitada. O ETL funciona como um tipo de *middleware*, tratando com os seus conectores as diferenças de tecnologia e disponibilizando a informação extraída em uma base homogênea.

A fase de transformação acontece por problemas similares aos da fase de extração. Como as tecnologias e sistemas diferem, muitas das informações relacionadas podem ser guardadas de formas e maneiras diferentes e podem, muitas vezes, nem utilizar as mesmas representações. Por exemplo, podemos supor que mesmo para informações simples, como o gênero de um cliente, pode estar guardado de diversas maneiras, desde o prosaico **M** para gênero masculino e **F** para o gênero feminino até um sistema de codificação mais complexo, como é utilizado no R/3, ferramenta de ERP da empresa alemã SAP. Desse modo, a camada de ETL deve homogeneizar as informações de modo que traduções não sejam mais necessárias (Inmon (ND b)). Obviamente esse trabalho depende de um estudo prévio do que será trazido dos sistemas fontes e o entendimento das estruturas de dados dos mesmos.

Por fim, o ETL faz a carga dos dados no Data Warehouse. Essa fase é responsável pela disponibilização final dos dados para uso direto ou para uso em uma ferramenta *OnLine Transactional Processing*, ou seja, ferramentas de uso de transações em tempo real.

Na Figura 3 podemos ver a diferença de comportamento entre a solução que usa o MOLAP e a que usa o ROLAP. Na solução MOLAP, a carga de dados é feita diretamente no Data Warehouse e nos cubos da ferramenta OLAP, de forma paralela. A integração entre essas dados depende de mais uma camada de processos. Essa estrutura faz com que a interface do OLAP e as ferramentas de *Queries* tenham que utilizar fontes diferentes. Já a

solução ROLAP cria um ambiente relacional que permite que as tanto as interfaces ROLAP quanto ferramentas de *Queries* acessem o mesmo ponto.

Uma característica interessante apresentada pelas soluções de Data Warehouse é o uso dos *Data Marts*. Um *Data Mart* é um subconjunto do Data Warehouse, que atende aos requisitos específicos de uma porção dos usuários (Inmomo (ND d)). Pensando no Data Warehouse como um repositório de dados de toda uma organização, um Data Mart seria uma construção para manter os dados que seriam mais úteis a um determinado nicho da empresa sem muita importância para os outros, como por exemplo o departamento financeiro. Essa divisão em Data Marts pode ter várias implicações e benefícios como, por exemplo, facilidade de definição das políticas de segurança, mais rapidez de navegação pelo usuário final e uso de informações mais refinadas dentro de contexto específico.

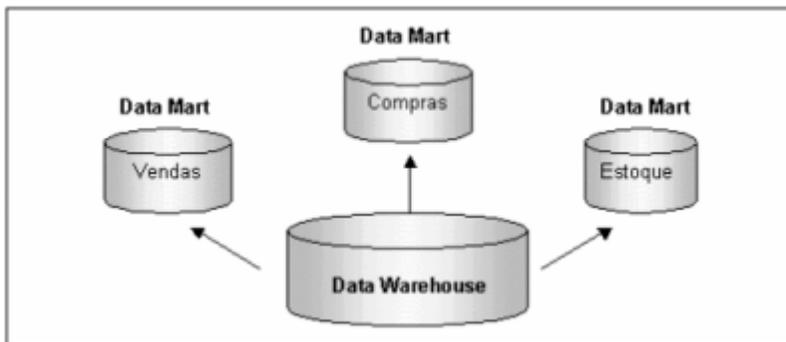


Figura 4 : Exemplo do uso de Data Marts (adaptado de (Paz, Cielo (1999-2000)))

As soluções de Data Warehouse têm como principal aplicação servir de base para sistemas de apoio a decisão. Os maiores interessados desse tipo de solução são grandes empresas que necessitam estudar e analisar grandes quantidades de informação a fim de traçar objetivos estratégicos e projeções futuras para os seus negócios. A maneira como o Data Warehouse é desenvolvido e quais as informações que ele deve conter dependem do tipo de negócio da empresa e da maneira como a organização trabalha seus dados e os transforma em projeções e objetivos.

Segundo Kimball (1998) uma empresa pode se beneficiar de um Data Warehouse para responder a uma série de questões sobre o negócio que uma ferramenta do tipo transacional não pode responder. Essas questões geralmente aparecem conforme a corporação entende que se a sua estratégia corporativa for bem apoiada por informações relevantes essa estratégia pode realmente alavancar os resultados da empresa. Mas a quantidade de informação existente na empresa dificulta um filtro razoável dessas informações, atrapalhando os analistas e por vezes escondendo as reais causas de um problema ou um fator de risco a ser considerado.

Desse modo, Kimball(1998) especifica seis objetivos principais de um Data Warehouse:

1. Fornecer acesso a dados corporativos ou organizacionais;
2. Manter os dados consistentes e confiáveis, segundo critérios da empresa;
3. Separar e combinar os dados de forma a facilitar qualquer visão possível do negócio;
4. Fornecer meios para consultar, analisar e apresentar informações;
5. Garantir a publicação de dados confiáveis;
6. Garantir qualidade de dados a fim de apoiar uma reengenharia de negócios.

Assim, podemos dizer que uma empresa que queira investir em uma solução de Data Warehouse está visando um melhor apoio de negócios baseado em informações tratadas e confiáveis. Para tanto, essa empresa deve ter um bom grau de maturidade tecnológica e de processos.

Para este trabalho, utilizaremos a representação tradicional de um Data Warehouse dimensional. O modelo dimensional de um Data Warehouse pode ser representado por meio do diagrama conhecido como Star Schema. Esse diagrama contém uma tabela principal, conhecida como *tabela-fato*, que contém todos os dados representativos da aplicação. Essa *tabela-fato* está associada com uma série de tabelas menores que

representam as *dimensões* do cubo de um Data Warehouse. A figura 5 exemplifica uma aplicação simples, com sua *tabela-fato* e suas *dimensões*.

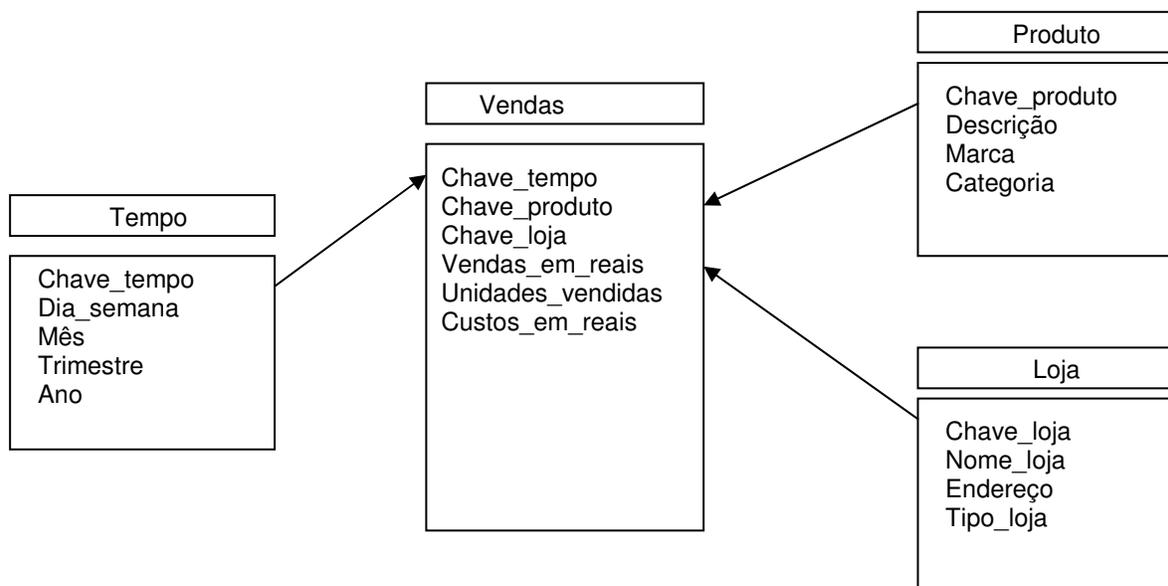


Figura 5 : Exemplo de um Star Schema (adaptado de Kimball (1998), p. 10)

2.2 Desafios da Elicitação de Requisitos para Data Warehouse

Uma solução baseada em um Data Warehouse difere de um sistema computacional padrão por uma série de fatores ligados diretamente à sua função e constituição. Um Data Warehouse contém informações transformadas e tratadas de uma empresa buscando melhor visibilidade de informações e a possibilidade de se tratar desses dados de maneira mais rápida e versátil.

Shanks e O'Donnell (ND) descrevem o principal problema com a elicitação de requisitos de um Data Warehouse da seguinte forma:

“Análise de requisitos para um Data Warehouse é problemático porque usuários executivos são incapazes de articular precisamente suas necessidades de informação. Além disso, o uso dos sistemas que disponibilizam os dados do Data Warehouse vão auxiliar o aprendizado do cliente e engatilhar a identificação de novos requisitos. O foco da

elicitação de requisitos deve ser a identificação dos tipos de problemas executivos precisam resolver e que tipos de decisão eles tomam. ”

Na visão de Inmom (ND c), as diferenças entre um sistema computacional regular e um Data Warehouse são as “anomalias”, e a principal delas é que o levantamento de requisitos de um Data Warehouse é extremamente difícil. A principal causa dessa dificuldade é que a audiência para a qual se constrói uma solução de Data Warehouse é extremamente diversa, e desse modo muito difícil de cobrir. Ainda segundo a visão de Inmom, a melhor maneira de se fazer essa análise é dividir e classificar os grupos de usuários segundo uma determinada categoria, e atuar de maneira diversa para esses grupos. São eles:

- **Fazendeiros:** são tipicamente gerentes de produtos, analistas financeiros e analistas de vendas. Esse tipo de usuário geralmente é bem organizado e pertence ao grupo de planejamento de uma corporação. Geralmente, o fazendeiro é o tipo mais fácil de entender e levantar requisitos, pois suas necessidades são claras e bem definidas. O tipo de relatório e dados são facilmente especificados, como por exemplo um analista de vendas que precisa de um relatório de vendas de produto por localidade e período. Além disso, suas exigências raramente mudam.

Outro ponto sobre fazendeiros é que eles geralmente já passaram muito tempo pesquisando dados em fontes esparsas, então eles compreendem e valorizam soluções de dados agregados e “limpos”. Pela mesma razão, esses usuários entendem de tecnologia e seu entendimento da ferramenta é muito bom.

O conhecimento do negócio de um fazendeiro também é maior e, portanto, ele é o típico “key user ” de uma implementação.

- **Turistas:** Os turistas estão entre os usuários mais críticos de um sistema de Data Warehouse, porque compreendem altos e médios gerentes e a diretoria. A alusão com turistas é dada porque eles querem uma visão geral dos dados e verificar a fundo apenas alguns pontos chave. São como turistas que têm pouco tempo para verificar uma cidade toda, então eles fazem um *city tour* e visitam apenas os pontos turísticos mais importantes. Como se trata de alta gerência, a satisfação dos turistas

pode causar um impacto muito grande no projeto. Esses usuários dominam uma perspectiva ampla do negócio, mas não têm conhecimento detalhado de todas as áreas. Eles necessitam de informações básicas sobre a saúde das áreas e se estão satisfeitos com os indicadores de uma área passam para outra, sem verificar a fundo os dados da anterior. Desse modo, percebe-se que os turistas precisam ter a informação que querem de maneira clara, de fácil acesso e com um grau de transformação bastante elevado.

Um outro ponto importante é que os turistas não são operadores de computador experientes, apenas usuários padrão. Desse modo, as interfaces e a maneira de disponibilização de dados deve ser feita de maneira a facilitar sua navegação. Historicamente, é o grupo de usuários que mais dá trabalho para o suporte.

- Exploradores: os exploradores são os usuários mais difíceis de contemplar. Geralmente são analistas estratégicos ou de *marketing*. O modo de encarar o negócio deles é diferente da maioria dos usuários. As suas consultas são randômicas, tentando discernir um padrão escondido nos dados de modo a achar uma correlação que possa dar alguma vantagem estratégica à empresa. Não raramente eles estão errados, mas quando acertam podem elevar em muito as margens de lucro da empresa. Para isso, uma solução que atenda a exploradores deve ser capaz de proporcionar certo nível de detalhes, informação história e capacidade de correlação de fatos.
- Mineradores: a alusão a mineradores vem do fato que esses tipos de usuários “escavam” as montanhas de dados em busca de raras “pepitas de ouro”, ou seja, informação preciosa. Eles analisam esses dados para verificar correlações e tendências. Esses profissionais são representados por profissionais de logística, estatísticos e comerciantes. Suas buscas se baseiam em grandes quantidades de dados detalhados e dados históricos. E o que eles buscam? Geralmente correlações que possam mostrar hábitos de consumo de uma determinada faixa de população, padrões que podem indicar um tipo de fraude. Essa massa de dados de pesquisa deve ser altamente confiável e a sua manipulação deve ser fácil. A principal

diferença entre os mineradores e os exploradores é que os mineradores utilizam métodos formais e matemáticos para fazer as suas buscas, enquanto os exploradores não têm um método sistemático.

- Operadores: Os operadores são os usuários mais comuns de uma ferramenta de Data Warehouse. São pessoas que necessitam gerar relatórios e consultas padrão, para conseguir detalhes. A grande vantagem de uma ferramenta de Data Warehouse para os operadores é que a informação de vários sistemas está contida nele, e assim ele não precisa fazer uma junção de vários relatórios, geralmente em uma planilha eletrônica. As principais necessidades de operadores são dados atuais, bons tempos de resposta, interface simples. Os requisitos desses usuários são fáceis de levantar, porque a sua maneira de uso é previsível e a necessidade de dados históricos é mínima.

Os dois grupos antagônicos mais representativos dessa classificação são os fazendeiros e os exploradores. Os fazendeiros apenas “colhem” os mesmos tipos de dados e fazem as mesmas tarefas, os exploradores reviram os dados em busca de algo interessante, como uma nova tendência de moda ou crescimento das vendas de um determinado produto em uma região específica. A pior consequência disso é que o sistema de Data Warehouse teria que espelhar as necessidades de dois grupos distintos como esses. Além disso, a maneira de se elicitar os requisitos desses grupos seria diferente porque, enquanto os fazendeiros sempre cumprem as mesmas tarefas e estão interessados nos mesmos tipos de dados, os exploradores não tem idéia do que querem ao iniciar uma análise, ou seja, não há um processo definido para seguir. Para Inmon, para os fazendeiros bastam algumas entrevistas e a aplicação de técnicas tradicionais para elicitar os requisitos, mas como podemos entender o que um explorador quer? Como podemos representar em um sistema computacional a maneira peculiar como ele trabalha e, o pior, como podemos garantir que essas necessidades estão bem entendidas. A solução para Inmon, é começar a construção do Data Warehouse e ir mostrando o resultado para o explorador, incluindo no modelo de dados as necessidades que forem sendo expressas por ele. Obviamente esse tipo de abordagem aumentaria o tempo de desenvolvimento, além de possivelmente forçar muito

retrabalho da equipe de desenvolvimento. Há ainda o risco das necessidades dos exploradores acabarem anulando as dos fazendeiros, tornando o sistema inútil para esse último grupo. Esses problemas não escaparam de Inmon (ND e), que descreve a maneira como vários departamentos diferentes de uma empresa descreveriam seu processo como seis homens cegos descrevendo um elefante. Cada um deles seria capaz de apenas dar detalhes das partes com as quais faria contato e essas dificuldades relacionadas ao tamanho do processo e a quantidade de pessoas e departamentos diferentes a cobrir seriam um grande obstáculo à fase de elicitação de requisitos.

Bruckner, List e Schiefer(2001) concordam com essa dificuldade vista por Inmon e também apontam como um grande problema da elicitação de requisitos as diferentes unidades organizacionais que precisam estar contempladas em uma solução de Data Warehouse. Também indicam um outro fator bastante preocupante: eles lembram que, durante a etapa de elicitação de requisitos, os responsáveis pelo levantamento escrevem a sua visão do que foi passado pelos usuários, mas os técnicos que irão construir a solução querem um modelo bem definido para seguir e atualmente não há um padrão para esse modelo. Os técnicos acham a descrição muito informal e acabam fazendo o Data Warehouse com base em suas próprias experiências e tentando fechar os pontos obscuros dos requisitos tendo por base seu próprio entendimento do determinado tipo de negócio. Ao final do processo, os usuários não entendem o que está representado, por acharem técnico demais e, com isso corre-se o risco de que todo o projeto falhe. Entretanto, como não há, aparentemente, uma solução, técnicos e usuários simplesmente seguem em frente. Bruckner ainda faz mais uma assertiva bastante interessante, de que esse tipo de modo de trabalho não pode funcionar pelo simples motivo de que usuários e analistas não falam a mesma “língua”. Pode parecer óbvio, mas esse tipo de constatação não é geralmente levado em conta quando se analisa os fatores de risco de um projeto. Na verdade, analisando-se a literatura sobre a construção de Data Warehouses, como o caso dos livros de Kimball, geralmente o caso é tratado com um modelo teórico preconcebido acerca do problema. Como pode ser visto na maioria dos livros de construção de Data Warehouse, há capítulos que indicam o que procurar em determinados tipos de indústria, e ao seguir esses livros o analista apenas se preocupa com quais os dados a empresa tem que levam ao seu modelo

prévio, ignorando as necessidades reais e as características próprias da entidade para a qual está se fazendo a solução.

Esses dois problemas são os que mais se sobressaem na literatura sobre a elicitação de requisitos de Data Warehouses. Podemos ainda acrescentar o problema causado pela falta de entendimento do que é uma ferramenta de Data Warehouse pelos usuários, e do entendimento de quais dados realmente devem constar na solução. Isso porque não são todos os dados da empresa que devem ir ao Data Warehouse e sim os que devem sofrer algum retrabalho para facilitar a análise e dar apoio ao processo de tomadas de decisão da empresa (Inmom (ND a)) . Informações transacionais e de cunho operacional não devem ser migradas ao Data Warehouse, e sim visualizadas nos sistemas transacionais da empresa. Não é raro ao fazer o processo de levantamento de requisitos para uma solução de apoio à decisão ser colocado como necessidade pelos usuários algum processo como o acompanhamento atual dos estoques da empresa. Isso não é do escopo de Data Warehouse e isso deve ser bem esclarecido para o usuário.

Outro ponto importante é que os usuários de uma solução Data Warehouse nem sempre são pessoas de fácil acesso. Presidentes, gerentes e diretores não tem tempo a “perder” com entrevistas, respostas a formulários e outras práticas normais de elicitação de requisitos e delegam a responsabilidade de definição para outras pessoas, que nem sempre entendem as reais necessidades da alta cúpula de uma empresa. Mesmo essas pessoas, como foi colocado por Shanks e O’Donnell (ND), não sabem exatamente o que querem. Elas sabem que tem uma enorme quantidade de informações em suas empresas e não sabem qual a melhor maneira de utilizá-la. Em alguns projetos de Data Warehouse que participei era comum os *key users* quererem superpopular as bases do Data Warehouse, trazendo toda a informação disponível dos sistemas legados, ou simplesmente resolver que o melhor é ter uma cópia de relatórios já existente na ferramenta. Obviamente nenhuma das soluções é satisfatória, porque impactam diretamente na qualidade da informação que está sendo trazida e na usabilidade da ferramenta de Data Warehouse.

Por outro lado, o analista deve se eximir de tentar entender mais do negócio do cliente do que o próprio cliente. Durante uma apresentação de um analista sobre um projeto de Data Warehouse, foi apresentada a questão do exagero de dados que um usuário queria no seu banco. Uma das informações exemplificadas era que o usuário gostaria de ter correlações com o número de acidentes de trabalho na empresa. O analista achou que não seria útil, já que se tratava de uma solução para a tomada de decisões estratégicas, ter um indicador que fizesse correlação com o número de acidentes de trabalho. Analisando o caso, verificamos que há necessidade de cuidado com certas idéias pré-concebidas quando vamos ajudar um usuário a determinar as suas necessidades. Talvez para uma indústria metalúrgica o índice de acidentes de trabalho não seja relevante, mas e no caso de uma indústria alimentícia, onde um acidente de trabalho, além da produção parada, pode significar a contaminação e perda da matéria prima na linha de produção? Talvez o impacto para o negócio seja bastante grande. A função do analista deve ser entender o problema do cliente, verificar sua forma de trabalho e desenhar a solução de modo a auxiliar o cliente utilizar os dados disponíveis.

É consenso nos artigos existentes sobre Data Warehouse que as técnicas tradicionais de elicitação de requisitos não servem para esse tipo de solução, sendo as ferramentas mais empregadas checklists e questionários como os sugeridos por Kimball(2001). Apesar disso, apenas a partir do ano 2000 começaram a aparecer artigos relacionados a esse tema. Mesmo assim, a maioria dos métodos sugeridos apenas cuida do ciclo de vida do projeto de Data Warehouse, e não indicam nenhuma ferramenta para apoiar nesse processo. Mesmo as poucas exceções não encontram muito respaldo e são criticados na maioria dos fóruns sobre o assunto. Mas o que então se usa atualmente para a elicitação dos requisitos de Data Warehouse? Segundo Inmom, o método mais utilizado ainda é a análise dimensional, do qual ele é o maior defensor, mas como dito mais acima até mesmo ele reconhece as fraquezas desse método.

2.3 Referências da Literatura

Muitos autores têm exposto suas opiniões acerca do desenvolvimento de soluções de Data Warehouse, entre eles os dois de maior expressão no mundo das ferramentas de apoio à tomada de decisão (Decision Support Systems), Bill Inmon e Ralph Kimball. Outros autores, como Price e Shanks (2004) têm também dado a sua contribuição, inclusive usando alguns dos preceitos da semiótica. Apesar do tempo em que se tem trabalhado algumas questões, a análise de requisitos de um sistema de DSS tem sido um assunto pouco explorado, com poucas metodologias e sugestões. Grande parte da literatura técnica segue a linha de receitas prontas, e geralmente as técnicas clássicas de engenharia de software são empregadas, mas sem um ajuste necessário para cobrir as peculiaridades dos sistemas de Data Warehouse.

Em seu livro de 1998, *Data Warehouse Toolkit*, Ralph Kimball versa uma série de instruções e maneiras de se construir um Data Warehouse para os mais diversos tipos de indústrias da atualidade. Segundo suas próprias palavras “*É um conjunto de ferramentas e técnicas de projeto que, quando aplicadas às necessidades específicas do usuário e do banco de dados, permitirá que planejem e construam um Data Warehouse de nível empresarial*”. Entre os objetivos descritos nesse livro, está a criação de um conjunto apropriado de requisitos para o Data Warehouse de sua aplicação. Esse livro é baseado em estudos de caso fictícios, que explicam por exemplos as necessidades da implementação de um Data Warehouse. A primeira verificação que se faz é que existe uma idéia pré-concebida de cada tipo de negócio e o autor determina o que o analista deve buscar e de que maneira ele deve proceder com os dados. Apenas no capítulo 12 da publicação há referências às necessidades de entrevistas com usuários, e para isso o autor sugere um *checklist* de perguntas. O público alvo dessas entrevistas seria uma mistura de DBA (administradores de bancos de dados) e responsáveis pelos sistemas legados (as fontes de dados) e usuários finais. Cada *checklist* está dividido em função da área da empresa e desse modo podemos ver que quem conduz as entrevistas é o analista. Tendo por base a sua visão dos processos, pré-concebida pela literatura técnica, o analista provavelmente irá direcionar

as respostas dos usuários e muitos problemas e necessidades ficarão sem ser conhecidos. Podemos dizer que grande parte da literatura técnica sobre Data Warehouses segue os mesmos preceitos.

Shanks e O'Donnel (ND) discutem algumas das técnicas e ciclos implementados até essa data, e propõe o seu próprio ciclo de desenvolvimento. O ciclo de Shanks destaca a importância da iteração da análise de requisitos na construção de Data Warehouses, bem como um bom entendimento do modelo de negócios da empresa alvo. Mesmo com essas definições, Shanks não sugeriu um modelo de apoio para a elicitação de requisitos, preferindo o uso de técnicas tradicionais de engenharia de software, entrevistando os usuários.

Continuando seu trabalho na área de qualidade de informação e Data Warehouses, Price e Shanks (2004) propõem um *framework* semiótico para avaliar a percepção de qualidade de dados em um ambiente de Data Warehouse pelos Stakeholders. Esse *framework* visava garantir a usabilidade e confiabilidade dos dados inseridos em um Data Warehouse, forçando assim entendimento de cada área sobre quais dados eles eram responsáveis e quais ações eles deveriam tomar para garantir que esses dados fossem úteis para as tomadas de decisão da corporação.

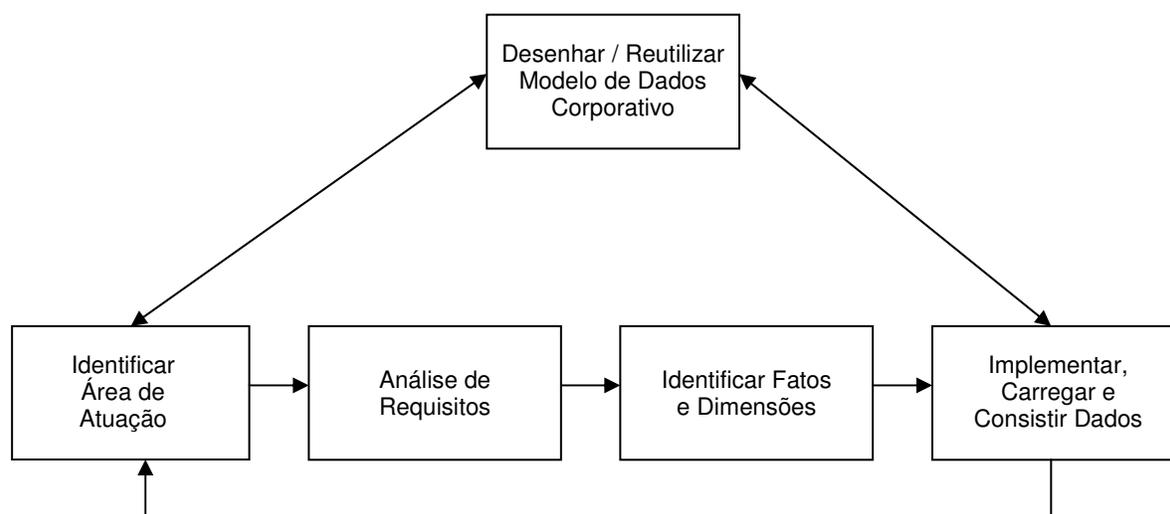


Figura 6 : Diagrama de Etapas de Projeto de uma Solução de Data Warehouse (Shanks e O'Donnell (ND))

Brucker (2001) propôs o uso de Use Cases para a elicitação de requisitos de uma solução de Data Warehouse. Os principais problemas, segundo ele, seriam a dificuldade de entendimento entre usuários e analistas, de modo que a comunicação entre eles deveria ser facilitada por um método de sinais simples que possibilitasse o entendimento mútuo. Com relação aos trabalhos existentes na área, ele considera que, apesar do uso de diversos métodos utilizados para o levantamento de requisitos de um ambiente de Data Warehouse, ainda não existia uma ferramenta formal de suporte à elicitação de requisitos. Sobre os trabalhos relacionados à área, ele faz a seguinte crítica:

- Inmon(1996) dizia que, como um ambiente de Data Warehouse não é direcionado a requisitos e sim a dados, os requisitos necessários apareceriam após o desenvolvimento do sistema com os dados dos sistemas legados. Isso seria feito com a transposição do modelo corporativo diretamente para o sistema de Data Warehouse.
- Anahory e Murray(1997) propuseram um catálogo para conduzir as entrevistas com usuário, da mesma maneira de Kimball.
- Kimball(1998) considerou que o principal passo para a construção de um Data Warehouse seria a escolha de um modelo de negócios e representá-lo. Segundo Bruckner, esse método seria o mais utilizado e mais funcional até a data de seus trabalhos, e assim baseou-se em algumas premissas de Kimball para o seu estudo.

O trabalho de Brucker divide as categorias de requisitos em quatro camadas, os requisitos de negócios, requisitos de usuários, requisitos detalhados de sistemas e requisitos de atributos. Para conseguir esses requisitos, ele propôs o uso de Use Cases em um processo com cinco passos de refinamento, a fim de conseguir uma boa definição das representações de processos de negócios a serem desenvolvidos.

As maiores críticas ao uso de Use Cases para o desenvolvimento de Data Warehouses vêm de especialistas da área. Segundo a maioria deles, o uso dos Use Cases seria atrapalhado pelo principal problema de elicitação de requisitos de um Data Warehouse: os usuários não

sabem o que querem, os processos não estão bem definidos e não haveria como refinar os requisitos ao nível que a metodologia indica (Moss, Barbusinski, Rehm, Oates (2004)).

Winter e Strauch (2003), colocam que soluções de Data Warehouse costumam ter um ciclo de projeto bastante longo e que nesse período os requisitos da solução podem mudar ou tornar-se ambíguos. Também indicam que os requisitos de informação são difíceis de elicitar devido à singularidade ou mesmo falta de estrutura dos processos de decisão, a relutância dos tomadores de decisão em explicitar os detalhes desse processo de análise e tomada de decisão e inconscistências conceituais dentro dessas grandes corporações. Em Winter e Strauch (2004) é colocada a necessidade de um método de apoio para a elicitação de requisitos de grandes projetos, pois consideram que problemas com essa fase do desenvolvimento da solução são responsáveis pela maior parte dos projetos fracassados. Além disso, afirmam que a natureza especial de soluções de Data Warehouse necessitam de uma metodologia de suporte específica, e que os métodos de análise de requisitos atuais cobrem apenas parcialmente as necessidades da elicitação de requisitos desses sistemas.

Em sua pesquisa, Winter e Strauch (2003,2004) entrevistaram vários *experts* no desenvolvimento de soluções de Data Warehouse e derivaram dessas entrevistas o que esses *experts* consideravam necessário que uma metodologia de elicitação de requisitos contemplasse. Entre esses requisitos estão:

- Ser orientada a demanda;
- Apresentar um modelo hierárquico e com várias camadas;
- Que cubra o processo completo de: determinar os requisitos de informação dos usuários, comparar os requisitos com as fontes de informações existente, avaliar e homogeneizar os requisitos de informação, estabelecer prioridades sobre os aspectos não cobertos, formalizar a especificação dos requisitos para as próximas fases do desenvolvimento.

A metodologia proposta por Winter e Strauch (2003) é composta por seis fases, cada uma cobrindo um aspecto do ciclo de elicitação e formalização dos requisitos para soluções de Data Warehouse. Cada uma dessas fases é composta de outras sub-fases, num total de doze

tarefas para a complementação de todo o ciclo. Algumas dessas fases apresentam laços de iteração, nos pontos considerados chaves para a consistência para os requisitos levantados.

A proposta da metodologia de Winter e Strauch (2003) foi parcialmente aplicada em algumas corporações, com resultados variantes em cada etapa. Em sua própria análise da metodologia, Winter adverte que ainda é necessário que algumas ferramentas de apoio sejam testadas e desenvolvidas, pois as técnicas usuais de entrevistas, questionários e análise de tarefas são aplicáveis, mas apenas parcialmente.

Paim (2003) apresentou em seu trabalho uma metodologia de acompanhamento de projeto de Data Warehouse com ênfase no aspecto de engenharia de requisitos e acompanhamento de projeto. É proposto um modelo em espiral para a fase de elicitação de requisitos, contendo quatro fases: Elicitação, Análise e Negociação, Documentação e Conformidade dos Requisitos. Para a fase de elicitação de requisitos, quando os usuários irão indicar suas necessidades aos analistas, são propostas as seguintes técnicas clássicas: entrevistas, workshops, prototipação e desenvolvimento de cenários em Use Cases. Há ainda a preocupação com o que foi chamado de Análise de Requisitos Não-Funcionais, que são requisitos não ligados diretamente à função de um Data Warehouse, como por exemplo performance e segurança de dados. Para a elicitação desses requisitos não funcionais foi proposto um *framework* chamado de DW-ENF (*Data Warehouse Extended NFR Framework*). Esse *framework* consiste de um *checklist* com os principais aspectos a ser discutidos. Mais uma vez o maior problema desse trabalho é a falta de uma ferramenta específica para apoiar a fase de elicitação de requisitos. A maior contribuição desse trabalho é uma metodologia de controle de projeto de Data Warehouse, com ênfase no planejamento da engenharia de requisitos.

Dale (2004) descreve um processo de análise de requisitos em uma grande corporação utilizando-se os processos de qualidade Six Sigma (metodologia de qualidade na qual se pontua o nível de maturidade de um processo de produção baseado em sua quantidade de erros de produção. Quanto maior o número Sigma, menor a quantidade de erros da produção) e DMAIC(Define, Measure, Analyse,Improve, Control), um processo adaptado baseado nessas fases. O ciclo de elicitação baseou-se em entrevistas e verificação as necessidades dos principais usuários, sem o uso de ferramenta específica de apoio. Ao final

do ciclo de elicitação ele pode inferir que o seu maior problema foi conseguir separar o que os usuários realmente necessitavam e não o que eles gostariam de ter no sistema. A necessidade de separar esses requisitos aparece porque a tendência natural dos usuários é querer todas as informações possíveis em um sistema de Data Warehouse, inclusive as que devem ser consultadas nos sistemas transacionais normais. Levar todas as informações da empresa para o Data Warehouse irá causar queda de desempenho e problemas de tempo de carga, impossibilitando o uso correto da solução.

A Semiótica também foi estudada por Shanks e Corbitt (1999), como ferramenta de apoio para soluções de Data Warehouse, mas não para ajudar na elicitação de requisitos. Shanks propôs uma variação do Framework Semiótico para garantir o entendimento das necessidades de qualidade de dados de uma corporação. As conclusões desse estudo remetem a práticas que devem ser tomadas, em grande parte, após a implementação final do modelo do Data Warehouse, como por exemplo, a necessidade de auditoria dos dados pelo departamento responsável antes de sua publicação final.

Ainda como base referencial, grandes empresas de consultoria propõem métodos próprios de implementação de soluções de Data Warehouse. Esses métodos, em grande parte, são decorrentes da experiência dessas consultorias nesse tipo de soluções, mas contêm como principais problemas o fato de não terem apoio de ferramentas específicas, utilizando-se em grande parte técnicas tradicionais como entrevistas e questionários e de forçarem a implementação segundo os seus interesses de venda de ferramentas como determinados tipos de bancos de dados e programas de interfaces com os usuários.

2.4 Considerações Finais

Resumindo, então, podemos dizer que as maiores dificuldades encontradas para a análise de requisitos de uma solução de Data Warehouse são:

- A necessidade de se espelhar os requisitos de muitos departamentos diversos na mesma ferramenta. Apesar disso poder ser minimizado com o uso dos Data Marts, ainda assim todo o processo precisa ser coeso;

- A falta de entendimento entre os analistas técnicos e os usuários, dificultando o entendimento das necessidades uns dos outros;
- Falta de entendimento do usuário sobre o que exatamente deve ir ao Data Warehouse;
- Idéias pré-concebidas dos analistas, por causa da maneira como são tratados os Data Warehouses pela literatura técnica.
- Dificuldade dos próprios usuários em entender e expressar o que precisam. Como colocado por Inmom, muitas vezes um usuário do tipo explorador não tem uma seqüência lógica de busca, não sabe realmente o que quer, e o próprio uso do Data Warehouse mostra novos caminhos a trilhar.
- Falta de apoio de uma técnica específica para a elicitação de requisitos de uma solução Data Warehouse;
- Enorme variedade de técnicas e tecnologias de implementação que, dependendo do uso desejado e do tipo de usuário podem variar de uma solução baseada numa ferramenta MOLAP ou apenas num repositório de banco de dados;
- Enorme quantidade de dados, que devem ser filtrados e agregados de forma correta, a fim de garantir qualidade de informação aos usuários.

Ao estudar os principais problemas existentes com a elicitação de requisitos durante o estudo de viabilidade e posterior implementação de soluções de Data Warehouses e de ferramentas baseadas no mesmo, notamos que, além de uma falta de especificação formal para a elicitação, há também falta de uma linguagem formal de descrição da solução. Dessa maneira, resolvemos trabalhar seguindo o fluxo de criação de um Data Warehouse, esperando que o produto final nos fornecesse uma especificação que pudesse ser compreendida facilmente pela equipe técnica de implementação. Desse modo, decidimos que a melhor maneira de especificar a solução seria criando o mapa das *tabelas fato* e das *dimensões*, que formam os cubos de dados, representadas na forma de *Star Schema*, modelo

de representação de dados amplamente utilizado para o desenho de soluções de Data Warehouse. Com esse modelo das dimensões e dos cubos, qualquer tipo de tecnologia de implementação escolhida, seja ela através de banco de dados relacionais ou de banco de dados multidimensionais, bem como com qualquer ferramenta existente no mercado, poderia facilmente ser ajustada e parametrizada para espelhar essa realidade específica.

Os desafios para a elicitação de requisitos de uma solução de Data Warehouse são conhecidos e entendidos, mas podemos dizer que até o momento as pesquisas para resolver esses problemas estão apenas engatinhando. Ao olharmos com mais atenção os problemas relatados podemos ver que alguns pontos são parelhos com o que a Semiótica Organizacional e os métodos e ferramentas desenvolvidos pelo MEASUR têm a oferecer, e desse modo acreditamos que o uso desse ferramental possa dar algumas respostas ou ao menos algum direcionamento para essas questões.

Capítulo 3

Objetivo e Metodologia

A proposta deste trabalho remete às dificuldades encontradas e enfrentadas na prática de minha carreira profissional, como apresentado anteriormente. Percebi que havia ainda uma grande dificuldade por parte de analistas e usuário e entender o que realmente se precisava para uma solução de Data Warehouse coerente e que aparentemente muito do desenvolvimento dessas soluções eram feitas sem uma fase inicial de análise de requisitos apropriada, inclusive com grandes dificuldades de se determinar o real público alvo do sistema.

Como pôde ser visto no Capítulo 2, as técnicas tradicionais de levantamento de requisitos não atendem completamente as necessidades especiais de uma solução de Data Warehouse, e ainda não existe uma técnica comprovadamente eficaz e com um bom apoio de ferramentas. Desse modo, ao iniciarmos esse trabalho, a principal motivação era tentar identificar um conjunto de ferramentas que pudesse apoiar de modo eficaz a etapa de elicitação de requisitos de uma solução de Data Warehouse. Por causa da afirmação de quase todos os autores de que alguns dos principais problemas do levantamento de requisitos era a incapacidade dos usuários de expressar corretamente o que precisavam e a incapacidade dos técnicos de entender o modelo de negócio das corporações, as ferramentas da Semiótica Organizacional, em especial das ferramentas do MEASUR, aparentemente cobriam essas dificuldades e poderiam servir como uma metodologia eficaz para o desenvolvimento de soluções de Data Warehouse.

As dúvidas principais que tive ao trabalhar com uma solução Data Warehouse estão espelhadas e explicadas nesse trabalho. Ao iniciar o planejamento desse estudo eu visava responder algumas perguntas sobre as reais necessidades de um usuário a serem migradas para uma solução de Data Warehouse e se possível definir se o uso das ferramentas da Semiótica poderiam ajudar nessas definições.

Como foi discutido no Capítulo 2, as referências literárias sobre a elicitación de requisitos para um ambiente de Data Warehouse são escassas e nem sempre cobrem todos os aspectos necessários. Quase sempre são variações de modelos clássicos, mas que não cobrem todas as reais necessidades de um Data Warehouse. Após verificar esse cenário, algumas dúvidas foram somando-se às iniciais, formando a seguinte lista de questões:

1. Quem realmente é o público alvo da solução de Data Warehouse?
2. Quais as pessoas imprescindíveis para a elicitación de requisitos?
3. Quais os reais problemas/intenções da corporação em relação a uma solução de Data Warehouse?
4. A empresa possui maturidade suficiente para operar e usufruir de uma ferramenta tão sofisticada e específica?
5. Qual o perfil dos usuários que irão usar a solução? Quais os seus modos de trabalho e operação?
6. A empresa já dispõe de todos os dados necessários para as análises que deseja?
7. Como fazer as pessoas confortáveis o suficiente para quebrar barreiras e permitirem-se ensinar aos analistas os segredos do seu trabalho?
8. Como garantir que esses analistas realmente entenderam e não apenas interpretaram os resultados segundo os seus direcionamentos pessoais e de suas empresas de implementação?
9. A qualidade dos processos atuais é boa? Quais processos são os mais problemáticos e deveriam ser ajustados primeiro?

Ao estudarmos as ferramentas da Semiótica Organizacional e em especial o MEASUR, notei que esse método poderia ser promissor na análise de requisitos de um Data Warehouse. O *Framework Semiótico* aparentava cobrir todos os grandes aspectos necessários para uma solução de Data Warehouse, incluindo os aspectos técnicos de Tecnologia de Informação, os aspectos sintáticos que ajudariam na homogeneização dos dados dos sistemas fonte e em especial o nível semântico, que remete à necessidade de entender as intenções de quem usaria uma ferramenta de apoio à decisão.

Assim sendo, o propósito inicial desse trabalho é verificar se as ferramentas de Semiótica Organizacional podem ser utilizadas nas fases iniciais de planejamento do desenvolvimento de uma solução de Data Warehouse, isto é, se com o apoio dessas ferramentas poderíamos elicitar os requisitos iniciais da solução, levantar suas principais fontes de dados, seus principais usuários, em quais processos de decisão a solução deve ser focada e quais as falhas nesses processos. Ao final do estudo, o objetivo é ter mapeado os processos na forma de um gráfico Star Schema, que daria a visão dos “cubos” de dados necessários. Essa representação baseada no Star Schema é livre de plataforma de tecnologia e por isso torna-se adequada a qualquer análise dimensional, pois a sua adaptação às plataformas de tecnologia conhecidas, OLAP’s ou ROLAP’s, é feita de forma rápida e transparente.

O estudo do MEASUR indicou que para o objetivo inicial algumas ferramentas contidas no método PAM eram as mais indicadas, pois esse método se presta a uma fase do processo de construção da solução em que o problema a ser resolvido não está muito claro. Para o estudo de caso o ideal era buscar uma empresa que tivesse as seguintes características:

1. Cujo trabalho necessitasse de grande quantidade de informações;
2. Que tivesse uma estrutura fixa e bem delimitada, com funções e cargos diferenciados;
3. Que concordasse em dispor de algum tempo para as atividades sugeridas.

Das opções que foram discutidas, a que mais se adequou às necessidades foi a Vigilância Sanitária do Estado de São Paulo – Diretório 1 – Capital. Além das características requeridas, os contatos iniciais para verificar a possibilidade do trabalho indicavam mais alguns pontos que poderiam por à prova a capacidade das ferramentas da Semiótica para a elicitação de requisitos. Entre essas características estão:

1. Como se trata da área de saúde, seus termos técnicos e necessidades seriam totalmente desconhecidos por mim;
2. Histórico anterior de casos de tentativas de construção de soluções diversas sem sucesso;

3. Algum apoio de ferramentas computacionais, que com certeza seriam integradas à solução;
4. Muitos processos e projetos em andamento, criando um grande fluxo de informações.
5. Pessoal pouco familiarizado com sistemas de informação e em especial Data Warehouses.

Após as negociações iniciais e dada a anuência do Governo do Estado de São Paulo, foi delimitada a metodologia a ser utilizada para o estudo de caso.

Para garantir a confiança do grupo estudado de que nenhuma informação confidencial referente ao trabalho da Vigilância seria disponibilizada, ficou acertado que eles teriam acesso a todo o material publicado e produzido. Isso é necessário por causa das informações contidas em processos de empresas e informações de interesse à saúde.

A principal preocupação nesse caso era não influenciar os participantes do estudo de caso e evitar que eles tentassem “adivinhar a resposta esperada”, ao invés de discutir seus reais problemas. Foi definido que o estudo de caso deveria transcorrer da seguinte forma:

1. As atividades seriam realizadas nas dependências da Vigilância, e eu levaria o material de apoio necessário;
2. Seriam feitas ao todo quatro reuniões, uma para cada técnica sugerida, a saber: Análise de Stakeholders, Quadro de Avaliação, Framework Semiótico e a Análise Colateral;
3. Não seria feito um treinamento formal das técnicas com os usuários e nem sobre o Data Warehouse. A princípio estaríamos levantando seus fluxos de informação e problemas com seus processos;
4. Para a primeira reunião foi convidado um representante de cada área, diretores e assistentes. Caso durante a análise de stakeholders fosse notada a falta de uma figura importante, essa pessoa seria convocada nas próximas reuniões;
5. As reuniões seriam *workshops*, onde os usuários seriam expostos às técnicas sugeridas e os resultados anotados;

6. Todas as reuniões seriam gravadas e notas seriam tomadas. Além disso, material de apoio em forma de cartazes seria utilizado nos *workshops*. Esses cartazes seriam transcritos após cada reunião para a análise dos dados. As gravações foram autorizadas e o gravador sempre ficou visível.

Em cada reunião a ferramenta a ser utilizada era apresentada e explicada, com o apoio de literatura e alguns exemplos, como o trabalho de Liu (2000) e Simoni (2003). Alguns ajustes foram necessários conforme o estudo de caso transcorria. Por causa das dúvidas encontradas pelo grupo no primeiro *workshop*, utilizamos parte do segundo para explicar mais a fundo o que era uma ferramenta de Data Warehouse. Após o final da segunda reunião decidimos também não mais utilizar o gravador, pois ele claramente trazia desconforto aos usuários.

O material transcrito seria compilado e analisado e um projeto de uma solução de Data Warehouse seria feito, com um alto grau de abstração. Também como resultado um modelo de uso das ferramentas da Semiótica seria produzido, provendo a possibilidade de aplicação da técnica em outros cenários.

Capítulo 4

Semiótica Organizacional como Referencial para a Elicitação de Requisitos

4.1 A Semiótica

A forma como uma pessoa vê o mundo está diretamente relacionada a suas próprias experiências. A sua linguagem falada, a sua maneira de expressão, sua maneira de vestir e o significado de suas ações estão diretamente relacionadas com o meio em que esse indivíduo cresceu, quais valores e informações foram-lhe passados por outras pessoas e a sua maneira pessoal de ver o mundo. Cada interpretação da realidade pode ser considerada única e particular.

Cada grupo particular de pessoas tem a sua própria linguagem, suas gírias e códigos de condutas. Policiais, professores, médicos, cada pessoa interage com um determinado grupo social e adquire parte do comportamento desse determinado grupo, começa a compartilhar a sua maneira de ver o mundo e absorver novas experiências que podem modificar as suas próprias, levando a novas conclusões e talvez até mesmo a modificações no modo de viver e de agir. Essa comunicação, seja falada, escrita, existe através de símbolos, utilizados para passar alguma mensagem para o grupo. Para ser entendido, o símbolo deve ser conhecido e ter um significado próprio e que faça sentido para a pessoa que recebeu a mensagem. Apesar de parecer simples, a maneira como os símbolos são utilizados pelas sociedades é extremamente complexa e alvo de estudos há longo tempo. Uma das linhas de estudo desses fatores é a Semiótica, ou Semiologia. A palavra Semiologia é formada pelos radicais gregos *logia*(estudo) e *semeîon*, que quer dizer signo. A principal diferença de uso entre essas palavras é que a primeira, Semiótica, refere-se à Charles Sanders Peirce(1839-1914), filósofo estadunidense que acreditava que a semiótica era a doutrina formal dos símbolos, relacionado a uma Lógica Formal. Já a Semiologia remete ao linguista suíço Ferdinand de Saussure(1857-1913), que definiu que a semiologia deveria investigar os signos e as leis que os governam.

Vários outros estudiosos de diversas áreas de conhecimento também trabalharam com o conceito semiótico, como Roland Barthes(1915-1980) e o romancista Umberto Eco. Barthes escreveu, em 1964: “a semiologia visa conseguir, em qualquer sistema de signos, o que for sua substância ou limites; imagens, gestos, sons musicais, objetos e qualquer associação complexa de todos esses elementos. Isso constitui, se não uma linguagem, ao menos um sistema de significados.” (Barthes (1964)).

Morris (1970), responsável pelos primeiros desenvolvimentos da semiótica comportamental, declarou, baseado nos trabalhos de Pierce, que a semiótica é baseada em três fatores, sendo eles:

- Semântica: o relacionamento dos signos com o que eles significam;
- Sintática: a relação entre signos, seja formal ou estrutural;
- Pragmática: a relação dos signos para o interpretante.

Mas o que é um signo? Qualquer coisa que tenha um significado pode ser considerado um signo, seja uma imagem, palavra ou símbolo. Segundo Pierce(1931-58), “nada é um signo até ser interpretado como um signo”. Ou seja, a partir do momento que alguém recebe uma mensagem através de algo e interpreta essa mensagem de determinada maneira, o signo se torna válido. Com isso, notamos que para haver a existência do signo é necessário que ele tenha um significado. Saussure sugeriu um modelo de duas partes para a representação do signo, sendo que :

- *Signifiant²*: é a forma do sinal;
- *Signifié*: é o conceito que ele representa.

Vemos então que a figura do intérprete do signo não pode ser descolada do mesmo, ou seja, apenas após um processamento mental o signo pode ser relacionado ao seu significado.

² Os termos foram mantidos em sua forma na língua francesa para identificá-los como definições de Saussure em contraponto às definições de Pierce, da vertente anglo-saxã.

Na mesma época em que Saussure fazia sua interpretação com o modelo de duas partes, Pierce apresentou seu próprio modelo, contendo três diferentes entidades:

- *Representamen*: a forma que o signo tem, não necessariamente material;
- Interpretante: não um intérprete, mas a quem o signo leva algum sentido;
- Referente: a o que o signo se refere.

Assim, segundo Pierce (1931-1958), “um signo, na forma de um *representamen*, é algo que significa algo para alguém em respeito ou relação a algo”. Ou seja, ao ver um signo (*representamen*), um intérprete usará de seus conhecimentos prévios e experiências para evocar um sentido, que irá remeter ao referente. Essa interrelação Pierciana entre essa tríade é chamada de semiose.

O processo pode ser ilustrado da seguinte maneira. Imagine que haja uma caixa com a palavra CPU escrita nela. Ao ler essa palavra, um interpretante pode imaginar, conforme seu processo cognitivo, um computador. Desse modo, o *representamen*, a palavra CPU, foi remetida a um computador pelo processo mental do intérprete. Por outro lado, se a pessoa que ler a palavra CPU for versada em eletrônica, provavelmente seu processo cognitivo vai pensar na *Central Processing Unit* de um computador, ou seja, apenas um um componente do mesmo em não no todo. Com isso, o mesmo *representamen* levou a dois referentes diferentes, por causa dos respectivos intérpretes.

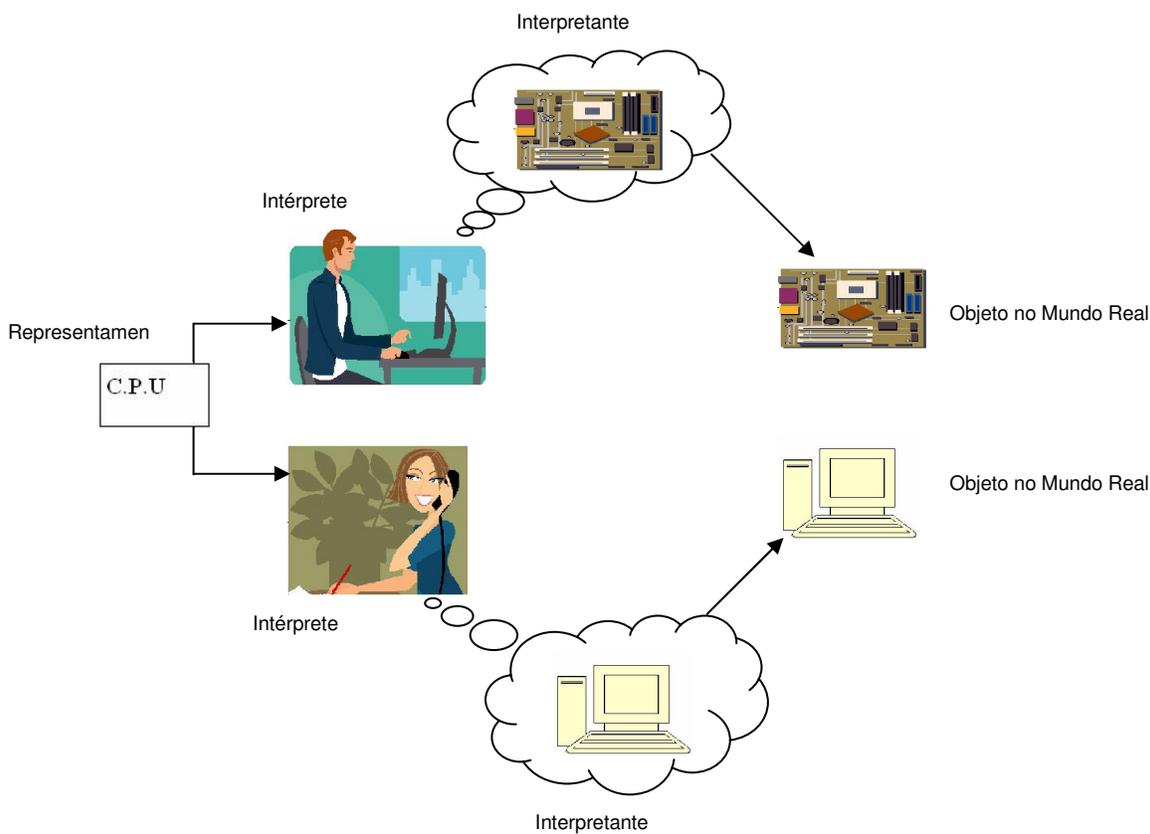


Figura 7: A interpretação de representamen por interpretantes diferentes

Pierce também trabalhou, diferentemente de Saussure, em uma taxonomia para os tipos de signos, forçando uma classificação. Infelizmente, ele era extremamente metódico e elaborado e sua classificação final de signos tem mais de 59000 subcategorias. As mais conhecidas são o símbolo, o ícone e o índice:

- **Símbolo:** é uma representação em que o *representamen* não tem características que o associem diretamente, por aparência ou qualquer semelhança, ao referente. O significado é puramente convencional e precisa ser aprendido ou combinado. É o caso de códigos, letras, números, entre outros.
- **Ícone:** é uma representação que apresenta características que ligam o *representamen* diretamente ao referente. É o caso de um desenho, fotografia, caricatura, onomatopéia ou qualquer forma de representação que não necessite de

conhecimento prévio para a sua codificação

- Índice: é uma representação não arbitrária, mas diretamente conectada ao referente de alguma maneira. É como se fosse uma “pista” ou indicação para o referente. Podem ser pegadas indicando passagem, fumaça indicando fogo, sintomas de doenças, etc.

A Semiótica têm se espalhado em várias outras áreas de conhecimento, e tem se aproveitado das novas mídias providas pela tecnologia. A maneira de visualização e interpretação mudam com o tempo e com a sociedade, e essa evolução é refletida na sua maneira de comunicação e em suas necessidades.

4.2 A Semiótica Organizacional

A Classificação corrente de Semiótica Organizacional foi dada em um *workshop* que envolveu vários estudiosos de várias áreas de aplicação da semiótica (OSW 1995). Ela pode ser definida como: o estudo das organizações utilizando-se os conceitos e métodos da Semiótica. A premissa por trás do uso da Semiótica para o estudo de empresas, sejam elas públicas ou privadas, é que como todo organismo ou organização cada uma delas têm a sua maneira própria de criar, lidar e transformar informações, utilizando um conjunto de signos e códigos próprios. O estudo dessas formas de comunicação pode indicar como é o real comportamento da empresa e quais seus valores e formas de trabalho, em um nível mais profundo.

O uso da Semiótica no campo de Tecnologia de Informação fica claro quando se analisa essa premissa inicial. A necessidade de entendimento de uma empresa, das suas reais metas e objetivos, sua cultura e modo de trabalho deve ser espelhado em um sistema de informação que seja construído para essa organização. Um sistema genérico, que não leve em conta a cultura da empresa e não seja passível de adequação às necessidades reais da organização invariavelmente será subutilizado e gerará insatisfação para os usuários. Mesmo soluções que servem para determinado propósito, como as ERP's, devem ser

adequadas à maneira de trabalho do grupo que as utilizará. A Semiótica pode ajudar a abrir um novo panorama sobre uma organização, facilitando o entendimento do que a entidade realmente necessita para atingir seus objetivos, respeitando a sua cultura, diretrizes e moral.

As ferramentas e métodos Semióticos desenvolvidos por Stamper (1993), com base no trabalho de Pierce, foram geradas a partir de um paradigma subjetivista, antagônico a uma postura objetiva em voga. A principal diferença entre esses dois paradigmas é que a visão objetivista indica que a realidade é única, sendo que os indivíduos inseridos nessa realidade a constroem com suas experiências e sensações, mas ainda assim a realidade é a mesma para todos os referenciais. Todas as idéias e experiências de um indivíduo seriam passadas por outros indivíduos, e as idéias seriam as mesmas para todos. Não haveria espaço para outras interpretações de realidade, essa seria única para todos. Qualquer diferença de visão da realidade seria uma anomalia, pois ela é sempre a mesma para qualquer referencial. Assim, todo o conhecimento e visões da realidade poderiam ser previstos por simples extrapolação e o todo representado por equações matemáticas imutáveis. A própria evolução cultural poderia ser medida e prevista.

Por outro lado, o paradigma subjetivista acredita que a realidade depende da interpretação de um ser para existir, e que cada realidade é pessoal e única. Por causa dessa diferenciação, uma postura subjetivista parece mais adequada para a construção de sistemas de informação, pois como cada realidade é diferente e apenas compartilhada em alguns aspectos, cada nova implementação de sistemas, mesmo que seja utilizando-se a mesma ferramenta e em empresas de nichos bastante similares, é nova e única. Os requisitos dos usuários mudarão, as suas formas, métodos e ferramentas de trabalho serão diferentes para cada nova instância.

A tabela 1 sumariza e exemplifica as principais diferenças entre os dois paradigmas dentro das necessidades dos sistemas de informação.

Tabela 1 : Diferenças entre os paradigmas Objetivista e Subjetivista. (adaptada de Liu (2000, p. 25)).

Conceito	Objetivismo	Subjetivismo
Realidade	Dada objetivamente, a mesma para todos e composta por entidades, suas propriedades e relacionamentos	Criada subjetivamente e socialmente com diferenças entre grupos dos agentes conhecidos
Dados	Um meio de representar a verdade sobre a realidade	Um meio de indicar intenções e coordenar ações
Verdade	A correspondência correta entre entidades reais	Um consenso alcançado, temporariamente, como base por uma ação coordenada
Significado	A relação entre um signo e alguma entidade real	A relação entre o signo e algum padrão de ação estabelecido como norma pelo grupo
Sistemas de Informação	Um tipo de sistema de escoamento pelo qual dados fluem	Um sistema semiológico, principalmente informal, mas suplementado com mensagens formalizadas
Papel do Analista	Especificar as estruturas de dados e funções necessárias aos usuários	Ajudar aos usuários a articular seus problemas, descobrir suas necessidades de informação e desenhar uma solução sistêmica.

Como podemos ver, a principal diferença no que concerne a sistemas de informação é o papel do analista, que é responsável por ajudar o usuário a definir as suas necessidades, no paradigma subjetivista, ao invés de apenas focar nas características técnicas e funcionais do sistema, segundo o paradigma objetivista. Assim, potencialmente um analista que esteja usando o paradigma subjetivista terá ao final de seu trabalho um entendimento maior das necessidades da entidade estudada, e será capaz de especificar soluções mais afinadas com a realidade dos mesmos.

A metodologia semiótica utilizada nesse trabalho baseia-se no programa de pesquisas iniciado nos anos 70 por Ronald Stamper, conhecido como MEASUR (Methods for Eliciting, Analysing and Specifying User's Requirements). O principal objetivo desse programa era a definição de uma série de ferramentas e metodologias para possibilitar a pesquisadores e usuários um melhor entendimento, desenvolvimento, gerenciamento e uso de sistemas de informação(Liu 2000). O paradigma adotado por esse programa foi o de subjetivismo radical, que entende que a realidade é vista como uma construção definida pelo comportamento dos agentes envolvidos. Conforme os agentes interagem, conversam compactuam, tomam decisões, o mundo muda. A realidade é, além de única e pessoal, mutável e adaptável. E essas mudanças no mundo irão forçar novas evoluções, novos comportamentos e novas realidades: desse modo as mudanças serão contínuas.

Tradicionalmente Pierce classificou a semiótica com três divisões, a sintática, a semântica e a pragmática. A fim de abranger todos os aspectos necessários para a elicitação de todo o ciclo de requisitos de um sistema, Stamper adicionou outros três e formou seu Framework Semiótico, em forma de escada. A Figura 8 mostra esse *framework*. Podemos ver a subdivisão entre dois aspectos, o de plataforma de tecnologiae as funções de informação Humanas.



Figura 8 : O Framework Semiótico (adaptado de Liu (2000, p. 27))

Cada um dos degraus da “escada” formada pelo framework indica um ramo da Semiótica necessário para o desenvolvimento de sistemas. Cada um deles tem uma determinada função e representatividade:

- **Mundo Físico:** indica as características e signos físicos, que podem ser medidos por análises físicas e de engenharia. Seus representantes são os sinais elétricos, marcas e outros meios reais. No caso de um sistema de informação, podem ser representados pelo computador utilizado, pela fibra ótica que transmite os impulsos de luz, por esses próprios impulsos de luz que transmitem um código.
- **Empírico:** Esse ramo da semiótica estuda as propriedades dos sinais. Por exemplo, quando um sinal elétrico é enviado por um fio de cobre, qual a tensão nominal que indica um determinado tipo de sinal? Do mesmo modo ele estuda as necessidades dos canais, como os sinais fluem no meio, qual o nível de redundância necessária, entre outras informações.
- **Sintática:** responsável pela apresentação formal da informação. Indica a linguagem, gramática, método de formatação, ou seja, regras de composição dos signos.
- **Semântica:** Segundo Liu (2000), o termo semântica normalmente é considerado

para descrever a relação entre um signo e seu significado. Como vimos, o significado de um signo depende de uma série de fatores, e é regido por normas pré-estabelecidas entre as entidades que estão se comunicando.

- **Pragmática:** Toda a comunicação tem subjacente uma intenção. Sempre que uma pessoa passa uma ordem, uma instrução, há a vontade que isso seja acompanhado de uma ação ou entendimento. Segundo Liu (2000), a pragmática é um ramo da Semiótica que estabelece o estudo do relacionamento entre um signo e o comportamento dos agentes envolvidos. O entendimento e interpretação, mais uma vez, dependem do ambiente e de experiências prévias dos agentes, de forma que para a mensagem surtir o efeito desejado, deve existir um contexto básico e o entendimento entre o teor da mensagem deve ser mútuo entre os participantes das interações. Como podemos perceber, ao se fazer um sistema computacional a pragmática deve ser levada em conta, pois se não for clarificada e implementada poderá acarretar em um fracasso de desenvolvimento.
- **Mundo Social:** A comunicação entre duas ou mais pessoas, segundo (Liu 2000), provoca uma modificação em nível social. Isso é aderente com a idéia de que a realidade é subjetiva e mutável. Quando informações são repassadas, ordens e instruções dadas, quando uma mensagem é interpretada e passa pelo processo cognitivo de uma entidade, o mundo social sofre uma pequena mudança. O estudo semiótico do mundo social evoca a necessidade de entender as normas sociais que regem as comunicações em grupos, e interpretá-las de modo a entender como essas normas e interações entre os grupos funcionam.

Um bom exemplo de análise utilizando-se a escada semiótica é apresentado por (Liu 2000, p.35-36). Como poderíamos analisar uma conversa telefônica por meio da semiótica?

1. Em nível físico, é necessário que os aparelhos de telefone estejam conectados a uma linha telefônica de prestadoras de serviço.
2. Em nível empírico, o sinal de voz é modulado e transmitido em forma de sinais

eletrônicos ou óticos pelo cabeamento.

3. Em nível sintático, as duas ou mais entidades envolvidas na conversa devem falar a mesma língua e devem utilizar convenções gramaticais válidas a ambos.
4. Em nível semântico, as palavras, termos técnicos e não técnicos, e as coisas que são referenciadas durante a conversação precisam ser conhecidas e entendidas por todos. As sentenças e o conteúdo da conversação precisam fazer sentido para todos.
5. Em nível pragmático, existe a preocupação com a intenção, e pode haver mensagens subliminares na comunicação. O exemplo pode ser uma negociação de preço, onde uma diz que gostaria de comprar, mas acha o preço um pouco alto, tendo claramente a intenção de negociar um desconto.
6. Em nível social, obrigações e acordos são feitos, por causa do resultado da conversa. Seguindo o exemplo, o desconto pedido ser dado ou não.

O MEASUR utiliza-se desses conceitos para prover os métodos necessários para o uso da Semiótica para uma solução computacional. Ele é composto de cinco métodos principais, que são:

1. Método de Articulação de Problemas (PAM): compreende uma série de métodos para serem utilizados em fases iniciais do projeto, enquanto o problema a ser resolvido não está bem entendido e ainda expresso de forma vaga.
2. Método de Análise Semântica (SAM): esse método tem como entrada um ponto focal ou problema determinado, e tem como principal objetivo ajudar ao usuário a elaborar e representar seus requisitos de forma clara e precisa.
3. Método de Análise de Normas (NAM): visa especificar formalmente os padrões gerais de comportamento de uma determinada entidade, e está focado nas normas culturais, sociais e organizacionais que regem o comportamento e ações dos agentes em seus ambientes de negócios.

4. **Análise de Comunicação e Controle:** ajuda na análise das comunicações entre todos os agentes envolvidos e o sistema focal
5. **Análise de Meta-Sistemas:** trata todo o desenvolvimento como um objeto de estudo, e ajuda no planejamento e gerenciamento do projeto.

O objetivo desse estudo é verificar se as ferramentas definidas pela MEASUR poderiam ser aplicadas para a elicitação de requisitos para um ambiente de Data Warehouse, em uma determinada entidade. Devido ao tamanho e complexidade do projeto, esse trabalho foi focado na fase de elicitação inicial e por isso a ferramenta mais adequada seria o PAM.

4.3 O Método de Articulação de Problemas (PAM)

O PAM consiste de um conjunto de ferramentas que deve ser utilizado no estágio inicial do desenvolvimento da solução, enquanto há apenas uma idéia do problema, estando esse representado de forma vaga.

Segundo (Liu, 2000), essa fase inicial pode ser considerada a fase de análise de infraestrutura, e pode ser pesquisada com o uso combinado do Framework Semiótico e o PAM. Na sua visão, Liu considera essa primeira análise importante para entender contextos sociais, organizacionais, técnicos e culturas de uma organização e, com isso, identificar problemas de processos de negócios e técnicos.

4.3.1 Análise de organização e contexto

Essa etapa permite conhecer, através da análise de Stakeholders, todas as pessoas que estão envolvidas em um determinado processo, bem como fazer uma avaliação total da organização e contexto para o sistema proposto.

A análise de Stakeholders deve estabelecer as partes interessadas no processo, a fim de garantir a representatividade de todos os envolvidos. Isso é extremamente importante

porque geralmente não há pessoas com o conhecimento total de todos os aspectos de um processo ou estrutura, e a falta de uma parte das informações pode provocar a implementação de um sistema falho e inoperável, ou pior, pode ser um sistema que resolva o problema errado. Para essa análise, podemos utilizar a estrutura presente na Figura 8, que representa todas as camadas de interação possível com um sistema, da mais interna à mais externa. As representações dessas camadas são:

- Operação: indica o sistema focado na avaliação;
- Contribuição: representa toda as pessoas com parte ativa no processo de desenvolvimento, com conhecimentos sobre os processos e necessidades a serem analisadas e automatizadas pelo novo sistema;
- Fonte: Todas as entidades relacionadas como clientes e fornecedores, que impactam ou são afetados pelo sistema;
- Mercado: entidades que servem como referencial para o sistema, sendo por comparação ou por competição.;
- Comunidade: o mundo social, como por exemplo legisladores, jornalistas, todos aqueles com potencial para influenciar indiretamente o sistema.

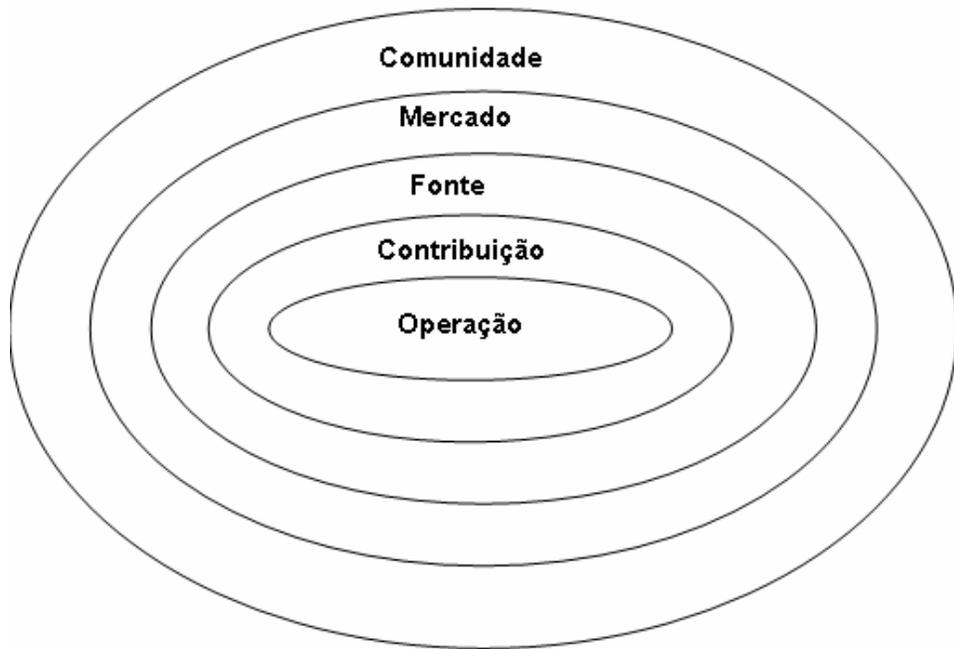


Figura 9 : Camadas da Análise de Stakeholders

Outra ferramenta é o Framework Antropológico, que ajuda a definir aspectos do sistema de informação facilitando a análise dos grupos sociais envolvidos. O Framework Antropológico avalia os seguintes aspectos:

- Interação: relativo à comunicação;
- Associação: grupos, relações, organizações;
- Subsistência: Questões econômicas, financeiras, financiamento, investimentos;
- Taxonomia: entendimento das classificações empregadas;
- Espaço: tamanho necessário, localização;
- Tempo: histórico, datas, período;
- Aprendizado: cognição, treinamentos, conhecimento necessário;
- Criatividade: processo de inovação, cultura;
- Defesa: como se preparar de ataques, como responder a questionamentos,

segurança;

- Exploração: utilização de materiais, processos, perfis.

Com base nas informações prévias levantadas, principalmente da Análise de Stakeholders, pode-se montar os Quadros de Avaliação. Esse quadros levam a transposição das pessoas e entidades interessantes para o processo a elicitar, por meio de um *brainstorming*, quais os problemas encontrados na solução ou processo atual, quais as condições e efeitos necessários e quais as soluções sugeridas pelos próprios usuários.

4.3.2 Análise da Morfologia Funcional

A Análise da Morfologia Funcional ajuda a representar a estrutura de cada unidade da entidade estudada. É considerada por Liu bastante similar ao método estruturado de análise de requisitos, “top-down”. A principal diferença está numa visão diferente, sendo a análise da morfologia funcional focada na identificação das normas que regem o comportamento das pessoas dentro de uma unidade de sistema. Cada unidade pode ser dividida em três subpartes:

- Substantivo: esse comportamento é regido por tarefas, regras e normas, derivadas dos ojetivos organizacionais dentro de uma determinada estrutura. As ações tomadas dentro dessa categoria devem levar a empresa a teoricamente alcançar seus objetivos, ou seja, é o que o funcionário deve fazer no seu trabalho a fim de adicionar valor agregado aos produtos ou à própria empresa;
- Comunicação: essa subparte é ligada a signos, está diretamente relacionada com as entradas e saídas dos processos de comunicação. As normas de comunicação direcionam a passagem de instruções e mensagens entre os agentes da empresa, dando subsídios de informação para o comportamento substantivo da empresa, indicação de medidas a serem tomadas, suporte a tomada de decisões, etc. Os sinais são utilizados entre agentes que os entendam, a fim de denotar uma intenção.
- Controle: Esse comportamento é diretamente ligado à idéia de controle e existe para

garantir que os dois comportamentos anteriores alcancem seus objetivos. É governado por normas, que estipulam regras e regulamentos, contratos, ameaças ou acordos. Caso um agente falhe em cumprir suas obrigações, ou então não consiga, a parte de controle é responsável pela aplicação de sanções, a fim de garantir que o processo seja continuado.

A figura 10 representa a decomposição de uma unidade X através dos três comportamentos. Nota-se que, mesmo dentro do ramo de um comportamento ele pode ser dividido em ações representativas dos outros, e assim até atingir-se uma tarefa atômica dentro do diagrama.

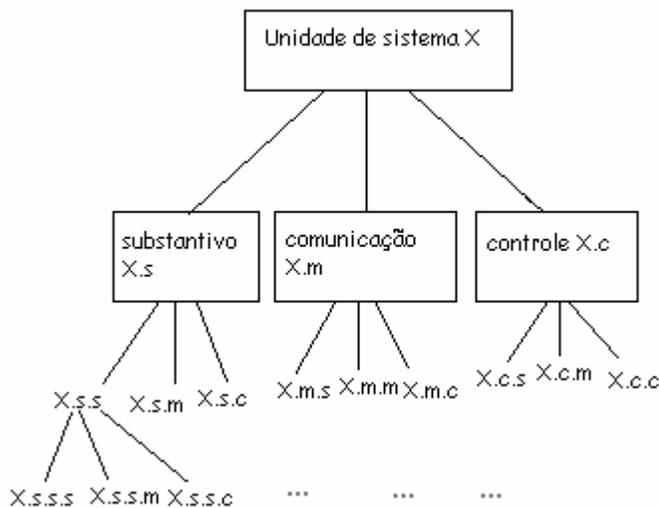


Figura 10 : Representação da Análise de Morfologia Funcional

4.3.3 Análise Colateral

A análise colateral ajuda a entender um sistema complexo em unidades ou ciclos menores, permitindo a visualização de fronteiras em um sistema mais complexo. É como se a técnica de “divisão e conquista” fosse aplicada em um sistema tecnológico total, dividindo-o em subcomponentes tecnológicos que facilitariam a sua implementação. O escopo da análise é

um sistema como ponto focal e então estuda-se todas as unidades de sistemas que interagem com ele, formando sistemas colaterais. Esses sistemas têm, dentro de si, um novo ponto focal, e assim por diante, até que seja garantido que todos os pontos de implementação estejam cobertos. O objetivo é garantir que mesmo os pequenos detalhes da implementação sejam cobertos, de modo a evitar ao máximo obstáculos não previstos durante o planejamento.

Os subsistemas de uma solução geral podem ser divididos nos seguintes ciclos:

1. Ciclo de vida: representa uma realidade temporal, indicando o sistema focal, seu predecessor e seu possível sucessor.
2. Funcionamento: Indica quem fornece o quê ao sistema, e como as saídas do sistema focal influenciam o ambiente, sendo realimentadas no mesmo.
3. Operação: indica os ciclos de construção e desmonte do sistema, quais os recursos disponíveis ao projeto, quais serão absorvidos, quais serão liberados, o que sobrar após o desmanche.
4. Análise e projeto: Documentação das análises e projetos, plantas, recomendações de uso, representação do sistema.
5. Manutenção: Ciclo de manutenção do sistema, quais as janelas de operação, quais as janelas de desenvolvimento, quais os alertas, quais as operações de prevenção?
6. Backup: quais os procedimentos em caso de queda, quais os recursos de recuperação, o que é necessário para manter a saúde do sistema e manter o nível de SLA (Service Level Agreement) da operação?
7. Aprendizado: Melhorias durante a implementação conjunta com correção de falhas, evolução real do sistema.

A figura 11 representa o ciclo proposto.

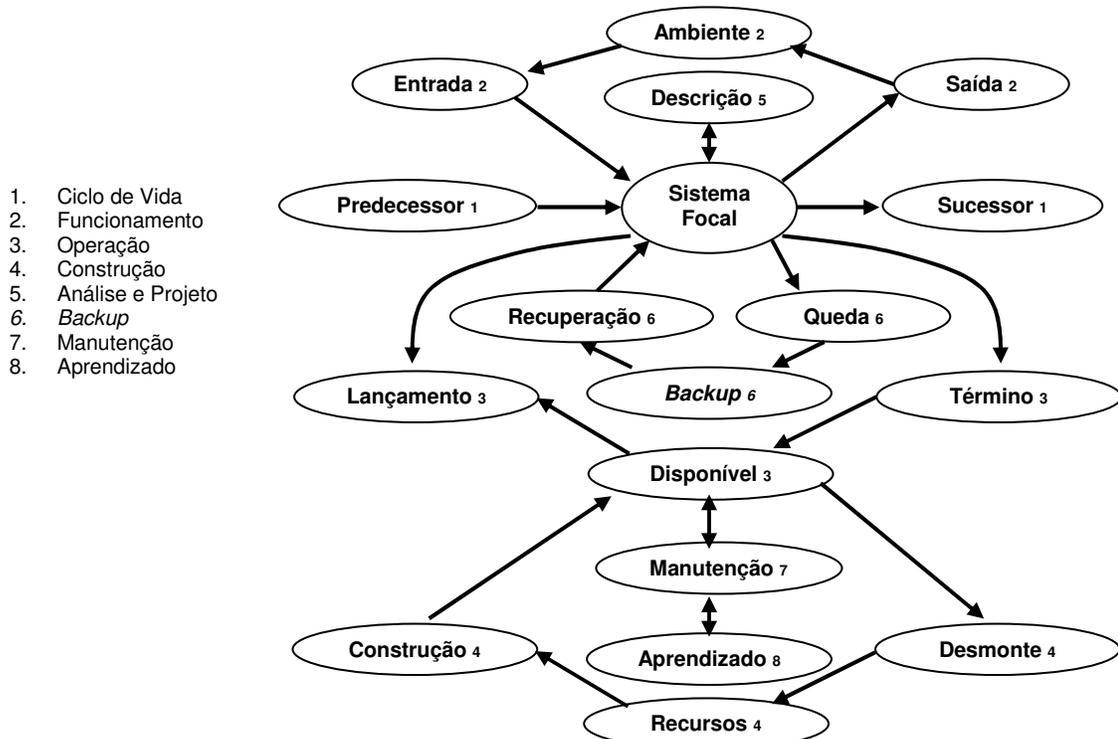


Figura 11 : Ciclo da Análise Colateral. (Simoni, 2003, adaptado de Liu, 2000).

4.4 Considerações Finais

Neste capítulo apresentamos o referencial teórico da Semiótica que serviu como base para a criação da Semiótica Organizacional, objeto deste estudo. Foram apresentados os conceitos básicos da Semiótica, sua extensão para a Semiótica Organizacional e como os pesquisadores do projeto MEASUR utilizaram esses conceitos para gerar uma série de ferramentas e métodos para a elicitación de requisitos. Este trabalho utilizará, no próximo capítulo, as ferramentas apresentadas para verificar sua validade de emprego para a elicitación de requisitos em um ambiente de Data Warehouse, em especial o método PAM.

Capítulo 5

Buscando uma solução preliminar de Data Warehouse: Um Estudo de Caso

5.1 Planejamento do Uso das Ferramentas da Semiótica Organizacional.

Durante a fase de planejamento do estudo de caso foram escolhidas quatro técnicas da Semiótica Organizacional. Cada uma delas apresenta características que consideramos importantes para a elicitación de requisitos de uma solução de Data Warehouse.

A seguir apresentamos a justificativa para a utilização de cada uma das técnicas escolhidas e quais aspectos da elicitación de requisitos essas técnicas deveriam cobrir:

Análise de Stakeholders:

Durante projetos de implementação de sistemas em grandes empresas, não apenas de soluções de Data Warehouse, mas de qualquer novo processo, é muito difícil conseguir que usuários chave, com conhecimento dos processos e de normas das empresas possam ficar focados no projeto da solução. No caso de ferramentas de apoio a negócio e da base de Data Warehouse que as apoiam, o cenário é mais complexo porque as pessoas que deveriam ser envolvidas geralmente são tomadores de decisão extremamente importantes para as empresas, com agendas de trabalho muito apertadas e que geralmente estão indisponíveis. Por esse cenário, muitas vezes o acompanhamento da implementação fica a cargo de colaboradores que detêm apenas visões particulares do processo, sem ter conhecimento do cenário como um todo.

Por isso, a análise de *stakeholders* poderia dar uma visão de toda a empresa, garantir que todas as áreas da mesma estariam bem representadas e que haveria conhecimento de todos

os processos exercidos por essas unidades. Além disso, a análise de *stakeholders* também dá uma visão da interação entre sistemas e sociedade, que poderia sugerir fontes de dados e do fluxo do mesmo entre os sistemas integrados.

Quadro de avaliação:

O quadro de avaliação, segundo as nossas expectativas iniciais, poderia evidenciar, além dos processos que necessariamente deveriam estar representados, os problemas enfrentados no dia a dia da empresa com a representação e com o fluxo de informações. Lembrando que um sistema de Data Warehouse não necessariamente tem uma maneira determinística de uso, essa fase ajudaria a evidenciar as necessidades dos usuários, que tipo de interação eles costumam ter com informações e qual a maneira que eles utilizam para manipular as informações. A análise do fluxo de informação pelos participantes também poderia ajudar a elucidar partes obscuras do processo e evidenciar caminhos mais otimizados para o fluxo de informação. A presença de colaboradores e representantes de todas as entidades evidenciadas na fase de *Análise de Stakeholders* seria de suma importância.

Framework Semiótico:

O Framework Semiótico foi considerado a principal ferramenta para a análise de requisitos de soluções baseadas em Data Warehouse. Isso porque as características do *framework* possibilitam tanto o estudo do fluxo e de necessidades de informação do ambiente que se estuda, bem como uma análise dos processos informais e intenções não registradas, principalmente por meio da camada pragmática da escada semiótica.

O uso do Framework Semiótico na fase inicial de estudo de viabilidade do projeto poderia nos indicar, através das camadas inferiores da escada, qual o tipo de tecnologia preferida da empresa e dar idéia de que aspectos de infra-estrutura computacionais seriam mais adequados no ambiente, facilitando assim a implementação e diminuindo o custo do projeto. Além disso, na fase de estudo de viabilidade, a parte superior da escada é de suma importância, pois pretendíamos que fossem evidenciadas as reais necessidades de fluxo,

métodos de acesso à informação dos usuários e o relacionamento desses usuários com a informação.

Análise Colateral:

A Análise Colateral foi considerada no método por prover o usuário com uma visão geral do que foi estudado, a fim de verificar se a compreensão do sistema por parte do usuário foi correta, bem como se os analistas responsáveis pela eliciação dos requisitos conseguiram compreender a dinâmica de trabalho da empresa, suas necessidades de informação e suas características próprias, como termos e jargões. Além disso, nessa fase poderia ser evidenciado o real desejo do grupo estudado em relação a uma solução computacional, quais as suas expectativas e, dependendo do resultado, fazer uma nova iteração de algumas das ferramentas, objetivando um ajuste fino das informações.

A Figura 12 mostra a sequência de técnicas sugerida.

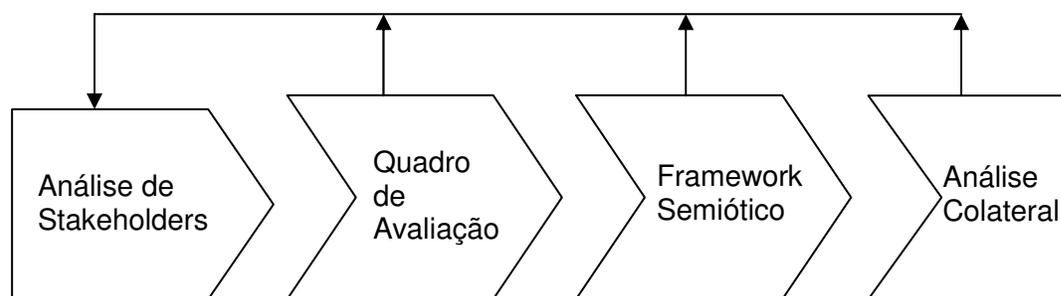


Figura 12 : Sequência de Técnicas sugerida

Nesse momento de análise de requisitos não estava prevista a eliciação de necessidades da camada de interface, apesar de que elementos dela já estivessem explícitos em algumas necessidades dos usuários que foram aparecendo conforme as ferramentas da técnica iam sendo utilizadas.

Descrevemos a seguir as seqüências de passos e informações esperadas em cada etapa do processo de elicitação de requisitos. Uma solução de Data Warehouse geralmente apresenta a estrutura mostrada na Figura 13 a seguir:

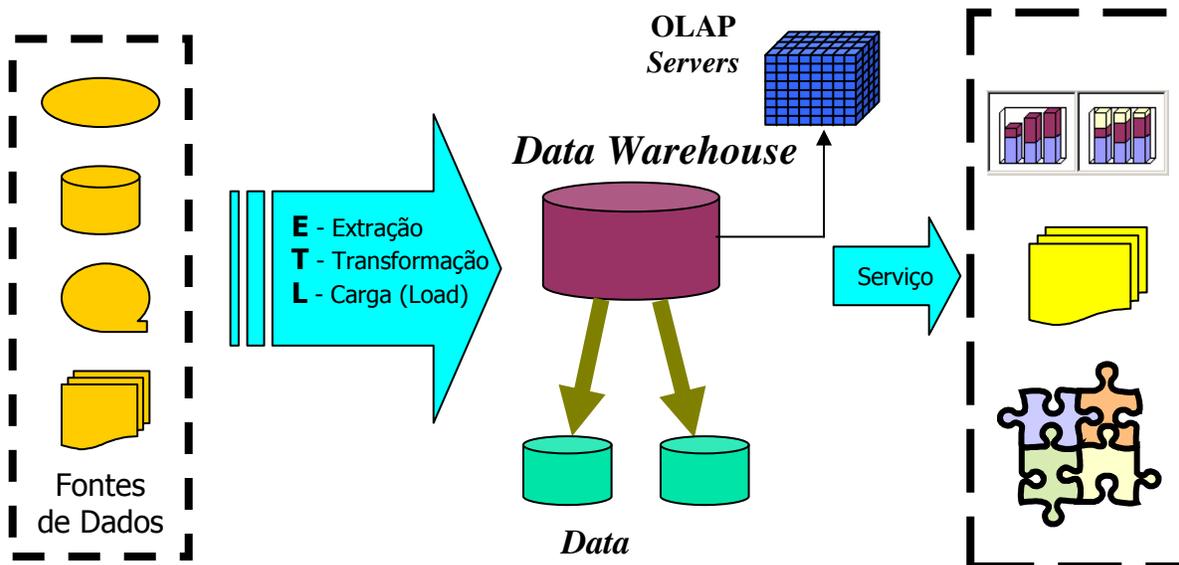


Figura 13 : Representação esquemática da solução final

O principal objetivo do emprego das técnicas da Semiótica Organizacional é o levantamento das necessidades dos usuários para a definição do esquema básico apresentado na figura 13.

O principal resultado esperado após a aplicação das técnicas da Semiótica Organizacional é um esboço da solução para o planejamento da implementação, bem como verificar se a empresa tem o nível de maturidade tecnológica e processual para poder implementar e usufruir de uma solução de Data Warehouse. Caso a empresa tenha problemas graves de processos, não haja definição do fluxo de dados ou problemas similares, o projeto de implementação de um Data Warehouse pode ser adiado até que os problemas sejam resolvidos e a empresa realmente tenha condições de fazer essa implementação. Dados os gastos de tempo e de recurso normalmente empregados em uma solução de Data Warehouse, uma definição precoce de problemas torna-se de suma importância.

Podemos representar a análise dimensional para o Data Warehouse através do modelo gráfico chamado *Star Schema*. Esse esquema gráfico mostra a determinação das informações que devem ser contidas nas *tabelas-fato* de um Data Warehouse, de onde as dimensões irão retirar suas informações e prover seus relacionamentos. Assim sendo, as dimensões necessárias elicítadas serão representadas dessa maneira, dando aos implementadores um caminho seguro a seguir. Nesse momento não pretendemos levantar todos os dados e metadados que irão compor os bancos de dados, mas sim evidenciar numa visão macro quais aspectos dos sistemas legados e de fonte os analistas seguirão para definir a estrutura final do banco de dados. A principal vantagem de se fazer esse diagrama utilizando o *Star Schema* é que essa representação não é atrelada a nenhuma das tecnologias em uso para Data Warehouse, podendo ser utilizada para a representação de sistemas que serão construídos tanto em Banco de dados Relacionais, OLAP , MOLAP ou qualquer outro.

A Figura 14 representa o fluxo de análise das ferramentas e os resultados esperados pelo emprego de cada técnica.

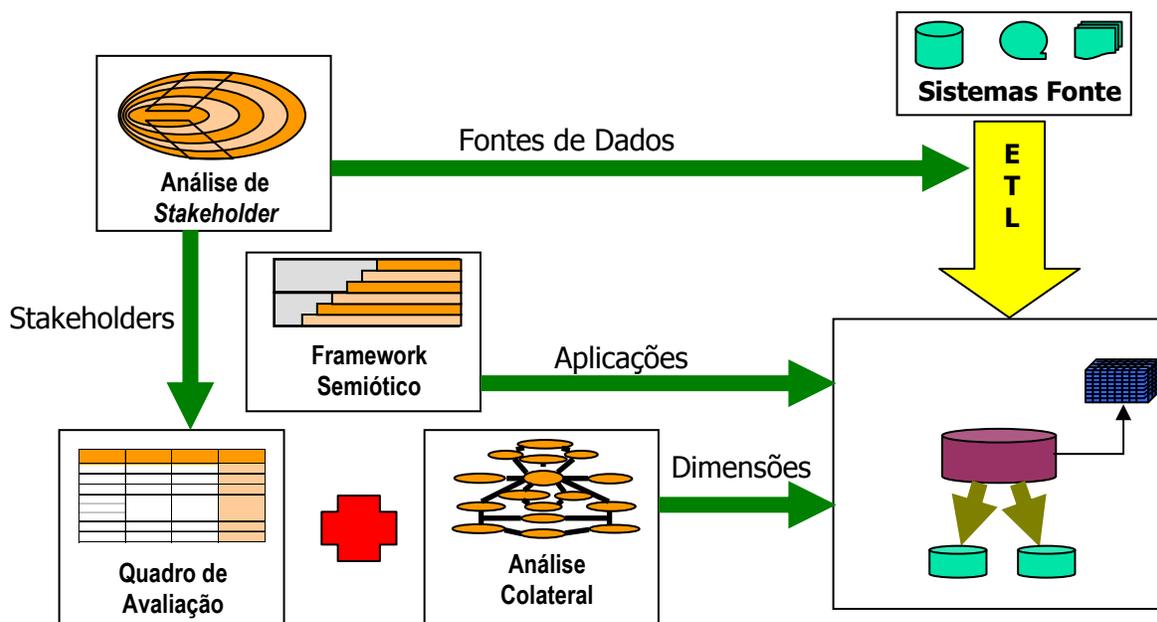


Figura 14 : Diagrama de uso dos resultados das técnicas empregadas

O processo inicia-se com o estudo dos dados levantados na etapa de *Análise de Stakeholders*. As entidades representadas na camada de contribuição serão as pessoas ou unidades que deverão ser consultadas para os levantamentos das necessidades do sistema. São desses usuários que os métodos de trabalho e as idéias de qualidade de dados e necessidades de informação serão levantadas, a fim de representá-las no sistema final.

Ainda avaliando as Partes Interessadas, temos as camadas de Mercado, Comunidade e Fonte. As camadas de Comunidade e Mercado podem dar pistas de que tipo de concorrência e atuação são esperados da empresa, mas o ponto mais importante é a camada de Fonte. As fontes de dados em uma solução de Data Warehouse são de extrema importância e qualquer falta de informação, inconsistência ou falta de homogeneidade entre os metadados das informações podem causar efeitos catastróficos durante o projeto e o uso do produto final.

As entidades presentes na camada Fonte podem ser levadas diretamente ao projeto do ETL. Com essas informações iremos popular a primeira camada da solução e saberemos o que verificar e perguntar em etapas mais aprofundadas do levantamento. Geralmente essas fontes são sistemas computacionais legados, mas podem ser qualquer tipo de informação existente na empresa. Esse levantamento precoce dos sistemas envolvidos é de grande importância para evitar que a falta de informações cause impacto em fases posteriores. Como principais vantagens do uso do método nesse momento e os resultados esperados são:

- Levantamento de todas as fontes de informação existentes para o sistema.
- Definição de quais informações não estão em ambientes computacionais integráveis (relatórios em papel, jornais, outros tipos de mídias), que precisarão definições de uso e carga
- Visibilidade das necessidades de homogeneização de dados e de metadados das aplicações
- Possíveis dificuldades com dados necessários mas inexistentes

- Quantidade de fontes diferentes e volume de tráfego de informações
- Pré-visualização de quais informações são importantes para um Sistema de Apoio a Decisão

O último item remete à necessidade de se definir granularidade e universo de abrangência dos dados. Um Sistema de Apoio a Decisão não deve nem pode conter todos os dados dos ambientes relacionais, pois senão seria apenas um repositório central de dados, descaracterizando a real vantagem desse tipo de sistema. Pessoas que necessitem de um grau de informação muito detalhado devem ir direto aos sistemas de informação relacionais, isso é fato. Um Data Warehouse com excesso de informações não terá a agilidade necessária a tomadas de decisão de momento, praxe nos ambientes empresariais de hoje. O impacto no desempenho de excesso de dados que não serão utilizados, tanto durante o uso, quanto nas janelas de carga, podem aumentar muito o custo da aplicação e inviabilizar o processo.

Os dados levantados nessa fase devem ser utilizados como base para a próxima etapa do processo. É importante frisar que a qualquer momento podem surgir novos dados, processos, fontes e usuários no processo, de modo que iterações podem ser necessárias, a critério do profissional que está utilizando a técnica. Essas iterações devem ser bastante comuns e não devem ser encaradas como problemas, pois a tendência é conforme a técnica é utilizada há um entendimento maior do processo pelo próprio usuário e idéias, sugestões e problemas que nunca haviam sido levantados podem ser reconhecidos.

A próxima técnica a ser empregada, o Quadro de Avaliação, invoca as entidades elicitadas na fase de *Análise de Stakeholders*, colocando em debate qual o seu papel nos processos, quais os problemas encontrados nesses processos e quais possíveis soluções os usuários gostariam de ver implementadas. Nesse ponto, devemos ter cuidado mais uma vez com o principal ponto de um sistema de suporte à decisão: ele é orientado a dados, logo o que devemos perguntar nessa fase é:

- Quais as reais necessidades de dados da empresa?

- Quem são as pessoas que dependem desses dados e para que eles os utilizam?
- Os dados apresentados hoje são completos?
- Os dados presentes nos sistemas hoje têm boa qualidade?
- Qual a interpretação da empresa para “Qualidade de dados”?
- Há alguma ruptura no fluxo de informações?
- Quais as áreas problemáticas em relação ao envio de dados?
- Como as áreas entendem e manipulam seus dados?
- Qual o perfil de uso e de análise desses dados?
- Quais as dificuldades encontradas para essa análise?

Essas perguntas não devem ser colocadas diretamente para não influenciar as respostas, mas devem servir como guia para que a atividade cumpra seu papel essencial. Dependendo do tipo de público e da experiência dos usuários com projetos de informática ou com processos computacionais, pode haver uma tendência a se entrar em discussões técnicas a respeito das atividades. Não é esse o objetivo da técnica nesse momento. O empreendedor dessa análise deve tentar entender quais os processos essenciais e as dificuldades correntes neles, bem como tentar homogeneizar o entendimento dos setores, em especial nas áreas de interface. Se possível, o perfil das pessoas que farão o uso da solução final deve ser analisado.

Depois do estudo dos resultados dos problemas levantados no Quadro de Avaliação, o analista deverá ser capaz de começar a enxergar os padrões, problemas e necessidades dos usuários, ainda de forma não estruturada.

A próxima etapa é o uso da Escada Semiótica para a avaliação geral. Essa etapa é a mais importante do processo de elicitação, pois dela sairão as aplicações que serão transformadas em *tabelas-fato* e regerão a criação das dimensões. Como foi visto na **seção 4.2**, o *Framework Semiótico* é dividida em seis camadas, e as que mais nos interessam são as camadas **Pragmática** e **Social**. A camada Pragmática representa as intenções, determinações das pessoas e entidades, podendo inclusive evidenciar fatos que não seriam considerados relevantes num método tradicional de elicitação e requisitos. As camadas

inferiores, relacionadas ao mundo físico, podem ser trabalhadas, geram informações extremamente relevantes, mas para a verificação das necessidades de dados da empresa elas não influenciariam muito no resultado.

As reais necessidades da empresa devem ser evidenciadas nessas duas camadas. O consultor deve identificar e dividir as necessidades conforme grupos bem definidos, que serão utilizados para as *tabelas-fato* e para a análise *dimensional*.

Podemos exemplificar o processo com o seguinte caso fictício. Digamos que o processo esteja ocorrendo numa empresa fabricante de determinado produto. As etapas vão se sucedendo e chega-se ao ponto do uso do Framework Semiótico. Dentro da camada Pragmática, entende-se que a maior necessidade da empresa é o controle do custo produtivo e o controle do fluxo de vendas. O analista, desse modo, teria em suas mãos os dois principais processos que orientariam as próximas tomadas de decisão da empresa. Esses dois processos seriam as bases para as *tabela-fato*.

Analisando agora os resultados dos Quadros de Avaliação, os problemas e soluções levantados deveriam ser distribuídos entre esses dois processos. Eles não são mutuamente excludentes, então uma informação pode ser parte dos dois. Quando todas as informações forem distribuídas, as análises e agrupamentos das mesmas irão nos informar as *dimensões* da aplicação e nos indicar quais as informações a disposição nas entidades fonte devem ser levadas a cada uma das aplicações. Essa definição é o resultado final da análise preliminar e objeto final do estudo.

Depois da utilização do Framework Semiótico, a próxima ferramenta utilizada é a Análise Colateral. Essa etapa deve dar uma idéia ao analista sobre o que a empresa enxerga como sua real necessidade, o que ele entendeu do sistema proposto e como ele acha que será a evolução do sistema e a sua interação com o meio. Além disso, em etapas mais técnicas, informações como necessidades de *backup* e *recovery* são explicitadas.

Para a confecção do desenho inicial do *Star Schema*, todos os dados levantados serão verificados, mas a Escada Semiótica deverá ser alvo de uma análise mais profunda. Os dois primeiros degraus, o que representa o nível pragmático e o que representa o nível social, deverão indicar quais “aplicações”, ou seja, quais *tabelas-fato* iremos utilizar, para quais processos específicos. Esses processos servirão de base para podermos fazer o estudo que indicará quais *dimensões* devem ser agregadas ao modelo. Cabe ao analista conseguir diferenciar e ter discernimento sobre isso, também com base no conhecimento adquirido durante as sessões de trabalho. Mesmo se um determinado processo não tenha sido explicitado pelos usuários, mas pelas conversas e dados anteriores o analista achar que o processo deve ser analisado, ele deve ser colocado em pauta e revisto.

Continuando o exemplo hipotético, digamos que o analista consiga duas aplicações a partir da análise da Escada Semiótica. Essas duas aplicações devem ser colocadas no centro de um gráfico, representando a *tabela-fato* do *Star Schema*. Cada aplicação gera um *Star Schema* diferente. Para melhor representação, os processos devem ter nomes relacionados com o processo ou intenção do grupo trabalhado. Apenas como exemplo, vamos dizer que o levantamento esteja sendo feito para uma rede de supermercados, com estoque central e com várias lojas. A verificação da Escada Semiótica indica que os diretores querem um melhor controle do estoque, para melhorar a logística e um melhor controle de vendas por supermercado. Assim, podemos dividir o trabalho por duas aplicações diferentes, ESTOQUE e LOJAS.

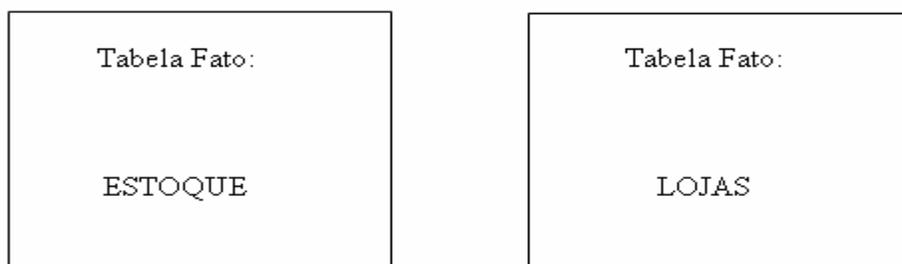


Figura 15 : Aplicações representadas por *tabelas-fato*

Após a definição das aplicações, devemos percorrer todos os dados levantados nas outras sessões para avaliar quais dimensões seriam necessárias para conter os dados requeridos para os cruzamentos de informação dos usuários. Podemos fazer isso percorrendo principalmente as colunas de solução do Quadro de Avaliação, que é a atividade que irá indicar como os usuários gostariam que as atividades fossem feitas. Desse modo, e sempre contando com a experiência do analista, podemos separar os problemas relacionados com o Data Warehouse, desprezando alguns que não tenham relacionamento com o âmbito de dados. Devemos lembrar que apesar dos processos serem extremamente importantes para a geração e coleta de dados, um Data Warehouse é orientado a dados e não pode ajudar em problemas exemplificados como “Aumentar o moral da equipe”. Ele pode sim ter um indicador que mais tarde em um sistema de avaliação, como por exemplo uma ferramenta de *Balance Scorecard*, possa ter isso como meta, mas não deve entrar como problema direto em um Data Warehouse.

Deve ser feita uma lista indicando qual a solução, a qual aplicação pertence, se a uma ou a ambos. A análise desses dados deve indicar ao analista quais as dimensões de importância, pois ela deve estar destacada na frase, expressando o sentido principal dela. Por exemplo, se a solução estiver em uma frase do tipo “Verificar quais lojas têm maior índice de vendas por período”, podemos notar que na aplicação LOJA deve ter uma dimensão VENDAS e TEMPO, para conter as informações desejadas. Após essa análise do Quadro de Avaliação principalmente, devemos estar aptos a montar o mapa inicial das *tabelas-fato* e *dimensões*.

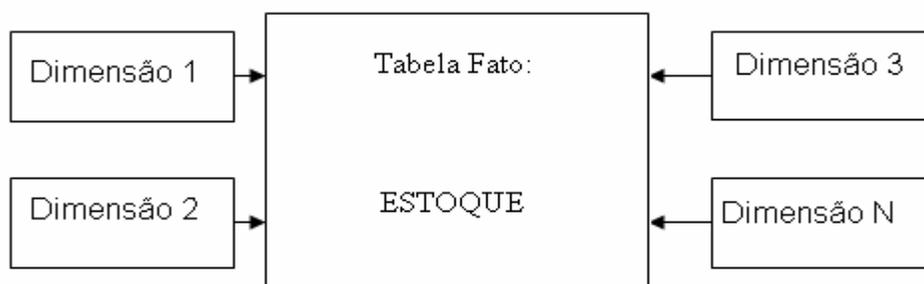


Figura 16 : Representação de uma aplicação e suas dimensões

Essa seqüência de atividades e de análises foi empregada no estudo de caso instanciado, descrito na seção 5.3.

5.2 Descrevendo a Organização e o Contexto

5.2.1 A definição da organização a ser estudada

A definição do objeto de análise para o estudo de caso levou em considerações alguns aspectos para se eleger a melhor entidade para estudar durante o trabalho. A entidade em questão não poderia ser muito pequena, pois o volume de dados de uma empresa influencia diretamente em uma solução de Data Warehouse e o acesso à informações deveria ser o mais livre possível. Essa última necessidade seria a mais difícil de cumprir, pois os dados de um Data Warehouse são dados estratégicos, que espelham a saúde e os objetivos da empresa. Sendo assim, o grau de acesso às pessoas e às informações seria relevante para a decisão de qual entidade seria utilizada nos estudos.

Surgiu a possibilidade de realizar esse estudo na Secretaria de Vigilância Sanitária do Estado de São Paulo – DIR I – Capital. DIR I significa Diretório Regional. Doravante chamaremos a entidade apenas como DIR. Para se chegar a essa conclusão, fiz uma série de visitas e verificações na maneira como eles trabalhavam a informação, o volume de dados e qual o tipo de problemas que poderiam ser encontrados na DIR e que poderiam demonstrar a capacidade das ferramentas da Semiótica Organizacional para a elicitación de requisitos de uma possível solução de Data Warehouse.

Na primeira reunião que tive com a Diretora da DIR ficou evidente como as diferentes interpretações de simples palavras e termos podem levar a resultados totalmente diferentes, mostrando assim a importância de um ferramental como o da Semiótica Organizacional. Nessa reunião eu estava tentando explicar para a Diretora quais eram as minhas intenções para o trabalho. Durante a conversa, sem dar maiores detalhes do que era um Data Warehouse, pois não cabia na situação e poderia influenciar nos resultados futuros, eu

tentei exemplificar a necessidade de informações de um núcleo de trabalho como o deles, utilizando exatamente essas palavras. Notei que a diretora ficou ofendida e, conforme ela me explicou, eles eram uma “unidade” e não um “núcleo”. Mais tarde entendi que no jargão deles um “núcleo” era um grupo menor, algo como apenas um centro de saúde e, ao chamá-los assim, houve uma conotação de diminuição do trabalho deles. Felizmente, a situação pôde ser revertida, mas esse caso serve bem para exemplificar como o sentido e significados das palavras e símbolos podem afetar de modo crucial a interação dos analistas com os grupos de pessoas cujo trabalho está sendo elicitado.

A decisão final da possibilidade de trabalho ficou com a representante do CVS (Centro de Vigilância Sanitária), órgão paralelo à Secretaria e que mantém os sistemas computacionais empregados pela DIR. Marquei com a analista de sistemas responsável pelo suporte e operação desses sistemas para conversarmos e para expor as minhas necessidades e ela repassou meus pedidos e explicações para a sua chefia imediata, que permitiu o trabalho.

Essa reunião com a analista de sistemas foi bem interessante, porque algumas de minhas suspeitas se confirmaram. A analista contou-me de casos prévios e consultorias de informática que haviam tentado fazer trabalhos de levantamento de necessidades de sistemas e processos, sempre sem resultado prático algum. O principal motivo, segundo ela, era a dificuldade de entendimento da complexa estrutura e das motivações da DIR.

Nessa mesma reunião ela se comprometeu a me ajudar com os levantamentos e eu mostraria a ela as técnicas da Semiótica e explicaria o que era um Data Warehouse, tecnologia que ela não dominava. Essa analista foi de grande valia ao processo e conseguiu ter um bom entendimento das ferramentas utilizadas.

Foi feita ainda mais uma reunião, na qual a analista me mostrou as ferramentas utilizadas por eles para o acompanhamento do ciclo de vida dos processos. Eram ferramentas feitas pelo próprio CVS, utilizando a linguagem de programação *Visual Basic*, sem qualquer tipo de integração entre essas ferramentas e com a base de dados em *Access*. Esses sistemas, conhecidos como SIVISA (Sistema de Vigilância Sanitária) e SIAP (Sistema de Informação de atendimento a Público) não dispunham de ferramentas computacionais otimizadas para ajudar no fluxo de informações das diferentes DIR's, e nem para a análise e

garantia da qualidade dos dados, servindo quase somente como um banco de dados de processos.

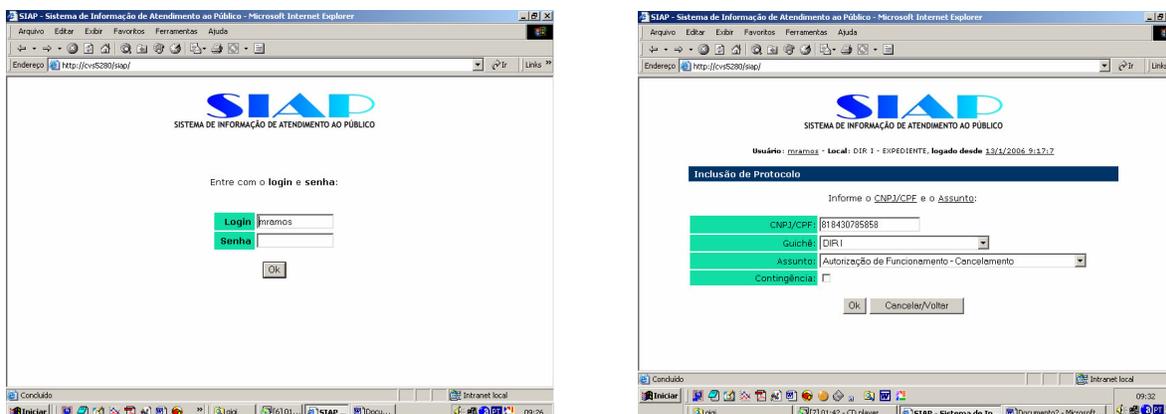


Figura 17 : Exemplos de Telas do Sistema SIAP

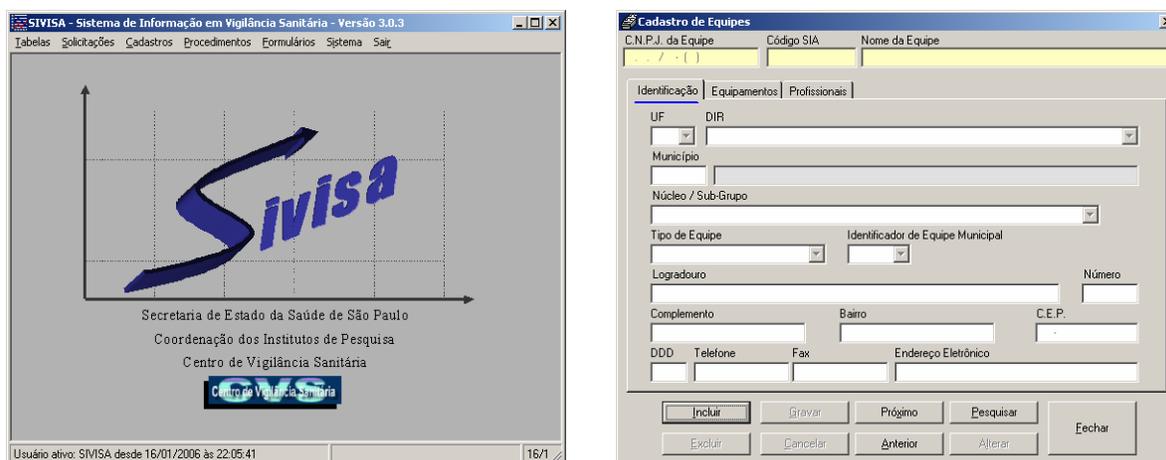


Figura 18 : Exemplos de Telas do Sistema SIVISA

Ainda assim, o volume de dados que transita na DIR é enorme, devido ao tamanho de sua operação. Com certeza existe a necessidade de um maior apoio computacional, e a minha pesquisa poderia mostrar qual é esse apoio e quais as reais necessidades de qualidade de informação e como essa informação poderia ser manipulada.

5.2.2 A DIR I e o seu papel na área de Saúde

A regulamentação dos órgãos de saúde do Estado de São Paulo depende de uma série de regulamentações em instâncias superiores, entre as quais a Constituição Federal, as Leis Orgânicas de Saúde, dos códigos de saúde do Estado de São Paulo e de uma série de outras decisões e normas, todas elas geralmente editadas através do Diário Oficial da União. Nos últimos anos, o setor Saúde vem passando por um profundo processo de transformação processual e organizacional, através da consagração e estabelecimento do Sistema Único de Saúde – SUS.

O SUS é uma série de medidas que servem para garantir que haja um conjunto de ações e intervenções que possam prover as condições necessárias para que a saúde da população seja protegida, promovida e, se necessário, recuperada. As ações da Vigilância Sanitária são de suma importância para esse processo, em todas as três esferas de governo, federal, estadual e municipal, sendo as suas ações e atribuições descritas na Constituição.

Na Lei Orgânica da Saúde, Lei Federal 8080, de 1990, está prevista a competência da Vigilância Sanitária - VISA. O artigo 6º, inciso XI § 1º declama:

“Entende-se por vigilância sanitária um conjunto de ações capaz de eliminar, diminuir ou prevenir riscos à saúde e de intervir nos problemas decorrentes do meio ambiente, da produção e circulação de bens e da prestação de serviços de interesse da saúde.”

Ou seja, a Vigilância Sanitária é mantenedora de toda a responsabilidade do Governo, nas três esferas de atuação, sobre as ações que incidam sobre a saúde da população. A ação da Vigilância está bem acima do que apenas a simples aplicação de leis, ela é responsável também pela prevenção formal e material da saúde pública. Entre as suas atividades, mas não resumidas a elas, estão a expedição de licenças de funcionamento, registro de produtos, autorizações de funcionamento de empresas, autos de infração, penalizações e multas por problemas relativos à saúde. Esses documentos são muito importantes para a garantia de funcionamento de empresas competentes e que cumpram a lei de saúde, mas como os

próprios participantes do trabalho frisaram, as ações da DIR não são meramente cartoriais, não apenas distribuem licenças e documentos. A função mais conhecida na ação da vigilância é seu aspecto fiscalizador e de “polícia” das questões de saúde, sempre procurando pontos onde tenham ocorrido falhas no cumprimento das leis sanitárias. Somente por essas funções colocadas, já podemos ter idéia do volume de dados, processos e informações que navegam pela DIR todos os dias e a sua importância para a vida da população como um todo.

Nos últimos anos o enfoque do modelo de saúde no Brasil tem mudado. Ao contrário do modelo anterior, que era puramente reativo, atualmente tenta-se tomar ações de cunho preventivo, priorizando ações e intervenções baseadas no risco inerente à saúde. Essas ações partem do pressuposto de que os serviços e produtos de saúde devem ser oferecidos em quantidade e qualidade adequadas às necessidades da população.

Podemos dizer, partindo das determinações formais existentes na Constituição e nas Leis estaduais e municipais, que o campo de atuação da vigilância sanitária está dividido em quatro áreas que se inter-relacionam e que compõem as unidades da DIR:

1. Controle de produtos que direta ou indiretamente se relacionam à saúde;
2. Controle da prestação de serviços que se relacionam direta ou indiretamente com a saúde;
3. Controle sobre o meio ambiente;
4. Controle sobre o processo de trabalho de qualquer natureza (ações de saúde do trabalhador).

No item 1 estão incluídas as indústrias de medicamentos, de produtos médicos hospitalares, de alimentos, farmácias de manipulação, drogarias, distribuidoras de produtos (medicamentos, produtos médicos hospitalares, alimentos), importadoras de produtos, venda direta ao consumidor de alimentos (supermercados, mercados, bares e restaurantes); as cozinhas privativas de empresas e as hospitalares isoladas, serviços de esterilização.

No item 2 estão: os hospitais, os ambulatórios médicos, as clínicas de cirurgia ambulatorial (em especial as clínicas de cirurgia plástica e estética), os consultórios médicos isolados, as clínicas odontológicas, os laboratórios de análises clínicas, os laboratórios de patologia clínica, os hemocentros, as agências transfusionais, os serviços de hemoterapia, os serviços de diálise e hemodiálise, os serviços de quimioterapia, os centros de diagnóstico por imagem, os serviços de medicina nuclear, as casas de repouso, os serviços de transporte de pacientes, os serviços de *home care*, os consultórios de psicologia, nutrição e dietética, as lavanderias hospitalares isoladas.

No item 3 estão: gerenciamento de resíduos de serviços de saúde, áreas contaminadas, programa da qualidade da água, criação de animais, depósitos de sucatas, telefonia celular, cemitérios, aterros sanitários (deposição de lixo como o aterro Bandeirantes), usinas de tratamento de resíduos (incineração, microondas).

No item 4 estão: ambiente de trabalho (insalubridade), verificação se existe Comissão Interna de Prevenção a Acidentes, SESMAT (Serviços Especializados em Engenharia e Medicina do Trabalho), controles de saúde como exames periódicos, vacinação, queixas frequentes, ergonomia, locais potencialmente prejudiciais à saúde do trabalhador (amianto, marmoraria, etc).

Para garantir o respeito às normas de cada órgão e situação acima, várias medidas são tomadas pelas equipes das DIR, entre elas análises de documentos, vistorias no local, análises laboratoriais, pesquisas em textos científicos, entre outros. A autoridade sanitária pode escolher quais métodos e ações serão tomados caso a caso e a sua autoridade é total em sua resolução.

Ainda segundo a Constituição, os atos de um servidor da saúde devem ser sempre corroborados por lei, e seus atos justificáveis por necessidades maiores da população. Assim sendo, ele sempre deve obedecer a princípios que garantam que as suas ações são legais e necessárias, sob pena de prisão, inclusive. Além disso, ele é obrigado a denunciar e registrar quaisquer possíveis causas de dano à saúde da população, podendo ser julgado por prevaricação se não o fizer. Dessa forma, notamos mais uma vez a necessidade de um

controle rígido de informações e processos, pois pode ocorrer a necessidade do técnico de saúde se defender contra acusações sobre os seus procedimentos, do mesmo modo que pode existir necessidade da geração de provas sobre um determinado ato lesivo à saúde da população, e essas decisões devem ser tomadas no menor tempo possível, pois agravos à saúde da população podem aumentar exponencialmente tornando-se casos de epidemias graves.

Após os contatos iniciais, foi pedido ao grupo que iria começar as atividades que fizessem uma descrição sucinta do histórico da DIR e de uma rápida opinião a respeito dos seus problemas em relação à informação e suas críticas aos sistemas atuais. O material gentilmente produzido por eles está reproduzido no Anexo A. Os principais problemas levantados por esse documento são:

- Falta de recursos operacionais e de apoio;
- Pouca confiabilidade da infraestrutura de informática atual;
- Falta de políticas de recuperação de dados;
- Histórico de dados confuso por problemas de consistência de dados e de carga;
- Registros atuais ainda feitos sem padronização;
- Sistemas de informática atuais sem troca de informações (SIVISA e SIAP);
- Várias instâncias dos sistemas, gerando diferentes arquivos de dados distribuídos que nem sempre são agregados de forma correta ou em tempo hábil;
- Dificuldade de planejamento estratégico causada pela falta de acesso à informação clara.

5.3 O Estudo de Caso

Utilizamos para as reuniões, ou *workshops*, do estudo de caso as instalações da própria DIR, localizada na tradicional Avenida São Luiz, no centro velho de São Paulo, número 99, no prédio cordialmente conhecido como “Palácio da Saúde”. Nesse prédio funcionam

diversos órgãos que cuidam de vários aspectos relativos à saúde da população e é para onde deve se dirigir uma pessoa em busca de ajuda, que deseje fazer uma denúncia ou que queira orientação sobre um determinado procedimento.

Ao todo seriam feitas quatro reuniões em dias diferentes, cada uma delas para a utilização de uma das técnicas da Semiótica Organizacional. Essas reuniões seriam feitas nas dependências da própria DIR para evitar transtornos aos participantes.

As instalações físicas se mostraram adequadas ao processo, possibilitando algum nível de privacidade durante as reuniões e com bom espaço físico para o material de apoio. Geralmente utilizamos a sala de reunião presente no andar, mas quando não foi possível nos alojamos na sala da Diretora.

A seguir descrevemos cada uma das reuniões, seus objetivos, técnicas utilizadas e resultados.

Reunião 1 – Apresentação da proposta aos participantes e Uso da Análise de Stakeholders

A avaliação da empresa que quer implementar uma solução de Data Warehouse deve começar com a utilização da análise de *stakeholders*. Para essa sessão, devem ser chamados representantes de todos os setores iniciais que estejam envolvidos nos processos e fluxos de informação que a empresa quer sistematizar. Aqui há dois pontos a serem considerados:

1. Apesar de ter uma idéia do que quer, a empresa pode não estar familiarizada com as necessidades de uma solução de Data Warehouse e o departamento que teve a iniciativa de começar os estudos para verificar a viabilidade do projeto pode não ter conhecimento de todas as interfaces de comunicação envolvidas nos processos. Assim sendo, é muito provável que nesse primeiro momento não estejam presentes todas as pessoas chaves para um bom levantamento do processo.

2. Uma solução de Data Warehouse é direcionada para os tomadores de decisão de uma empresa. Isso quer dizer que geralmente uma solução dessas será útil para pessoas de nível hierárquico elevado ou que tenham cargos que exijam uma visão generalista da empresa, como por exemplo, consultores estratégicos. Esse tipo de público geralmente não está prontamente disponível, sendo que conflitos de agenda são típicos. É muito importante que, até a definição de quem realmente detêm o conhecimento dos processos e para qual perfil de uso a solução será desenvolvida, as sessões sejam agendadas com antecedência e que se tente um compromisso com instâncias superiores dentro da corporação, a fim de garantir a presença das pessoas chave ao processo de elicitação.

Com o exposto acima, podemos ver que a primeira reunião de análise de stakeholders é de suma importância. É durante essa sessão que poderemos avaliar todas as interfaces de comunicação e tentar que as peças fundamentais assumam compromisso com o andamento da análise. Com base nessas considerações, para essa reunião ficou decidido que:

- Seriam convidadas pessoas de todo o organograma básico da DIR, no nível gerencial;
- Não seria feita referência ao Data Warehouse diretamente;
- A reunião começaria com uma explicação dos meus objetivos;

Participaram dessa primeira etapa ao todo seis usuários, representando as diversas camadas hierárquicas da DIR:

- Diretora;
- 2 Assistentes Técnicas da Diretora, uma da área jurídica;
- 3 diretores de área, sendo elas projetos (assistente técnica), meio ambiente e produtos;
- A analista de sistemas do CVS.

Com essas pessoas, esperava ter a total representatividade dos grupos de trabalho da DIR e a Análise de *Stakeholders* poderia confirmar isso.

Inicialmente, nenhuma das pessoas presentes tinha conhecimento prévio sobre soluções de Data Warehouse e nunca havia participado de sessões de elicitação de requisitos onde tiveram sido utilizadas as técnicas da Semiótica Organizacional. Notei que alguns deles, incluindo a própria analista de sistemas do CVS que me acompanhava, já haviam participado de algum tipo de levantamento sobre sistema de informação, pois durante o começo dos trabalhos notei certa “contaminação” de alguns usuários, caracterizada por perguntas sobre quais questionários iríamos utilizar, qual o tipo do programa, em que computador rodava, ou seja, perguntas que davam a entender que já haviam participado de algum tipo de projeto de software.

A reunião aconteceu no período da manhã e teve duração aproximada de duas horas. Essa reunião foi gravada com um gravador portátil, com conhecimento dos participantes.

Os primeiros momentos dessa reunião foram utilizados para apresentar aos participantes quais seriam as atividades envolvidas e porque elas seriam feitas. Houve uma pequena explicação da técnica da Semiótica utilizada e que gostaria que nos concentrássemos nos problemas existentes hoje com o fluxo de dados e problemas com acesso a informações. Uma explicação mais detalhada sobre Semiótica Organizacional, Data Warehouses e sobre a técnica que utilizaríamos no dia foi dada apenas para a analista de sistemas do CVS, conforme acordo prévio.

Foi utilizado como material de apoio a figura em camadas representando a “cebola” dos *Stakeholders*. A Análise de *Stakeholders* é representada em forma de "cebola" por causa dos diferentes níveis de aninhamento do sistema técnico, do nível mais interno até o nível mais externo, onde as implicações sociais do sistema técnico podem refletir em alguns pontos que indiretamente influenciariam nas tomadas de decisão.

Essa estrutura organiza o levantamento das "partes interessadas" no sistema de Data Warehouse. Como apresentado anteriormente, soluções de Data Warehouse são utilizadas para auxiliar no estudo dos processos dentro de uma entidade e de análise de seu negócio e objetivos. Levando em consideração esses aspectos, a avaliação das camadas nesse artefato deve ser tomada da seguinte forma:

Comunidade: Essa camada mais externa nos mostra agentes que influenciam indiretamente nos sistemas, na empresa e/ou no seu resultado. Exemplo disso pode ser o Governo, Imprensa, etc;

Mercado: Como visto no Capítulo 4, a camada de mercado evidencia entidades que servem como referencial para a organização, sendo por comparação ou por competição. Ela pode indicar entidades de servem como modelo, concorrentes ou empresas similares. Apesar de poder servir como entradas para elementos de *benchmark*, não podem ser consideradas como influenciadoras diretas em um sistema de Data Warehouse.

As camadas mais internas, que efetivamente irão fazer parte do fluxo de informação são:

Fonte: Essa camada é a principal área de interface do fluxo de informação da empresa. Os sistemas, organizações ou departamentos que fazem parte desse nível da “cebola” tomam parte importante no processo, sendo fontes ou sorvedouros de informação, ou então disponibilizando dados vitais para o fluxo e transformação das informações. Geralmente são agentes que pedem relatórios, absorvem informação ou, no caso dos próprios sistemas transacionais da empresa, que fornecem a informação em níveis de granularidade bem baixos, informação essa que será trabalhada, agregada e transformada numa forma de apresentação que possibilitará a interpolação de informação na forma desejada pela empresa.

Contribuição: São as áreas que efetivamente irão utilizar e alimentar o sistema, o seu principal público alvo. Mesmo que uma entidade das camadas superiores necessite de uma informação do sistema, é um representante dessa área que irá suprir essa informação retirada do sistema, bem como serão as necessidades desse grupo que serão espelhadas no sistema final.

As principais informações retiradas dessa primeira sessão, pela camada de contribuição, são todas as pessoas-chave para o processo, de forma que todas as áreas fiquem bem

representadas. Nesse momento, devemos tentar descobrir quais as pessoas que cobrem todo o conhecimento do fluxo de informações, quais trabalham com que tipo de resultado e quais as interfaces necessárias. Durante a análise, é esperado que os participantes notem áreas que não estão sendo representadas, e com isso processos dos quais não se pode ter a informação total. Nesse momento, deve-se fazer uma anotação e tentar uma reunião futura, convidando-se os representantes das novas áreas e explicando-lhes a sua importância no processo. Essas iterações são necessárias para garantir a representação de todas as áreas e evitar ausências sobre o processo. Deve-se tomar nota, também durante esse processo, quais as pessoas que podem ser "substituídas" por outras com maior disponibilidade de tempo.

A camada de fonte contribui de forma essencial ao projeto, mas de maneira diferente. Podemos considerar a camada de fonte realmente como onde ficarão as entidades originadoras e sorvedouras de informação. Dessa camada apresentará o começo de todo o fluxo de informação que irá passar pela camada de ETL (camada de homogeneização e transformação dos dados), ou seja, a partir dela que serão elicitadas todas as fontes principais de dados do sistema, quais as transformações necessárias e quais tipos de dados deverão ser inseridos no Data Warehouse. Apesar dos metadados serem muito importantes para a camada de ETL, não será essa análise inicial que irá levá-los, mas sim dará uma pista de quais sistemas deverão ser estudados para garantir que os dados sigam homogêneos aos repositórios do Data Warehouse.

Mesmo após o término dessas sessões de análise de *stakeholders*, a qualquer momento ela pode ser revisada. Caso ocorra, durante as fases posteriores, o consenso de que faltou alguma entidade fonte ou contribuinte, é aconselhável reiniciar o processo a fim de garantir a representatividade dessas áreas na solução final. Esse cuidado deve ser tomado por causa da forma como modificações posteriores ao projeto são dificultadas pelas próprias características das soluções de Data Warehouse como, por exemplo, grande quantidade de dados, grande número de cálculos e transformações de dados e interfaces entre sistemas.

Após as explicações iniciais, os participantes começaram a descrever cada uma das camadas de informação que eles necessitavam, e num primeiro momento indicavam que tudo o que eles queriam estava presente no sistema interno *SIVISA* (**S**istema de **I**nformação em **V**igilância **S**Anitária). Entregaram-me uma apostila com uma portaria de lei com a criação desse programa, com a descrição de algumas de suas funções, códigos e tabelas internas. Essa apostila indica basicamente os códigos de operação e as tabelas internas, sem maiores explicações de uso e objetivo.

Houve certa dificuldade de manter o foco no levantamento das necessidades do fluxo de informação. As pessoas presentes começaram a discorrer sobre uma série de problemas que não faziam parte diretamente do problema de fluxo e uso de informação como, por exemplo, problemas de treinamento, instalações físicas e dificuldades de interação com outras esferas de governo. Alguns ajustes no rumo da reunião foram necessários, a fim de evitar que a atividade fosse prejudicada.

De uma maneira geral, essa primeira reunião me forneceu uma idéia muito boa da maneira de trabalho deles e conforme a atividade foi sendo feita os próprios usuários notaram uma série de problemas e falta de informações, principalmente de acompanhamento do ciclo de vida de um processo e de seu histórico. Há pontos em que se percebem alguns problemas com a qualidade das informações apresentadas, como por exemplo a falta de homogeneização dos relatórios dos técnicos.

Ao final da atividade, o resultado foi documentado conforme ilustra a Figura 19.

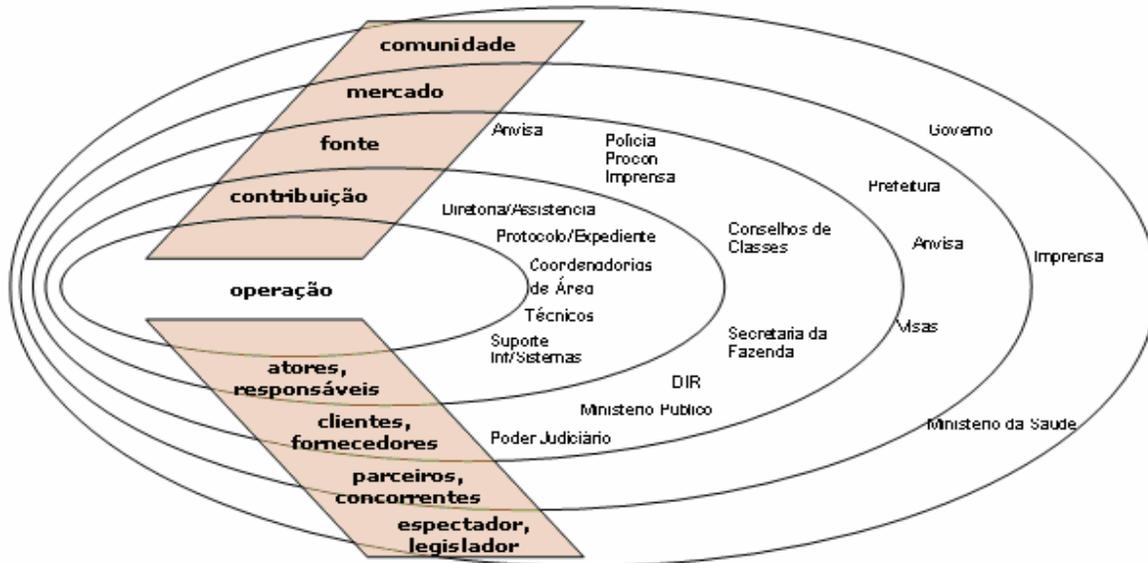


Figura 19 : Resultado da Análise de Stakeholders

Cada uma dessas partes interessadas interage de uma forma específica com o trabalho final da DIR. Dessa forma, a representação de cada um deles é extremamente importante. A análise desse primeiro resultado indicou que as unidades da DIR estavam bem representadas, pois a área de protocolo, que não tinha um participante direto, poderia ser representada pela Assistente Técnica Jurídica, que é a consultora da área. Assim, foi caracterizado que o grupo estava consistente e apto para a continuação do trabalho.

As principais entidades elicitadas e sua importância no trabalho da DIR, conforme levantamento utilizando-se a técnica de Análise de Stakeholders são:

Na Camada de Contribuição:

- Diretoria/Assistência Técnica: responsáveis pelas tomadas de decisão relativas a todos os trabalhos da DIR. Sua responsabilidade é controlar e responder perante instâncias superiores e à própria população sobre assuntos de interesse da saúde pública. Todas as informações coletadas durante o trabalho da DIR passam por essas pessoas, que necessitam de informações claras para tomadas de decisão. Essas

decisões podem ser de cunho operacional, como por exemplo, quem irá fazer uma vistoria, ou de cunho gerencial, como o aumento da criticidade de um processo por motivos diversos.

- **Protocolo/Expediente:** São os responsáveis pela entrada, encaminhamento e saída das solicitações do público para a DIR. Essa unidade é a interface com o público, devendo ter todas as informações de uma determinada requisição. Também é responsável pelo cadastro das informações nos moldes atuais, ou seja, nos sistemas existentes, principalmente o *SIAP* (Sistema de Informação de Atendimento ao Público).
- **Coordenadorias de Áreas:** São os responsáveis pelo controle direto dos técnicos e por determinado nicho de trabalho da DIR. São divididas em três: coordenadorias de processos, projetos e meio ambiente. Ao lado da Diretoria e Assistência técnica é a área que mais necessita de um bom fluxo e controle de dados, devem fazer controle de seus recursos humanos e de seus recursos materiais, quase sempre escassos, da melhor maneira possível.
- **Técnicos:** são os responsáveis pelas vistorias e pareceres técnicos sobre os processos que dão entrada na DIR. Seu trabalho influencia diretamente todos os resultados. Como são muitos, a sua presença foi substituída pelos coordenadores de áreas, que conhecem bem o trabalho deles e poderiam demonstrar com certeza as necessidades dos mesmos.
- **Suporte à Informática/Sistemas:** representadas pela analista de sistemas do CVS (Centro de Vigilância Sanitária), são responsáveis por toda a implementação e suporte de soluções computacionais dentro da esfera estadual da Vigilância Sanitária. Eles seriam os operadores finais do sistema e as suas necessidades de implementação deveriam ser espelhadas na elicitação de requisitos.

Na camada de Fonte:

- **Anvisa:** A Agência Nacional de Vigilância Sanitária é a responsável por determinar normas, regras e procedimentos para as agências de saúde do país. Assim sendo,

suas resoluções devem ser espelhadas e constar do sistema de Data Warehouse.

- Polícia/Procon/imprensa: Esses órgãos geralmente entram como fornecedores de denúncias e de dados sobre determinados problemas. No caso da polícia, ela serve ainda como apoio operacional em algumas situações de risco.
- Conselhos de Classes: Conselhos de classe, como o CRM (Conselho Regional de Medicina) e o COREN (Conselho Regional de enfermagem), entram com denúncias e ajudam no processo regulatório de algumas entidades como hospitais e clínicas.
- Secretaria da Fazenda: O cadastro das empresas na Secretaria da Fazenda dita o que a empresa pode ou não fazer, sendo assim, essas informações são cruciais para a tomada de decisão a respeito de licenças de funcionamento, por exemplo.
- Ministério Público: O Ministério público geralmente pede apoio em algumas situações, geralmente pedidos de informação sobre casos sobre a saúde pública que tomaram vulto.
- Poder Judiciário: Toma decisão a respeito de processos remetidos para/contra a DIR. Suas decisões devem ser analisadas em situações futuras e as informações de processos da DIR para defesa devem ser bem resguardadas.
- Outras DIR: Como cada DIR é responsável por uma região, informações de empresas, processos e licenças entre eles devem ser compartilhadas.

Na Camada de Mercado

- Prefeitura: A área da saúde, bem como outras áreas de interesse da população, está passando por um processo de municipalização. Com isso, uma parte do trabalho da DIR foi transferida para as prefeituras, que servem então como comparativos diretos dentro de determinados parâmetros.
- Anvisa (Agência Nacional de Vigilância Sanitária)/Visas: A própria Anvisa e os outros escritórios das Visas pelo país servem como fonte de comparação do trabalho da DIR.

Na camada de Comunidade

- Governo/Imprensa/Ministério da Saúde: Dentro de um determinado nível, esses órgãos afetam de forma indireta o trabalho e o fluxo de informações dentro da DIR, servindo em grande parte como sorvedouro de informações. Por isso, os participantes entenderam que eles seriam parte da comunidade.

Analisando esses dados e comparando-os com as anotações feitas durante a reunião, notamos que os usuários comentaram sobre algumas fontes importantes mas não as indicaram. Apesar de não terem colocado explicitamente o principal sistema de acompanhamento de processos da DIR, o SIVISA, como uma fonte, com certeza ele é uma importante fonte para o Data Warehouse, pois ele é o principal sistema legado cujas informações seriam carregadas. É de extrema importância que o analista que esteja utilizando a técnica preste atenção a esses detalhes, pois mesmo nas conversas paralelas entre os usuários informações importantes foram elicitadas.

Essa primeira reunião mostrou algumas características do grupo, durante a execução da técnica. As pessoas se sentiam um pouco nervosas pela presença do gravador. Além disso, a participação das pessoas foi diferenciada. Algumas delas tomaram a frente, participaram e deram opiniões, enquanto outras permaneceram quase todo o tempo caladas ou dando alguma negativa a uma afirmação dos colegas. Essa situação se confirmou nas outras etapas do estudo de caso.

Reunião 2 – Uso do Quadro de Avaliação

A próxima Ferramenta utilizada foi o Quadro de Avaliação. Para isso, como material de apoio, os dados levantados durante a análise de *Stakeholders* foram transferidos para cartazes que continham as linhas com cada entidade declarada, mais colunas com as funções, problemas encontrados e com soluções sugeridas pelos usuários.

Foram utilizados três cartazes, sendo cada um deles continha uma das camadas do diagrama da Análise de Stakeholders, ficando apenas as camadas de Mercado e Comunidade no mesmo cartaz. Em cada cartaz havia quatro colunas, sendo elas:

- Parte interessada: a entidade que é transportada a partir da análise de *stakeholders*. Cada uma dessas entidades deve ter um representante para ajudar na elicitação dos processos e de necessidades de informação;
- Condições/Efeito: Essa coluna representa qual a interação da entidade no fluxo de informações ou de tomada de decisão, bem como as informações necessárias para sustentar essa tomada de decisão;
- Questões/Problemas: Nessa coluna, os participantes descrevem os seus problemas mais comuns relacionados ao suporte de tomada de decisão, quais os conflitos gerados, necessidades não atendidas e problemas no processo;
- Possíveis soluções: Como os representantes das áreas enxergam os seus problemas, e que medidas poderiam tomar para solucioná-los.

Apesar do desconforto notado na reunião anterior, ainda assim tentei utilizar o gravador. Essa reunião foi a mais demorada de todas, começando na parte da manhã e se estendendo pela tarde. Foi também uma das mais proveitosas em relação a problemas de processos e do fluxo de informações da DIR.

Antes da última reunião havia sido tomada a decisão de que eu não iria explicar aos usuários exatamente o que era um Data Warehouse, pois havia a possibilidade de que isso acabasse influenciando o resultado. Ao invés disso, essa falta de conhecimento atrapalhou um pouco o andamento da reunião anterior e, por isso, resolvemos que na primeira parte da reunião seria explicado, de uma maneira acessível a leigos, o que era um Data Warehouse. Apenas para a representante da área e informática do CVS havia sido feita essa explicação previamente.

O número de pessoas participantes foi o mesmo da reunião anterior, mas um dos participantes não pode comparecer e foi substituído por um outro coordenador de área, que

nos acompanhou por todas as reuniões até o final do estudo de caso. Um dos fatos mais interessantes do dia ocorreu por causa desse participante. Ao ser informado das atividades que estávamos fazendo e dos resultados da reunião anterior, essa pessoa começou a fazer uma série de questionamentos a respeito de como o “programa” funcionava, que tipo de Data Warehouse seria necessário e uma série de perguntas que me levaram a pensar que ele já havia passado por algum tipo de atividade de levantamento de requisitos. Nesse momento os outros participantes da atividade tomaram a direção e começaram a explicar não somente a parte do Data Warehouse, mas também da Semiótica Organizacional. Foi uma ótima oportunidade para verificar o que pessoas leigas pensam quando apresentados para essas ferramentas e foi bastante proveitoso para o estudo de caso, pois consegui verificar que os participantes conseguiram entender da proposta e suas opiniões sobre as técnicas até aquele momento.

A frase a seguir faz parte da transcrição da fita gravada, como um exemplo do que foi colocado por um dos participantes:

“Quando a gente desenvolve o sistema o que acontece? Tinha lá o formulário, todo mundo palpitou, aí desenvolvemos o programa e todo mundo usa. Só que foi feito uma coisa entre aspas de cima para baixo. Gostou ou não gostou, vai ter que engolir. Algumas regionais participaram, só que não a grande maioria. Hoje a proposta seria mudar isso. Antes do computador, antes do sistema, antes de qualquer coisa, vamos sentar e ver o que a gente efetivamente precisa. Não se preocupe com a máquina. Se a gente tem equipamento, se não. Agora, a gente vai ver primeiro como é que funciona esse relacionamento das pessoas, como eu me relaciono com os outros públicos... Porque aqui eu tenho esse público, só que ali naquela salinha tem um outro público, tem o ministério...”

O sistema ao qual esse participante se referiu foi o sistema original da DIR, o SIVISA. Podemos notar pela primeira sentença que o método de elicitação de requisitos foi o método tradicional de entrevistas, sem o apoio da maioria dos órgãos que mais tarde utilizaria o sistema. A parte do relacionamento entre as pessoas, a maneira como as pessoas se comunicam e trabalham, como flui a informação e as suas relações, foi bem entendido

pelo grupo que era esse tipo de informação que eu buscava com o uso das ferramentas da Semiótica Organizacional.

Nesse momento também houve, de um outro participante, um *feedback* a respeito da atividade anterior, a Análise de *Stakeholders*:

“O que a gente fez da outra vez, foi levantar os fatores que envolvem nosso trabalho... Por isso que é legal, porque ele conseguiu amarrar o que cerca nosso trabalho, aonde nós estamos, quais os atores estão envolvidos”

A reunião continuou assim por algum tempo, com opiniões e relatos de casos de levantamentos para a construção de sistemas, que quase sempre não atenderam às expectativas dos usuários. A pessoa que originalmente perguntou a respeito de Data Warehouse já havia realmente passado por algumas situações dessas, e tinha um filho estudante da área de informática, vindo daí a “contaminação” pelo assunto.

As tabelas com a transcrição dos cartazes utilizados como material de apoio durante a atividade, com os problemas e soluções indicados pelos usuários para a camada de contribuição, encontram-se no Anexo **B**. Apesar de alguns pontos dos problemas não se relacionarem diretamente com aspectos do uso e do fluxo de informações, que é o mais importante para a elicitação de requisitos para um sistema orientado a dados como as soluções de Data Warehouse, esses problemas são importantes para conseguirmos avaliar aspectos gerais da empresa, quais seus pontos fortes e fracos, quais problemas poderemos enfrentar durante um eventual desenvolvimento da solução e quais poderiam ser as maiores lacunas nos processos. Nessa fase para a qual estamos estudando o emprego da ferramenta, quanto mais conhecermos dos meandros do objeto de estudo, melhor. E são esses detalhes que nos mostram a real saúde dos dados da empresa e sua percepção de qualidade.

Uma verificação simples desse resultado nos evidencia alguns fatores muito importantes para a confecção de uma solução para Data Warehouse:

- As maiores dificuldades em torno de informações estão na Diretoria/Assistência Técnica e nas Coordenadorias de Área. Isso é natural, pois os tomadores de decisão são os que têm maiores dificuldades para conseguir as informações e as inter-relações de dados que precisam.
- As dificuldades apresentadas são de várias fontes diferentes, como volume de dados, falta de recursos e problemas de controle de pessoal;
- A incapacidade atual de se fazer verificações de *status* atual de um projeto ou requisição.

O resultado do Quadro de Avaliação para fontes pode ser visto no Anexo B2.

Como já discutido em sessões anteriores, a camada de fonte é muito importante para uma solução de Data Warehouse, pois nela deverão aparecer os sistemas legados e as fontes de informação que deverão ser tratadas, homogeneizadas, calculadas e carregadas no Data Warehouse. A quantidade de fontes indicadas pelos usuários no Quadro de Avaliação, bem como os problemas relativos às suas informações dão uma idéia das dificuldades e pontos de ajuste que a empresa deverá cuidar antes da implementação final de uma solução de Data Warehouse.

Mais uma vez, seguindo padrão da Reunião 1 , muitos problemas não são relacionados diretamente com fluxo e uso de informações, mas dão pistas sobre a forma de trabalho da entidade estudada e de sua interação com as outras entidades que fazem interface com ela.

Para as camadas de Mercado e Comunidade não houve muita discussão, as pessoas que participaram da atividade, apesar de terem indicado as entidades como parte atuante do processo de dados, não conseguiram indicar como elas se encaixariam no processo como um todo, além do fato de algumas vezes requisitarem informações transacionais, fora do escopo de um Data Warehouse. Assim sendo, a maioria das colunas ficou sem comentários. Esse efeito pode ser isolado nesse caso específico estudado, mas não causa surpresa, pois já

acreditávamos que as camadas mais externas da Análise das Partes Interessadas não poderiam realmente influenciar de forma direta um sistema de apoio à decisão. Isso porque as camadas mais externas não têm contato direto com o sistemas, sendo a sua interação com o mesmo feita indiretamente e através de um agente das camadas mais internas. Caso perceba-se que um agente considerado das camadas mais externas atua de forma direta com o sistemas, seja fornecendo dados como fonte ou até mesmo operando o sistemas, ele deverá ser migrado para camadas mais internas.. Podemos ver nos Anexos B3 e B4 os resumos das atividades para Mercado e Comunidade.

A Reunião 2 foi a que mais durou e que demonstrou como as pessoas interagem durante o uso das ferramentas estudadas. A longa reunião, que começou no meio da manhã e se estendeu até o fim da tarde, foi bastante cansativa e no final dela os participantes já estavam bastante esgotados, com o nível de participação bastante baixo. Apesar disso, fatos importantes foram evidenciados.

Conforme a reunião ia fluindo, em alguns momentos houve algumas divergências de opinião entre os participantes, o que era esperado. Após a pausa para o almoço, algumas pessoas voltaram mais rápido, e começaram a discutir alguns assuntos que haviam sido tratados na etapa da manhã. Nessa hora ficou evidente o desconforto de algumas pessoas criticarem o trabalho e apontar as falhas de outras pessoas. Durante essa conversa alguns pontos relativos a divergências de processos entre os participantes foram tratados, bem como alguns pontos relativos à hierarquia. Notamos que a divisão em grupos menores pode ser necessária durante algumas tarefas, a fim de evitar desconfortos e maior naturalidade e honestidade nas respostas, principalmente quando há diferentes níveis hierárquicos presentes.

A análise dos dados levantados, feita antes da última reunião, mostrou que já havia bons requisitos para a construção de um Data Warehouse, mas, além disso, esses dados já mostravam algumas informações importantes como, por exemplo, os pontos de falha nos processos, as áreas pouco informatizadas e as necessidades gerais da DIR. Essa avaliação dos problemas dos processos e da qualidade e do fluxo de dados poderia evitar que, caso o

sistema de Data Warehouse fosse implementado, pontos de falha importante não fossem percebidos apenas em fases avançadas do projeto, provocando mais custos, demandando maiores prazos e provocando desconfiança prévia dos usuários com o sistema.

Ao final dessa reunião, ainda houve alguns comentários por parte dos participantes. Em especial dois, um que demonstrou o possível perfil de uso de uma ferramenta de Data Warehouse pela DIR e o outro que demonstrou como o envolvimento de um usuário que usa uma das técnicas da Semiótica é grande.

Um dos participantes estava explicando a necessidade de controle de um talonário fornecido a médicos para a prescrição de substâncias controladas. Cada um desses talonários tem um número de série e cada médico que pega um desses talonários pode ser rastreado. Esse participante tinha interesse e achava bom fazer algumas correlações entre o tipo de droga prescrita e os médicos, saber que determinado volume era prescrito, de qual tipo, para quem, ou seja, gostaria de ter informações para levantar se algum médico estaria fazendo prescrições indevidas, em grande quantidade. Outro participante acredita que isso não é necessário, pois a função de fiscalização havia saído das mãos deles e eles não tinham a obrigação de ver isso. Aqui podemos notar que, embora um participante esteja querendo manipular as correlações e tentar extrapolar novos dados, o outro não achava isso necessário para a execução dos trabalhos, preferindo um método de trabalho mais formalizado e rígido. Assim, o perfil de uso da ferramenta por essas duas pessoas seria bastante diferente. Não estamos querendo dizer com isso que haja uma postura errada por parte de algum dos usuários, mas apenas que pessoas encaram as informações disponíveis de maneiras diferentes. Algumas pessoas, de posse de uma ferramenta de apoio a decisão, irão apenas verificar alguns dados em relatórios padrão, enquanto outras irão realmente navegar nos dados e tentar achar novas correlações de dados. As necessidades de todos os usuários, inclusive como eles fazem a navegação dos dados, é um requisito bastante importante para a confecção de uma boa solução de Data Warehouse.

O outro fato foi um comentário de um participante ao final da reunião. Ele ficou parado, olhando para os cartazes e depois de algum tempo se declarou surpreso, pois não esperava

encontrar tantos pontos de problema. Esse tipo de conclusão é muito bom, pois dá idéia do grau de auto crítica que se consegue com esse tipo de trabalho. Somente esse tipo de choque de realidade pode fazer com que as pessoas reajam e tenham desejo de perceber os seus erros e trabalhar para corrigí-los. Numa outra técnica de elicitação de requisitos usualmente empregada, esses problemas passariam a largo e afetariam com certeza a percepção de qualidade final do sistema.

Normalmente o emprego das técnicas do PAM é rápido, mas o próprio perfil do grupo estudado dificultava uma maior velocidade na elicitação dos requisitos. Por isso, resolvemos tentar manter um horário e tempo fixo para as reuniões, para evitar que o trabalho da DIR fosse prejudicado.

Reunião 3 – Uso do Framework Semiótico

Nessa reunião estavam presentes as pessoas consideradas mais importantes, a Diretora, as Assistentes técnicas e a analista de sistemas. Havia ainda mais uma participante, uma coordenadora de área. Como havia sido imposto que a reunião deveria acabar em aproximadamente duas horas, tentei ao máximo que não houvesse desvios como os das outras reuniões, para manter o foco do assunto. Como material de apoio eu utilizei a escada Semiótica, e nos focamos nas três camadas superiores, a semântica, a pragmática e a social.

Há alguns cuidados necessários para o uso do Framework Semiótico. O analista que está participando dos workshops com os usuários das técnicas anteriores, ao chegar nesse passo deve ter uma boa idéia de como o trabalho da organização é executado e quais os seus maiores problemas. A Escada Semiótica é responsável pela unificação das informações abrange todos os aspectos informais ainda não descritos, possibilitando assim a confecção das *tabelas-fato* pelas separações dos processos descritos na mesma.

Tendo em vista o universo de uma implementação de sistemas de apoio à decisão (DSS, uma outra terminologia para a classe de sistemas entre as quais se enquadram as soluções de Data Warehouse), podemos considerar cada um dos degraus da Escada Semiótica da seguinte forma:

Os três degraus inferiores, o físico, o empírico e o sintático são de suma importância para a análise de requisitos de software existentes na empresa, para a implementação final da solução. Com base nessas camadas, podemos definir quais as tecnologias preferidas da empresa, quais os tipos de recursos computacionais que mais os agradam e quais sistemas existentes na empresa podem facilitar a implementação. Essas informações podem ser de grande valia no momento de se definir qual tecnologia de implementação de Data Warehouse deve ser utilizada, se ROLAP, MOLAP, ou qualquer outra.

Por outro lado, durante o processo de análise inicial do sistema, quando se estuda a viabilidade do sistema e suas características de negócio, o uso dessas três camadas não é necessário, podendo confundir o usuário leigo da área de informática. Assim sendo, nessa fase inicial deve-se focar nos três degraus superiores da escada; são eles:

Semântico: Essa camada explicita os significados, valores, cultura das entidades estudadas. Consideramos essa camada muito importante para entendermos o âmbito de linguagem e de significados, tentando evitar, através de uma compreensão mais afinada das particularidades de linguagem e de simbologia do grupo estudado, que o escopo de entendimento da realidade estudada esteja errada. Em qualquer sistema computacional, mas principalmente numa solução como as soluções de DSS, qualquer anomalia introduzida por uma falta de entendimento pode causar danos irreparáveis ao projeto, sistema e solução, como um todo. Após as camadas de ETL e Dados estarem parametrizadas, qualquer modificação será extremamente custosa, incluindo custos operacionais, temporais, transacionais e, talvez o mais importante, o custo introduzido pela perda de confiança dos usuários na ferramenta e no processo e construção da solução.

Pragmático: Essa camada é a mais importante de toda a análise, pois trata das intenções dos usuários com o sistema. Quando perguntados, os usuários podem dar uma gama enorme de motivos para utilizar o sistema, diferentes formas de utilização, com que tipos de dados alimentar o sistema, etc. Mas nem sempre as reais intenções são mostradas. Muitas vezes há interesses políticos, hierárquicos e operacionais que não constam dos *modus operandi* oficial da empresa, mas fazem parte do dia a dia da operação. Podemos exemplificar com a necessidade de um determinado setor se proteger de outro, através de

logs de operação e curvas de tendência de resultados. Imaginemos que numa determinada empresa haja um setor de logística que falha em cumprir os prazos de entrega. De quem é realmente a culpa? Desse setor ou do responsável por embalar o produto e disponibilizá-lo para a entrega? Qual a real parcela de responsabilidade de cada setor? Qual a intenção de um diretor, o que ele realmente quer ver na sua empresa? Se a empresa visa lucro, de que maneira os responsáveis pretendem controlar isso? Como podemos imaginar, há muitas nuances de interpretação na maneira de trabalho das entidades e descobrindo as reais intenções podemos dispor das informações necessárias para cobrir essas necessidades.

Social: Essa camada representa a cultura da entidade, como ela se expressa e se comporta dentro do mundo social. Apesar de Liu (2000) considerar essa camada independente, Shanks e Corbitt (1999) consideram que ela é extremamente relacionada com a camada pragmática, tanto que eles defendem que a camada social deveria ser englobada pela pragmática. Isso porque as suas intenções definem seu comportamento para o mundo, seja esse comportamento real ou apenas interpretado como tal.

Cada uma dessas camadas foi dividida em duas colunas, representando a descrição de assuntos tratados na camada e possíveis soluções para problemas apresentados, indicadas pelos próprios *stakeholders*. Isso provoca uma análise profunda por parte das entidades, sempre lembrando que estamos estudando o âmbito de uma ferramenta de DSS, os principais problemas de interface de dados, as dificuldades mais comuns quando se necessita de uma informação ou os problemas com a qualidade das informações apresentadas. As soluções apresentadas pelos participantes do estudo representam a maneira como a organização pensa e trabalha. Os problemas são resolvidos com base na maneira de trabalho da entidade e não há determinação externa para adequação de forma de trabalho. Os dados e informações dessa etapa servem como base para a construção da solução de Data Warehouse.

As camadas inferiores, segundo minha previsão, foi apresentada e acabou não sendo discutida, pois o público presente não era parte de uma equipe de projetos de *software*, não tinham conhecimento técnico para tanto. Além disso, como a idéia do uso do PAM nesse

momento era ter independência de tecnologia, visando um mapa macro da solução e o estudo de possibilidade de implementação de uma solução de Data Warehouse na DIR, as camadas inferiores da escada eram de pouco relevância naquele momento. Isso não quer dizer que essas camadas não sejam importantes. Numa segunda etapa do projeto, após a decisão de implementação tomada, essas três camadas serão de suma importância para a decisão da tecnologia de apoio para Data Warehouse a ser utilizada.

Como já foi dito, o uso do gravador parecia melindrar os participantes. Nas duas etapas finais o gravador não foi utilizado, e foi feito um maior aproveitamento do material de apoio. Os dados levantados nessa etapa constam no Anexo C.

Como era de se esperar, a parte que mais demandou debate foi a pragmática. O degrau pragmático da escada gerou algumas dúvidas, porque as intenções sempre são as mais difíceis. Os problemas que as coordenações têm com os técnicos foram bem retratados, como podemos ver no Anexo C , mas houve muita dificuldade nisso.

Reunião 4 – Uso da Análise Colateral

Essa reunião foi utilizada para o uso da técnica da Análise Colateral. Mais uma vez foi utilizada uma versão simplificada daquela proposta por Liu (2000), pois não haveria como elicitar pontos operacionais como necessidades de *backup* e recuperação, nesse momento. Apesar de o modelo completo ser extremamente importante para o desenvolvimento final do sistema, durante a fase de elicitação inicial de requisitos e de estudo de viabilidade podemos nos ater aos ciclos superiores, compreendidos por:

- **Ciclo de vida:** Esse ciclo forma uma linha temporal com o sistema que está sendo proposto como ponto central. Ela indica os sistemas predecessores, contextualiza o sistema focal e indica um possível sistema sucessor. Esse ciclo vai nos indicar o que é utilizado atualmente, qual a percepção dos usuários sobre o sistema proposto e qual a intenção na implementação, onde eles gostariam de chegar no futuro.
- **Funcionamento:** forma um ciclo de funcionamento do sistema focal, indicando as

suas entradas, suas saídas e a maneira como ele interage e transforma o ambiente, recebendo como *feedback* as possíveis modificações exigidas por ele, na forma de novas necessidades ou novas exigências. Além disso, fornece uma descrição do sistema focal, fornecida pelos usuários, expondo a percepção deles sobre o sistema.

Com base em estudos anteriores, esperava mais uma visão geral de como os usuários enxergavam o que estava sendo proposto. O resumo dos resultados pode ser encontrado no Anexo D.

Apesar de não ser uma surpresa, a análise colateral acrescentou alguns fatos e confirmou outros. Um dos pontos que os usuários não colocaram no Quadro de Avaliação e que eu considerei importante: em momento algum eles citaram oficialmente como fonte de dados os sistemas existentes, o SIVISA e o SIAP. Apesar de falar de suas deficiências e problemas, o grupo não pensou neles como uma fonte de dados. Essa situação foi corrigida na Análise Colateral, pois eles incluíram como predecessor e como fonte de dados os dois sistemas.

Também somando ao resultado, o grupo pela primeira vez durante a atividade colocou dados mais aprofundados que desejariam ver, como os dados que gostariam de conhecer a respeito dos técnicos. Apesar de não ser extremamente relevante nesse momento do estudo de viabilidade, esse fato dá indícios de que o conjunto das técnicas pode ser bem utilizado em todos os momentos do ciclo de vida de um projeto de Data Warehouse, mesmo nas fases de verificação e validação mais avançadas.

Considero que além da facilidade de uso das ferramentas e do poder que elas têm para conseguir eliciar os desejos e intenções dos usuários de forma clara, as reuniões serviram como sessões de auto crítica para o grupo estudado, forçando um aumento da maturidade e do entendimento de seus problemas. Um dos participantes, num determinado momento, disse que as reuniões pareciam muito com sessões de planejamento, o que não deixam de ser. Mas para planejar é necessário reconhecer os seus objetivos, reconhecer seus defeitos e saber aonde se quer chegar. Por isso, ao fazer uma avaliação real da sua condição atual, e

muito bem apoiado pelas técnicas da Semiótica, o grupo pôde realmente entender as ações que eles deveriam tomar e quais seriam seus reais objetivos.

5.4 Análise dos resultados e desenho da solução

De posse de todos os dados, o analista pode começar a desenhar os planos para o seu Data Warehouse. A análise de tudo o que foi levantado até o momento visou um levantamento inicial da possibilidade de se fazer um Data Warehouse para uma empresa, tendo como principal produto as suas *tabelas-fato* e *dimensões*, possíveis entradas na camada de ETL e avaliação da qualidade de dados na empresa, perfil dos usuários e ações necessárias anteriores ao projeto. Com os dados levantados na DIR, foi possível elicitare seus requisitos para algumas dimensões e para as camadas de ETL, utilizando-se o modelo proposto. Para um maior entendimento da forma de uso das ferramentas, serão descritas passo a passo as avaliações feitas para o resultado final.

A primeira etapa foi a de Análise de *Stakeholders*. Essa etapa do processo visava dois objetivos principais:

1. Garantir que todas as áreas interessadas no processo participassem;
2. Conhecer as principais fontes de dados necessárias para as tomadas de decisão.

Analisando o resultado verificamos que, para elicitação inicial, era imprescindível ao menos um representante das seguintes áreas:

1. Diretoria
2. Assistência Técnica
3. Cada uma das três coordenadorias de área
4. Técnicos de cada área
5. Protocolo
6. Sistemas de informação

Durante a análise, todos concordaram que a presença dos técnicos poderia ser substituída pelos coordenadores de área, pois os técnicos não utilizariam diretamente o sistema, mas sim alimentariam os dados dos sistemas fonte do Data Warehouse.

As fontes de dados levantadas foram:

- SIVISA : Relatórios técnicos e dados sobre os processos
- SIAP: Dados sobre os processos e protocolo
- Procon: Informações sobre estabelecimentos e processos
- Polícia: Informações sobre estabelecimentos e denúncias
- Conselhos de classe: Informações sobre estabelecimentos e denúncias
- Ministério Público: Denúncias, resoluções e processos
- Secretaria da Fazenda: Informações sobre estabelecimentos
- Imprensa: Denúncias e Informações gerais
- Papéis Gerais de processos da DIR

Apesar de também não terem sido colocados como fonte na análise das partes interessadas, durante esse processo pude elicitare várias fontes de informação e dados que não estavam automatizadas em sistemas computacionais, como por exemplo as fichas de tempo dos funcionários e planilhas de disponibilidade de viaturas e outros insumos. Os próprios usuários recolocaram esses problemas durante o Quadro de Avaliação, mas uma análise profunda nesse momento ainda indica que os dados de entrada e saída de processos, no momento de entrada das requisições, é apoiado pelo SIAP, mas ainda assim há fatos que não são computados. Por exemplo, em momento algum há um registro de pessoas encaminhadas de forma errônea por outras entidades, como a Prefeitura. Desse modo, ainda como preocupação para a camada de ETL, teríamos essas novas entradas não formatadas. Essas informações faltantes impactariam a implementação do sistema final e devem ser trabalhadas em uma fase anterior ao Data Warehouse, para a adequação de processos e disponibilização desses dados. No caso específico da DIR, há uma necessidade de reengenharia de processos e um aumento do escopo de informações dos sistemas existentes.

No caso das planilhas de ponto e de insumos, é necessário que sejam criadas rotinas de carga, mesmo que essas cargas sejam apenas arquivos texto formatados. O importante no caso é criar um processo para a disponibilização dessas informações faltantes.

A Figura 20 representa as entradas dos dados na camada de ETL. Em fase posterior, caso o sistema de Data Warehouse seja recomendado e aprovado, a primeira coisa que o analista responsável pelas interfaces de carga de dados do sistema deveria fazer é um estudo aprofundado da sintática utilizada para cada um dos códigos desses sistemas, e a necessidade de transformações para as cargas dos mesmos. Durante as reuniões, quando falei a respeito da sintática, foi-me passada a regulamentação da portaria de lei que criou o SIVISA, e nela estavam todos os códigos e tabelas utilizados pelo sistema. De forma bastante precisa os usuários entenderam a necessidade sintática de um Data Warehouse e puderam me prover de informação bastante precisa nesse aspecto.

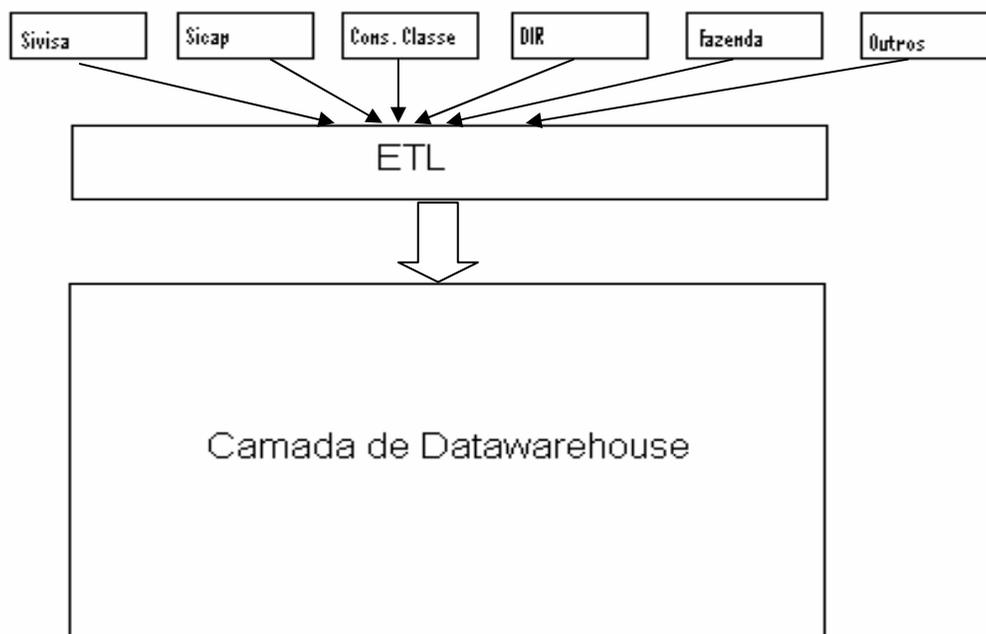


Figura 20 : Sistemas fonte e Camada de ETL

Esse primeiro mapa formam o direcionamento inicial ao analista que fará as interfaces para quais as fontes de dados e quais problemas com os quais ele deve lidar.

A seguir, o analista deve analisar os resultados do *Framework Semiótico*. Apesar de não ter sido a técnica imediatamente utilizada após a Análise de *Stakeholders*, essa ferramenta é a principal na análise de requisitos. Isso porque dela pode-se inferir as reais intenções dos usuários, aonde eles querem chegar com a análise dos dados, quais os problemas que mais dificultam seus resultados.

Analisando as intenções nos degraus pragmáticos e sociais, percebemos que a grande preocupação da DIR é a qualidade e informações de seus processos, que incluem pedidos de licenças, liberações de projetos, problemas investigativos de saúde da população, entre outros. A camada semântica nos dá várias pistas a respeito de vários problemas, principalmente do cunho de formatação e homogeneização de procedimentos e relatórios. Desse modo podemos dizer que a principal *tabela fato*, o processo que melhor representa as necessidades da DIR seria uma *tabela fato* onde ela pudesse acompanhar o ciclo de vida e angariar informações sobre os seus diversos processos. Assim sendo, temos a primeira figura no nosso mapa do *Star Schema*. Como a DIR abre um processo administrativo para todas as suas tarefas e pedidos de usuários, chamaremos a aplicação de “Aplicação Processos”.



Figura 21 : A Aplicação elicitada

A análise do *Framework Semiótico* não necessariamente deve sugerir apenas uma aplicação. Caso houvesse surgido mais algum processo crítico, mais processos seriam incluídos, até atingir a totalidade das necessidades da DIR.

Agora iremos analisar todas as informações das outras etapas do processo para descobrirmos que *dimensões* são requeridas para o Data Warehouse. Uma análise dos resultados deve ser feita principalmente no Quadro de Avaliação, que irá nos indicar os principais problemas enfrentados pelos usuários analisados.

Podemos elencar dos Quadros de Avaliação, problemas colocados pelos usuários que façam sentido no escopo de um Data Warehouse. O ideal é que esses problemas sejam categorizados a fim de evitar redundância de dados.

Com base no Quadro de Avaliação da Vigilância, podemos fazer a seguinte divisão de informações necessárias, sempre visando a coluna das soluções propostas pelo grupo de trabalho:

1. Com quem está o documento;
2. Quantos documentos estão com qual técnico/área;
3. A quanto tempo está sem solução/encaminhamento;
4. Onde ficou parado (gargalo);
5. Porque não teve encaminhamento (ex: por falta de viatura);
6. Melhor controle das empresas por projeto;
7. Melhor visibilidade dos casos por técnico e controle de retorno dos das demandas encaminhadas para os técnicos (frequência);
8. Disponibilização de agendas de insumos (recursos, como carros e equipamentos especializados);
9. Acesso aos registros de produtos, pendências e do histórico do produto;

Com essa lista de soluções propostas pelos usuários, podemos verificar quais *dimensões* serão necessárias para cobrir essas necessidades. É fácil perceber que existem pontos de mais atenção para a DIR, e podemos cobri-los com base nessas informações. Fazendo a análise ponto a ponto:

1. O primeiro ponto indica a necessidade de um cruzamento entre o que vamos chamar de **PROCESSOS** e os responsáveis pelo ciclo de vida desse processo, que iremos chamar de **RESPONSÁVEL**.
2. O segundo ponto indica a necessidade de um cruzamento entre **PROCESSOS**, **ÁREAS** e **TÉCNICOS**.
3. O ponto 3 indica a necessidade de histórico, precisando verificar o **TEMPO x PROCESSOS**.
4. A necessidade de “onde ficou parado”, pode ser inferida com base no resultado da query entre **PROCESSOS** e **RESPONSÁVEL**, não necessitando de uma dimensão nova para isso.
5. O porquê não teve encaminhamento depende de como vai ser a análise, se por um relatório ou pelo estudo de uma *query* entre **TEMPO**, **PROCESSO**, **RESPONSÁVEL** e **INSUMO**. Como a última opção possibilita maior flexibilidade em caso de necessidade de mudança, vamos escolher esse caminho.
6. O ponto número 6 indica a necessidade do cadastro dos **ESTABELECIMENTOS**, e de sua correlação com os **PROCESSOS** existentes.
7. O ponto de número 7 mostra a necessidade de uma correlação entre **TEMPO**, **PROCESSOS** e **TÉCNICOS**.
8. A agenda de **INSUMOS** deve ser correlacionada, pelo menos, com os **TÉCNICOS** e com **TEMPO**.
9. O ponto 9 especifica a necessidade de uma dimensão **PRODUTO**, que deve ser pelo menos o histórico desses produtos.

Após o levantamento desses itens, fomos capazes de inferir as dimensões necessárias para a aplicação de processos da DIR, que resumindo são:

- **PROCESSOS**
- **RESPONSÁVEL**
- **ÁREAS**
- **TÉCNICOS**
- **TEMPO**
- **ESTABELECIMENTOS**
- **INSUMO**
- **PRODUTO**

A representação dessas oito dimensões no *Star Schema* fornece o mapa de implementação inicial do Data Warehouse que queremos. A sua visualização gráfica, usualmente utilizada nos projetos de Data Warehouse ficaria da maneira representada pela Figura 22.

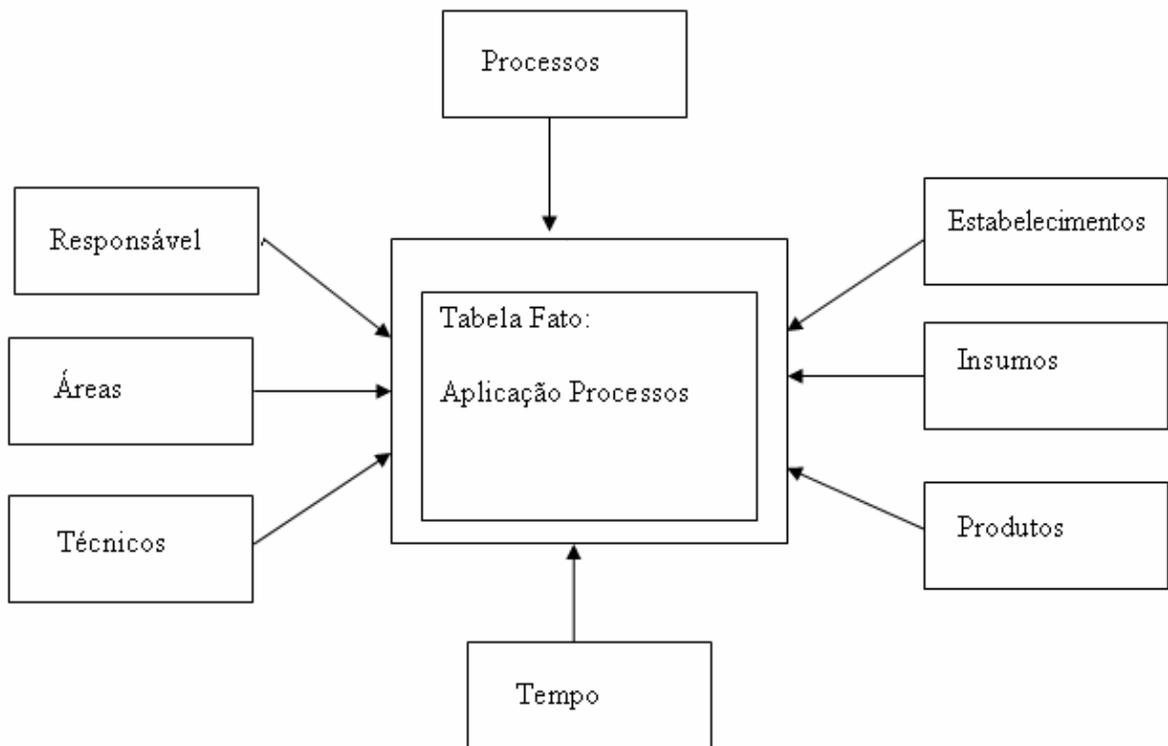


Figura 22 : A aplicação e suas dimensões

Assim, a representação básica do *Star Schema* está feita, com todas as dimensões escolhidas para a aplicação “Processos”.

Além da base do *Star Schema*, podemos também inferir algumas outras informações. A chave de cada tabela sempre precisa estar presente, para permitir a correlação entre as dimensões. O estudo do Quadro de Avaliação e da Análise Colateral também nos dá algumas pistas do caminho que o usuário deseja seguir. Verificando a parte de funcionamento da Análise Colateral, pode-se perceber quais dados realmente os usuários querem presentes e assim temos um indício para o projeto das tabelas do banco de dados ou da representação multi-dimensional. No caso da DIR, podemos associar várias informações desejadas no ciclo de entrada da camada de funcionamento diretamente à *tabela-fato* e às *dimensões*.

Exemplos desses dados são encontrados para estabelecimentos onde são pedidos Razão Social, Nome Fantasia, CNPJ, Endereço, Tipo de Estabelecimento/Atividade, CNAE fiscal, Responsável Legal, Responsável Técnico.

Todos esses campos de dados podem ser levados diretamente para a tabela fato, e depois distribuídos entre as dimensões conforme as necessidades das dimensões. A figura 23 mostra o *Star Schema* com algumas dessas informações complementares.

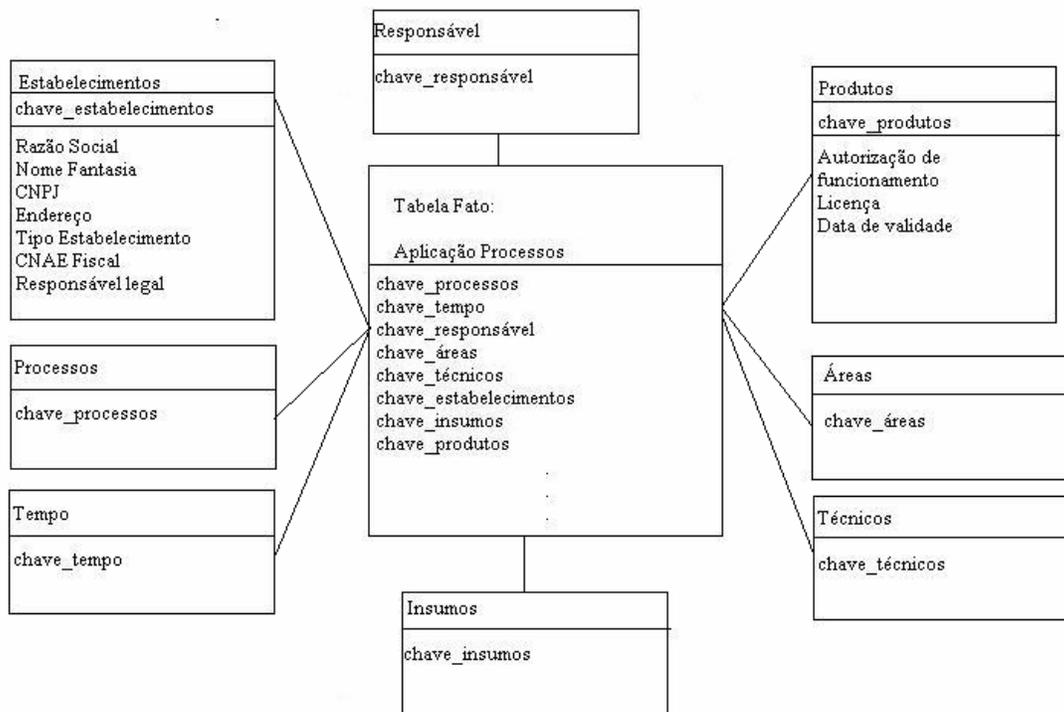


Figura 23 : As dimensões com chaves primárias e informações necessárias

O resultado demonstrado pelo mapa do *Star Schema* é bem acima do esperado, pois além das dimensões algumas necessidades de dados já foram preenchidas. As lacunas ainda existentes seriam sanadas numa etapa de projeto pela análise dos dados existentes nos sistemas fonte e por outras etapas de elicitação de requisitos mais aprofundadas, quando seriam definidos todos os dados necessários e seus metadados, formando assim a estrutura completa do *Star Schema*.

Baseando-se nesse resultado, também podemos inferir alguns potenciais problemas de falta de informação e de dificuldade de carga. Segundo o que foi elicitado no Quadro de Avaliação, as seguintes informações estão dispersas, não existem ou não estão imediatamente disponíveis:

- Relatórios de outras DIR: Os sistemas das DIR não são centralizados, todos eles baseados em sistemas locais. Quando há a necessidade de troca de informações, há

a centralização dos dados no CVS, após o envio dos novos arquivos do mês. Como um Data Warehouse não é um sistema transacional, ou seja, não há maiores problemas com um certo lapso temporal nas informações, isso não seria impeditivo ao sistema, mas um acordo prévio deve ser feito para a distribuição desses dados após a sua consolidação central. Essa necessidade fica clara no que foi explicado numa reunião, exemplificando que muitas vezes uma determinada indústria tem a central dentro da jurisdição de uma DIR, mas uma filial na jurisdição de outra. Desse modo, as informações relativas aos processos entre as empresas das DIR devem ser compartilhadas para possibilitar o estudo das situações das empresas.

- As informações dos dados da Fazenda, como o cadastro fiscal da empresa;
- Sem a implementação de um sistema de *workflow*, não é possível saber exatamente quais as áreas de gargalos dos processos. É possível com base em datas fazer um acompanhamento, mas ele não será totalmente confiável;
- Algumas das informações ainda não estão em sistemas computadorizados e ainda necessitariam de algum processo extra para a carga dos dados, como: cartão de ponto dos técnicos, agenda das viaturas e outros equipamentos, informações provenientes do expediente que ainda não são computadas, como a quantidade de pessoas com encaminhamento errado da prefeitura.

Além do desenho do *Star Schema*, a análise dos dados gerados pelo emprego das ferramentas nos indicou uma série de subprodutos que não eram esperados a princípio, mas extremamente desejáveis. Os resultados indicam uma forte tendência à auto-crítica, e esse exercício faz com que os usuários tomem consciência de problemas a serem resolvidos. Um fato muito interessante observado aconteceu durante a terceira reunião, e o resultado pode ser visto na parte de sucessores do ciclo de vida no efeito colateral, no Anexo D.

Durante a reunião 3, ao relatarem as dificuldades que eles têm com os técnicos, vários casos foram relatados, a maioria onde os técnicos tiveram dificuldades em termos de prazo e qualidade de relatórios. O grupo pensou em possíveis soluções e sem conhecimento prévio, deram a descrição de um sistema que desse uma tarefa a uma pessoa e avisasse quando essa tarefa não estava feita no prazo e que quando esse técnico mandasse o relatório

houvesse a possibilidade de uma etapa de revisão desse relatório. Ou seja, eles simplesmente descreveram um processo de workflow com edição de tarefas e com ciclo de vida de aprovação de documento. Esse resultado é surpreendente, principalmente porque eles nunca haviam ouvido falar de soluções de *workflow*. Expliquei a eles do que se tratava no final da sessão, tanto que nos resultados do Quadro de Avaliação aparece um ítem como solução “Organizar e inferir um sistema de “cobrança de resultados”. Eles notaram que não havia um padrão de avaliação para os relatórios, sendo as “notas” dos mesmos bastante subjetivas. Com isso, somente nessa reunião foi elicitada a necessidade de uma ferramenta de *workflow*, uma base de seu funcionamento e como seria o processo de recusa ou aprovação de um documento.

Após essa análise, verificamos que com a ajuda do ferramental Semiótico fomos capazes de:

- Definir quais os principais atores para a elicitação de requisitos inicial;
- Definir as principais fontes de dados para as camadas de ETL;
- Definir os processos necessários para os usuários;
- Definir quais as dimensões seriam necessárias para esses processos;
- Desenho do esquema macro do *Star Schema*;
- Definir alguns dados de entradas importantes nas dimensões do *Star Schema*;
- Definir quais dados necessários para a aplicação elicitada não são gerados por sistema algum.

5.5 Considerações sobre o resultado

Antes do início do trabalho com os usuários, a intenção principal era conseguir verificar se seria possível, utilizando-se as ferramentas propostas na Semiótica Organizacional, conseguir uma boa definição inicial para o desenvolvimento de uma solução de Data Warehouse para a entidade escolhida para o estudo de caso. Os resultados esperados, em um primeiro momento, eram:

- Definição do grupo de pessoas interessadas cujos interesses seriam elicitados;
- Definição das principais fontes de dados;
- Definição das aplicações e de suas dimensões.

Todos os objetivos acima foram alcançados. A *Análise de Stakeholders* mostrou-se uma ferramenta eficaz para essa tarefa, sendo capaz de evitar que grande parte de uma solução fosse desenvolvida sem que alguma entidade participante do processo fosse contactada. Do mesmo modo, essa mesma atividade dá uma boa idéia de quem realmente necessita das informações e deve participar, inclusive indicando substituições de pessoas e grupos, como aconteceu no caso dos técnicos sendo representados pelas coordenadorias de grupo. Esse resultado é muito importante para o ciclo de desenvolvimento de um Data Warehouse, pois como já comentamos, o público alvo dessas soluções geralmente é de difícil acesso, dificultando o processo de elicitação como um todo. Com essa possibilidade de substituição, pode-se garantir uma maior flexibilidade das reuniões e garantir o mínimo de andamento no projeto, sem ficar dependendo muito de um grupo específico de pessoas.

Da mesma maneira, a *Análise de Stakeholders* também influenciou muito num dos maiores problemas referentes a levantamento de dados de um Data Warehouse. Um Data Warehouse é um sistema para informações de apoio a decisão, ou seja, não necessita ter todas as informações de todos os sistemas, e sim um nível de informações consolidadas. Há um limite prático para a quantidade de informações que se deve carregar em um Data Warehouse, pois uma grande quantidade de informações pode provocar uma grande queda de performance e um tempo de carga muito grande. Qualquer dado não gerencial ou de baixa granularidade deve ser procurado nas fontes transacionais. Geralmente, quando se pergunta a um usuário que dados ele gostaria de ter no Data Warehouse, a resposta é algo do tipo “tudo que está no sistema X”. Com o uso da *Análise de Stakeholders* e do Quadro de Avaliação, essa pergunta não é colocada diretamente para o usuário, e o exercício de *brainstorming* fez com que uma crítica saudável ao sistema fosse feita e que se separasse as reais necessidades de dados. Com isso, uma solução mais enxuta do que se geralmente se consegue é possível.

Como colocado no capítulo 2, são consideradas duas as maiores dificuldades para a elicitação de requisitos para soluções de Data Warehouse: A dificuldade dos usuários em entender e articular as suas necessidades e a dificuldade dos técnicos e usuários se entenderem e fazerem um trabalho conjunto. Essas duas dificuldades foram claramente endereçadas pela Semiótica Organizacional nesse estudo de caso, pois os usuários ao mesmo tempo em que articulavam suas reais necessidades auxiliados pelas técnicas da Semiótica, também permitiam ao analista entender a sua real necessidade e compreender seus jargões e métodos de trabalho. O resultado final, apresentado como o modelo em *Star Schema*, é o que um analista necessita para dar andamento ao projeto de implementação, pois apresenta ali toda a base do desenvolvimento, ou seja, as aplicações e dimensões.

Todas as principais fontes de dados foram elicitadas com o auxílio das técnicas e também com o uso integrado delas foi possível a definição da aplicação necessária e de suas dimensões.

Além desses resultados, alguns outros também importantes foram conseguidos. Como colocado na seção de objetivos, o principal resultado desse estudo seria a verificação do uso das ferramentas da Semiótica para avaliar os requisitos iniciais de uma empresa para uma solução de Data Warehouse. Seria como uma verificação de pré-projeto, qualificando inicialmente a empresa para o uso da solução. Para isso não seria necessário realmente chegarmos a um desenho em profundidade do projeto, mas sim uma especificação inicial que servisse como guia para as etapas subsequentes. Esse objetivo foi plenamente alcançado e ainda obtivemos algumas informações não esperadas.

As principais foram sem dúvida o levantamento dos dados necessários mas não disponíveis e o nível de maturidade processual e computacional da empresa. A indicação de quais dados não estão disponíveis é um resultado colateral muito bom, pois eles indicam de forma bastante precoce quais serão os problemas de coleta de dados que serão encontrados durante fases mais avançadas do projeto. Desse modo pode-se planejar quais ações serão necessárias para que essas falhas sejam sanadas. Lembrando que, como um Data Warehouse é uma solução orientada a dados, esse tipo de problema, caso não tenha uma

solução simples ou em tempo hábil, pode inviabilizar um projeto. Dependendo da fase em que se encontre, isso pode significar o desperdício de anos de desenvolvimento e muitos recursos financeiros e de pessoal. Assim, podemos dizer que esse resultado, apesar de não ter sido um objetivo principal do estudo, é extremamente bem vindo e necessário. Também pôde ser facilmente percebido o contexto geral de uso da aplicação pelos usuários e, se fôssemos classificá-los segundo os tipo de usuários indicados por Inmon(1996), eles seriam basicamente usuários do tipo Fazendeiro, ou seja, que se mantém verificando os processos normais e não procuram correlações mais amplas. Essa informação seria bastante útil em etapas posteriores ao trabalho, para a confecção da interface para usuários.

Outro ponto bastante importante foi o aumento da crítica dos participantes pelo seu processo, exemplificado pela percepção deles da necessidade de um sistema de *workflow*. Apesar de os usuários terem colocado o *workflow* como um sistema sucessor ao Data Warehouse, na verdade ele deveria ser um antecessor. Os dados desse sistema deveriam ser carregados no Data Warehouse e seriam uma fonte de dados muito importante para o acompanhamento gerencial da DIR.

Com isso, a possibilidade de se fazer uma avaliação correta da possibilidade de sucesso da implementação é muito grande. No caso específico estudado, devido a problemas de dados e de processos, seria necessário ainda um ajuste para que a entidade tivesse maturidade o suficiente para a implemetação de um Data Warehouse. Os resultados do levantamento serviriam de guia para essa adequação e se evitaria assim o fracasso do projeto, o desperdício de recursos e o aumento da insatisfação dos usuários com as implementações dos sistemas de informática.

Verificando o resultado final, podemos ver que o objetivo inicial foi plenamente alcançado, inclusive sendo possível uma maior definição dos dados nas dimensões e na tabela fato do que era esperado. O uso das técnicas da Semiótica apenas não deram pistas em relação ao nível de detalhamento e granularidade dos dados requeridos pelos usuários, mas esse tipo de informação é necessária durante o modelamento final do banco de dados e das camadas de carga.

A maior dificuldade que foi sentida na implementação da técnica foi o foco do usuário nas questões de fluxo e qualidade de dados. Era comum, durante as reuniões, que o assunto fosse desviado para outros problemas ou facetas dos processos normais da entidade estudada, como por exemplo necessidades de novas contratações, treinamentos, posturas de profissionais e outros assuntos que apesar de influenciarem no processo de atendimento da DIR não seriam tratados por uma solução de Data Warehouse. Apesar desses fatos também trazerem importantes informações, isso pode causar um aumento excessivo do tempo de reunião e pode dificultar o alcance das metas principais. No caso estudado, isso pode ter sido ocasionado pelo perfil das pessoas que participaram do grupo, ficando esse ponto para um estudo futuro.

As quatro técnicas utilizadas foram de grande valia para o levantamento dos requisitos iniciais, mas um fator importante não foi detectado durante o estudo de caso. A camada de ETL, como foi explicado anteriormente, é responsável pela agregação dos dados e pela granularidade das informações que serão carregadas no Data Warehouse, e esse deve ser preparado para os níveis de *Drill-Down* necessários. Muito pouco se falou a respeito da necessidade dos históricos e dados e a única ferramenta que fez com que os usuários falassem dos níveis de granularidade foi a análise colateral, ainda que rapidamente. Uma maior verificação nesse aspecto torna-se importante para o uso do ferramental em fases mais avançadas do processo.

Apesar desses problemas apontados, as ferramentas são uma opção bastante válida para o levantamento de requisitos iniciais de um Data Warehouse. Outras ferramentas da Semiótica Organizacional, como o Framework Antropológico, Diagrama de Ontologia e Análise de Normas, que não foram utilizadas poderão úteis durante uma fase mais avançada da análise, quando será possível discutir aspectos técnicos e de projeto mais avançados.

Capítulo 6

Considerações Finais e Trabalhos Futuros

Após a análise dos resultados evidenciados no Capítulo 4, podemos dizer que os objetivos iniciais dessa monografia foram plenamente alcançados. O modelo de uso das técnicas da Semiótica Organizacional proposto pode facilmente ser reproduzido em qualquer projeto de soluções de Data Warehouse para a etapa inicial de elicitação de requisitos. Dentro dos resultados conseguidos estão:

- A escolha dos usuários necessários para uma boa elicitação dos processos a serem incluídos no Data Warehouse;
- Definição das fontes de dados necessárias para o apoio aos processos desejados pelos usuários;
- Definição dos diagramas do *Star Schema*, contendo as *tabelas-fato* e *dimensões* necessárias.

No decorrer do estudo de caso ainda fomos capazes de observar a maneira como os usuários reagiam em relação às técnicas da Semiótica Organizacional em diversos cenários, como, por exemplo, a maneira pela qual diferenças hierárquicas dentro dos grupos de trabalho poderiam influenciar no resultado. Outro importante ponto evidenciado durante o estudo de caso foi o crescente entendimento dos usuários em relação ao seus próprios processos, permitindo auto-crítica e evolução da maneira de trabalho, além de permitir que falhas do processos fossem mapeadas e discutidas.

Um resultado importante, além do desenho final do *Star Schema* e da sequência de técnicas sugeridas no capítulo 5, foi a determinação da maturidade da empresa para o uso de uma ferramenta de Data Warehouse, diminuindo o risco de uma implementação fracassada. No estudo de caso feito, constatamos que uma solução de Data Warehouse para DIR pode ser ainda prematura, dado o estágio em que se encontra em termos de fluxo de informação. Seu

processo poderia evoluir mais com a implementação de uma ferramenta de *workflow*, como notado pelos próprios usuários e melhor determinação de alguns fluxos de processos e aumento da qualidade de dados existentes nos sistemas atuais. Após alguns ajustes em seus processos atuais o desenvolvimento da solução de Data Warehouse seria mais fácil e menos custoso. Entre as sugestões de melhoria estão:

- Determinação de uma política de qualidade para a confecção de relatórios dos técnicos;
- Ampliação das informações contidas na ferramenta SIAP, para facilitar o estudo do perfil de público da DIR;
- Melhor acompanhamento do ciclo de vida dos processos administrativos, se possível com o auxílio de uma ferramenta de *workflow*;
- Uma auditoria de qualidade nas informações migradas para os sistemas de uso atual, para diminuir as inconsistências de dados;
- Implementação de controles computadorizados para insumos e horários de técnicos.

Como proposta de trabalhos futuro, está a implementação total de uma ferramenta de Data Warehouse utilizando-se como apoio as técnicas da Semiótica Organizacional, incluídas as que não foram utilizadas nesse trabalho como o NAM e SAM. Dentro de um projeto de implementação demorado e dinâmico como o de uma solução de Data Warehouse, acreditamos que essas técnicas serão de grande valia nas seguintes etapas do processo:

- Definição da melhor tecnologia de implementação para o escopo da empresa;
- Definição das interfaces de navegação para os usuários;
- Definição das regras de negócio formais da empresa;
- Definição das regras de extração e transformação da camada de ETL, permitindo a consistência e homogeneização dos dados;
- Definição das regras de *Drill-Down*;
- Definição dos níveis de granularidade de dados.

Ao final de um estudo de implementação de uma ferramenta de Data Warehouse, seremos capazes de verificar se as técnicas propostas pela Semiótica Organizacional serão tão efetivas para etapas mais avançadas no processo de desenvolvimento quanto foram para a elicitação inicial de requisitos.

Referências Bibliográficas

Anahory, S., Murray, D., Data Warehousing in the real World – a Practical Guide for Building Decision Support systems, Addison-Wesley Publishing 1997 apud Bruckner, R. M., List, B.,Schiefer, J., 2001, Developing Requirements for Data Warehouse Systems with Use Cases, Seventh Americas Conference on Information Systems, pg. 329.

Barthes, R., 1964. Elements of Semiology. London: Jonathan Cape apud Chandler, D., 2000, Semiotics for Beginners. Hypertexto disponível em <http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>. Acessado em novembro 2005.

Bruckner, R. M., List, B.,Schiefer, J., 2001, Developing Requirements for Data Warehouse Systems with Use Cases, Seventh Americas Conference on Information Systems, pg. 329.

Chandler, D., 2000, Semiotics for Beginners. Hypertexto disponível em <http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>. Acessado em novembro 2005.

Dale, M., 2004, Defining User requirements for a Large Corporate Data Warehouse: An Experiential Case Study. AWRE'04 9th Australian Workshop on Requirements Engineering, pg. 5.1.

Giannocaró, A.,Shanks, G.,Darke, P., 1999, Stakeholders Perceptions of Data Quality in a Data Warehouse Environment. Proc. 10th Australian Conference on Information Systems, pg. 344.

Gorla, N., 2003, Features to Consider in a Data Warehouse System. Communications of the ACM. Novembro 2003/Vol.46, pg. 111.

Inmon, W.H., 1996. Building the Data Warehouse. Wiley & Sons.

Inmon, W.H., ND a. Data Warehouse and Data Mining. [Web], disponível em <http://www.inmongif.com/> [Consultado em setembro 2004].

Inmon, W.H., ND b. Definition of a Data Warehouse. [Web], disponível em <http://www.inmongif.com/> [Consultado em setembro 2004].

Inmon, W.H., ND c. Gathering DSS/Data Warehouse Requirements. [Web], disponível em <http://www.inmongif.com/> [Consultado em setembro 2004].

Inmon, W.H., ND d. The Data Warehouse Evolution Into the Millenium. [Web], disponível em <http://www.inmongif.com/> [Consultado em setembro 2004].

Inmon, W.H., ND e. The Problem with Dimensional Modeling. [Web], disponível em <http://www.inmongif.com/> [Consultado em setembro 2004].

Inmon, W.H., ND f. What is a Data Warehouse? [Web], disponível em <http://www.inmongif.com/> [Consultado em setembro 2004].

Inmon, W.H., Terdeman, R.H., Imhoff, C., 2001, Data Warehousing, Berkeley Brasil
Kimball, R., 1998, Data Warehouse Toolkit, Makron Books do Brasil.

Liu, K., 2000, Semiotics in Information Systems Engineering. Cambridge University Press. Cambridge.

Mancuso, G., Moreno, A., 2002, The Role of OLAP in the Corporate Information Factory. DM review, disponível em www.dmreview.com.

McFadden, F. R., 1996, Data Warehouse for EIS: some Issues and Impacts. Proc. of the 29th Annual Hawaiian International conference on system sciences. Pg. 120.

Morris, C. W., 1970, Foundations of the Theory of Signs. Chicago: Chicago University Press. Apud Chandler, D., 2000, Semiotics for Beginners. Hypertexto disponível em <http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>. Acessado em novembro 2005.

Moss, L.,Barbusinski, L.,Rehm, C.,Oates, J., 2004, Ask The Experts, Forum público pela dmreview, disponível em <http://dmreview.com/>. [Consultado em julho 2005].

OSW (1995) "The circulation Document". Organizational Semiotics Workshop. Apud Liu, K., 2000, Semiotics in Information Systems Engineering. Cambridge University Press. Cambridge.

Paim, F., 2003, Uma Metodologia para Definição de Requisitos em Sistemas Data Warehouse. Dissertação de Mestrado. Universidade Federal de Pernambuco.

Paz, L.C., Cielo, I.,1999-2000, DW!, disponível em <http://www.datawarehouse.inf.br/> acessado em outubro 2004.

Peirce, C.S., (1931-58): Collected Writings (8 Vols.). (Ed. Charles Hartshorne, Paul Weiss & Arthur W Burks). Cambridge, MA: Harvard University Press apud Chandler, D., 2000, Semiotics for Beginners. Hypertexto disponível em <http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>. Acessado em novembro 2005.

Price, R. J.,Shanks, G.,2004, A Semiotic Information Quality Framework. Decision Support in an Uncertain and Complex World: The IFIP TC8/WG8.3 International Conference 2004, pg. 658.

Saussure, F.,1983, Course in General Linguistics (trans. Roy Harris). London: Duckworth
apud Chandler, D.,2000, Semiotics for Beginners. Hypertexto disponível em
<http://www.aber.ac.uk/media/Documents/S4B/semiotic.html>. Acessado em novembro
2005.

Shanks, G.,Corbit, B.,1999, Understanding Data Quality: social and Cultural Aspects. Proc.
Of the 10th Australian Conference on Information Systems.pg. 785.

Shanks, G.,O'Donnell, P.,Designing a Data Warehouse: Combining Entity Relationship and
Dimensional Modeling. White Paper, Monash University, Australia.

Simoni, C.A.C.,2003, A Prática do Desenvolvimento de Software e a Abordagem da
Semiótica Organizacional, Dissertação de Mestrado. Unicamp.

Simoni, C.A.C.,Baranauskas, M.C.C.,Da Análise de Requisitos ao Projeto de Interface:
uma Abordagem Subjetivista para Sistemas de Informação. Universidade Estadual de
Campinas, Unicamp.

Stamper, R.K., 1973, Information in Business and Administrative Systems. John Wiley and
Sons, New York apud Liu, K., 2000, Semiotics in Information Systems Engineering.
Cambridge University Press. Cambridge.

Stamper, R.K., 1993, Social Norms in requirements analysis – an Outline of MEASUR. In
Jirotko, M., Goguen, J. and Bickerton,M., Requirements Engineering, Technical and Social
Aspects. Academic Press, New York apud Simoni, C.A.C.,2003, A Prática do
Desenvolvimento de Software e a Abordagem da Semiótica Organizacional, Dissertação de
Mestrado. Unicamp.

Winter, R., Strauch, B., 2004, Information Requirements Engineering for Data Warehouse
Systems, ACM Symposium on Applied Computing, pg. 1359.

Winter, R., Strauch, B., 2003, A Method for Demand-driven Information Requirements Analysis in Data Warehousing Projects. Proceedings of the 36th Hawaii International conference on System Sciences, IEEE 2003.

|

Anexos

Anexo A – Documento da DIR

“A administração pública, em especial o setor saúde, desde a década de 80, vem passando por um processo de mudanças estruturais e conceituais, alcançando o ponto de maior representatividade do movimento da reforma sanitária na VIII^a Conferência de Saúde em 1986, fórum impar que reuniu todos os seguimentos representativos da sociedade que, preocupados com os resultados, forma da estrutura, funções que do sistema de saúde vigente (modelo autoritário e centralizador), conseguiram consolidar e aprovar propostas resultantes das discussões que aconteciam nas universidades e na sociedade organizada, que respaldaram a Constituição Federal de 1988, marco divisor dessas transformações, definindo que o setor saúde deve organizar-se num Sistema Único de Saúde – SUS.

O ano de 1983 marcou o início da democratização do país, no Estado de São Paulo assume como 1º governador eleito pelo povo o Dr. André Franco Montoro tendo como secretário da Saúde o Dr. João Yunes, com a seguinte proposta: “Não me proponho governar como se fosse possível e fácil resolver, da noite para o dia, a crise que atravessamos. Mas sei que é grande o potencial de recursos humanos e produtivos de nosso Estado, e conheço a capacidade de trabalho dos brasileiros que aqui vivem. Se unirmos São Paulo em torno da idéia generosa de um desenvolvimento baseado em nossos próprios recursos – um desenvolvimento cujo centro seja a pessoa humana – iniciaremos um movimento de transformações sociais e políticas que há de marcar uma geração, em nosso Estado e no País”. MONTORO (1987)

A Secretaria do Estado da Saúde em 1986 acreditando na reforma sanitária pregada há muito pelos sanitaristas e que ganhou força na 8ª Conferência de Saúde, fez uma reforma administrativa organizacional mudando radicalmente o Organograma da Secretaria, preparando-se para o Sistema Unificado e Descentralizado de Saúde atual SUS.

A estrutura da Secretaria do Estado da Saúde vigente na época datava de 1969 (Decreto nº 52.182, de 16 de julho de 1969) era fracionada em diferentes organizações,

vários centros de decisões que trabalhavam vinculados ao poder central, modelo centralizador do período de exceção. Compreendia: Conselho Estadual de Saúde, Gabinete do Secretário de Estado, Conselho Técnico-Administrativo, Grupo de Planejamento Setorial, Consultoria Jurídica, Departamento Técnico Normativo, Coordenadoria de Saúde da Comunidade, Coordenadoria de Assistência Hospitalar, Coordenadoria de Saúde Mental, coordenadoria de Serviços Técnicos Especializados, Departamento de Administração da Secretaria.

Os pressupostos da reforma sanitária já estavam colocados. As mudanças tinham que ser transformadoras e abranger toda a Secretaria da Saúde. O Decreto nº 25.519, de 17 de julho de 1986 trazia a nova estrutura com mudanças profundas, voltando a administração para uma lógica organizacional, estabelecendo órgãos centrais que fossem técnicos, operacionais. Acabava-se com as Coordenadorias e criavam-se os Escritórios Regionais de Saúde – Ersas, responsáveis pela integralidade da saúde em sua área de abrangência.

A Vigilância Sanitária padecia dos mesmos males da estrutura da Secretaria da Saúde, era uma vigilância departamentalizada, centralizadora, cartorial, burocrática, extremamente policialesca, não integrada, fracionada, não existia uma visão da ação integrada de processo de trabalho. Desde sua criação até 1985, mais de quarenta anos, muito pouco se sabe sobre este período, não existem registros e as informações perderam-se junto com os funcionários.

Em 1985/1986 conhecia-se a razão de mudar e o que se quer mudar. Criou-se o Centro de Vigilância Sanitária – CVS (Decreto 26.048, de 15 de outubro de 1986), órgão normativo subordinado diretamente a Secretaria da Saúde.

O Município de São Paulo dadas às dimensões territoriais foi dividido num primeiro momento em 8 (oito) ERSAS – Escritórios Regionais de Saúde. Em 1995 nova modificação estrutural dividiu a cidade em 5 núcleos regionais de saúde - os NRS de 1 a 5, subordinados a Divisão Regional de Saúde – I – DIR – I (Decreto nº 40.082 e Decreto 40.083, de 15 de maio de 1995).

Em 1995 há nova mudança de governo, sai o governador Fleury e assume o Dr. Covas. Na Vigilância Sanitária existiam muitas dificuldades, resquícios do velho modelo conviviam com o novo, ainda de maneira prevalente, havia uma forte reação a

descentralização, a municipalização, com forte resistência ao trabalho de equipe multidisciplinar.

Em particular na Capital de São Paulo, a Secretaria reconhecia em 1995 o resultado insatisfatório das ações de vigilância sanitária, com muitas denúncias e uma crescente insatisfação dos usuários do serviço, com graves problemas gerenciais dessas ações.

A VISA Capital passou por transformações, mudança de diretores, caminhou na mudança do modelo, na lógica das ações enfocadas na avaliação de risco.

Em 2003 a VISA Capital passou por uma modificação radical na organização física e funcional quando os cinco núcleos regionais de vigilância sanitária, cada um com um perfil sócio-econômico e uma vocação predominante foram unificados e centralizados num mesmo endereço. Cada Núcleo Regional dispunha de equipes técnicas auto suficientes que executavam as mesmas ações cada qual em sua área de abrangência.

Esta autonomia trazia muitas dificuldades, pois não existiam padronizações de procedimentos, as informações prestadas aos usuários eram diferentes de acordo com a região da cidade. Não existia supervisão sistemática pelo órgão central e as disparidades de situações encontradas indicavam a necessidade de reformulações.

A centralização exigiu a reorganização das equipes e uma nova organização e divisão de trabalho.

Num primeiro momento isto foi muito traumático, primeiro porque não houve preparação, todos foram pegos de surpresa e a mudança se deu de maneira improvisada, em tempos diferentes. A mudança física dos cinco núcleos demorou quatro meses para se efetivar e quando se consolidou os técnicos queriam a continuidade da situação anterior, o que obviamente não era possível.

A resistência às mudanças foi muito forte, a nova proposta de se trabalhar por área de interesse dentro da totalidade da área de abrangência (município de São Paulo) melindrou muitos profissionais que estavam acostumados com seu diretor e um espaço urbano delimitado, para romper a lógica de trabalho anterior foi muito desgastante e requereu muitas reuniões técnicas.

Não havia mais domínio do território, os grupos foram reformulados, e tiveram que se recompor em equipes para novos desafios.

Na Vigilância Sanitária todas estas mudanças que ocorreram nos últimos dez anos, tanto de paradigma como de localização física, prejudicaram a organização de dados, muitas informações foram perdidas ou polarizadas em arquivos sem controle. Cada Núcleo registrava as informações de uma maneira, alguns utilizavam um programa de cadastro e registravam aleatoriamente, outros utilizavam fichas e livros para fazê-lo. Diante da situação foi necessário pensar num sistema que trouxesse as informações mínimas necessárias. Foi desenvolvido o SIAP – Sistema de Informação ao Público que embora esteja em rede para consulta dos funcionários, padece de problemas, pois os antigos sistemas foram transportados para o novo com todas as discrepâncias, o que dificulta a consulta pela não padronização dos dados.

Somente em setembro de 2003 teve início a sistematização das informações com o cadastramento dos dados das empresas no SIVISA – Sistema de Informação em Vigilância Sanitária, ou seja, há dois anos a equipe está se familiarizando com o sistema e cadastrando todas as informações em rede.

Em Abril de 2004 as ações de baixa e média complexidade foram municipalizadas ficando sob a responsabilidade da VISA Capital as ações de alta complexidade, de acordo com o rito constitucional.

A Secretaria de Estado da Saúde passa por uma reestruturação a semelhança do que ocorreu em 1986, preparando-se para a função primordial de gestor estadual, uma vez que o maior município do Estado assumiu parte das ações de vigilância sanitária, todas as ações de vigilância epidemiológica, a integralidade da Assistência Básica. A nova estrutura tem a característica de centralizar a coordenação de atividades correlatas ou similares num órgão objetivando a otimização, transparência e o controle dos recursos disponíveis no atendimento as necessidades da população.

Isto muda novamente a hierarquia e a subordinação da VISA Capital. Tecnicamente a Vigilância Sanitária é subordinada ao CVS, administrativamente a DIR-I atualmente ao GSAE, presentemente se tem a notícia que a Secretaria estuda nova modificação com a recriação da DIR-I Capital e a possível vinculação da VISA a CCD – Coordenadoria de Controle de Doenças. (organograma atual)

As demandas chegam na VISA Capital por vários caminhos, seja através de denúncias de cidadãos, de atividades programadas, de solicitação de licença de funcionamento,

certificação em boas práticas de fabricação, ou solicitação de outros órgãos como Ministério Público, Autoridade Policial, Poder Judiciário entre outros. A interface da VISA Capital com outros órgãos é muito abrangente dadas às características do território que atua.

Qualquer que seja a demanda, o primeiro passo é protocolar o pedido na Central de Atendimento ao Público – CAP, onde a mesma é registrada no sistema SIAP – Sistema de Informação de Atendimento ao Público e recebe um número de protocolo e, se for o caso, um número de processo com o qual o interessado acompanhará sua solicitação. Verificada a natureza do pedido o mesmo é encaminhado para análise e providências de acordo com o status: todas as solicitações de empresa ou serviços como Licença Inicial ou renovação da licença são encaminhados para o setor de expediente para cadastramento no SIVISA - Sistema de Informação em Vigilância Sanitária. As solicitações de avaliação de projeto ficam cadastradas somente no SIAP.

A Licença de Funcionamento vale por um ano a partir da data de deferimento, devendo ser renovada a cada período de validade. A solicitação é acompanhada por documentos e quando se tratar de renovação deve ser pedida 60 dias antes do término da validade da licença. Na licença inicial são exigidos mais documentos inclusive projeto da edificação, pré-requisito para a obtenção da licença de funcionamento.

Após o cadastramento o processo ou protocolado será enviado para o setor competente para análise dos documentos e vistoria inicial por equipe multiprofissional quando serão avaliados os fluxos dos procedimentos, as rotinas e regras para asseguramento dos padrões de qualidade quando se tratar de prestação de serviços de saúde e de boas práticas de fabricação quando se tratar de produtos.

Da vistoria podem resultar várias situações que serão analisadas pela autoridade sanitária que proporá a melhor medida ou recomendação em sua função da fiscalizadora.

As dificuldades se multiplicam, desde a junção em 2003 a situação de viaturas se agravou, hoje o número máximo de veículos disponibilizados para a execução das ações são quatro viaturas por dia para fazer as vistorias quando nos núcleos existiam onze viaturas para o mesmo território, os equipamentos de informática disponíveis não dão conta da demanda, quebram com frequência ou cai o sistema, perdendo todo o relatório, há muita insatisfação da equipe.

Os registros são feitos no sistema sem padronização, muitos dados foram copiados de sistema antigos, com maneiras diferentes de registro, dificultando a consulta no sistema. Há momentos que o funcionário precisa ser “mágico” para conseguir a informação e muitas se perderam no caos das mudanças.

Os arquivos, importantes fontes de consulta estão parte no prédio da Av. São Luiz e estão sendo organizados e parte, espalhados por outros prédios sem controle e sem acerto de onde encontrar.

Não existe domínio dos dados, os dois sistemas SIAP e SIVISA não se comunicam, portanto não se tem cruzamento de informações, conseqüentemente as equipes dispõem mais esforço e tem mais trabalho para obter as informações. Em contrapartida a diretoria tem muitas dificuldades em gerenciar as atividades e planejar as ações.

Levando-se em conta as colocações acima e, acrescentando a isso as complexidades dos serviços e das empresas, é possível estabelecer um planejamento das ações de visa, desde que se conheça o universo da atuação, as pessoas envolvidas na execução das ações, e os interlocutores do processo.

Planejar as ações de visa é desejável e possível desde que alguns pré-requisitos sejam atendidos como sistemas de informação que permitam cruzamento de dados, capacitação ou reciclagem dos profissionais, pois as tecnologias em saúde são ágeis e os profissionais não são capacitados com a mesma agilidade pelo poder público.

Faz-se necessário zerar um conjunto de pendências que interferem muito na execução das ações, como a questão de ter ou não combustível, ter ou não motorista para se ter agilidade nas ações de visa.“.

Anexo B – Resultados do Quadro de Avaliação

1- Camada de Contribuição

Tabela 2 : Resultados do Quadro de Avaliação - Camada de Contribuição

Contribuição			
Partes Interessadas	Condições/Efeitos	Questões/Problemas	Possíveis Soluções
Diretoria / Assistência Técnica	Saber/conhecer as ocorrências cadastradas Controle do Trabalho dos técnicos	<p>Não há como encaminhar as ocorrências sem conhecimento e sem apoio das áreas técnicas</p> <p>Dificuldade em localização de documentos</p> <p>Falta de dados que permitam o gerenciamento e planejamento para cobrança de retorno e providências.</p> <p>Muito embora as certificações sejam ligadas às licenças, não há como cruzar as informações.</p> <p>Falta registro dos relatórios de certificação para subsidiar decisão.</p>	<p>Formalização de dúvidas e maneira de acesso à Diretoria, respeitando os níveis hierárquicos.</p> <p>Disponibilização de:</p> <ul style="list-style-type: none"> • Com quem está o documento • Quantos documentos estão com qual técnico/área • A quanto tempo está sem solução/encaminhamento • Aonde ficou parado(gargalo) • Porque não teve encaminhamento (ex: por falta de viatura) <p>Melhor controle das empresas por projeto</p>

<p>Coordenadorias de Área</p>	<p>Precisa saber o que o técnico faz, e como e quando. Cadastro provisório(deferido ou indeferido) pré-requisito para a licença</p>	<p>Não consegue monitorar auto de infração por irregularidades de estabelecimento Convocação dos técnicos pelo CVS para certificação, ficando a DIR sem pessoal para manter as rotinas Falta de apresentação de relatórios de vistoria (sivisa) por parte de todos os técnicos. Falta de apresentação do resultado de trabalho do técnico atingindo a meta solicitada, mostrando a falta de envolvimento do técnico com o tema abordado. Grande demanda de trabalho sem apoio administrativo, decorrendo na falta de tempo para análise, avaliação e planejamento do trabalho. Grande número de técnicos, grande número de casos com complexidades diferentes Localização de documentos Falta de padronização de trabalho Carga horária baixa e não cumprida Carga horária diferente por categoria profissional Falta de compromisso com o serviço e com as informações técnicas</p>	<p>Registro da hora técnica do funcionário disponibilizado para a certificação em Boas Práticas, para a negociação dos dias disponibilizados. Treinamento ou integração de todos para a demonstração de resultado homogêneo (relatório sivisa) através de reuniões de avaliação com o grupo. Reuniões com o grupo para a integração e uniformização do trabalho. Melhor visibilidade dos casos por técnico e controle de retorno dos das demandas encaminhadas para os técnicos(frequencia) Disponibilização de : <ul style="list-style-type: none"> • Com quem está o processo • Quantos documentos está com cada técnico/área • A quanto tempo está sem solução/encaminhamento • Aonde ficou parado(gargalo) • Por que não teve encaminhamento Trabalhar com planejamento e cronograma- preciso planejamento integrado das coordenadorias</p>
-------------------------------	--	--	--

Técnicos	<p>Conhecer normas regulamentares</p> <p>Instrumentos – viaturas, máquinas fotográficas, equipamentos medidas</p> <p>Coordenação – apoio técnico</p> <p>Vistorias – composição da equipe.</p> <p>Proposta de liberação ou não da licença</p> <p>Relatório de vistorias para a orientação das empresas e das coordenadorias.</p>	<p>Falta de equipamento de informática</p> <p>Falta de insumos</p> <p>Falta de técnicos de outras áreas para avaliar os projetos</p> <p>Falta de apoio para atividades administrativas</p> <p>Falta de treinamento e disponibilização de novas legislações</p> <p>Falta de apoio técnico para a decisão de viabilidade de uma solicitação</p> <p>Perda de dados informatizados.</p>	<p>Disponibilização de agendas de insumos</p> <p>Gravação de dados centralizados</p>
Suporte de informação/sistemas	Solicitação de informação (relatórios) pelas áreas	Saber entender as reais necessidades dos usuários	Emissão de relatórios facilitados por ferramentas
Protocolo/Expediente	<p>Interface (porta voz) da instituição – técnicos X usuários</p> <p>Porta de entrada / Saída</p> <p>Resposta à solicitação de usuários</p> <p>Registro e demanda para a área técnica.</p>	<p>Faltam condições de treinamento e padrões.</p> <p>Falta padronização de procedimentos para uniformidade de atendimento e respostas aos usuários (público)</p> <p>Autonomia para a decisão pela situação de regularidade da empresa por insuficiência operacional da administração</p>	Registro de todas os atendimentos (demandas), inclusive de balcão e telefone.

2- Camada de Fontes

Tabela 3 : Resultados do Quadro de Avaliação - Camada de Fonte

Fontes			
Partes Interessadas	Condições/Efeitos	Questões/Problemas	Possíveis Soluções
Anvisa	Define regras gerais de vigilância sanitária para todo o Brasil Autorizam funcionamento e registro de produtos e certificação	Falta de controle e atendimento dos prazos para encaminhamento de autorização de funcionamento Não ter acesso ao cadastro de produtos, o que causa retrabalho e dificulta o fluxo	Registro e controle dos prazos Acesso aos registros de produtos, pendências e do histórico do produto
Polícia	Demanda denúncias Solicita informações	Não se sabe de todos os casos que chegam através da polícia e os responsáveis não têm controle sobre os funcionários e destas ações	Fazer cadastro dessas informações
Procon	Solicita informações	Não ter relatório /registro da empresa ou produto	Fazer cadastro dessas informações
Assessoria de imprensa	Responde a reportagens e denúncias nos meios de comunicação Solicita informações de assuntos de nossa competência	Informação muito dispersa ou inexistente	Centralização das informações
Conselhos de classes	Fazem denúncias sobre estabelecimentos sob regime da vigilância Informam sobre penalidade profissional	Extrapolam a competência – Não tem poder de polícia	Mediação pelo ministério do trabalho sobre conflito de competência
Secretaria da Fazenda	Valores e taxas Cadastro nacional Cadastro para efeito fiscal Fornecem número Cnae	Não têm acesso a informações como cadastros fiscais, CNPJ e tipo de estabelecimentos.	Estar cadastrado (replicado) como informação da receita

Dir	Dir regionalizadas Troca de informações dependendo da localidade do estabelecimento	As Dir's não tem compartilhamento de informação	Ter um acordo de informação ou informações compartilhadas
Ministério Público	Solicitam informações sobre empresas sujeitas ao regime de vigilância	Controle das solicitações (ofícios) Demora no retorno com várias reiteraões	Melhora no controle das demandas (ofícios)
Poder Judiciário	Solicitam informações para a instrução de processos que envolvem nossa competência legal	Compilar as informações Obter a informação	Refinamento do cadastro e melhor cruzamento de informações.

3 – Camada de Mercado

Tabela 4 : Resultados do Quadro de Avaliação - Camada de Mercado

Mercado			
Partes Interessadas	Condições/Efeitos	Questões/Problemas	Possíveis Soluções
Prefeitura	Prestam mesmo serviço em complexidades diferentes definidas em Resolução 30/94	Falta de conhecimento por parte da prefeitura de certos aspectos da competência legal de Estado/Município	Esclarecer melhor casos encaminhados errados
Anvisa	Agência reguladora, com competência e foco diferente	Sem comentários	Sem comentários
Visas	Utilizam a DIR 1 como foco de referência	Sem comentários	Sem comentários

4 – Camada de Comunidade

Tabela 5 : Resultados do Quadro de Avaliação - Camada de Comunidade

Comunidade			
Partes Interessadas	Condições/Efeitos	Questões/Problemas	Possíveis Soluções
Governo	Sem comentários	Sem comentários	Sem comentários
Imprensa	Formação de opinião da população	Manter “boa imagem” perante o público	Bom trabalho de base da assessoria de imprensa
Ministério da Saúde	Sem comentários.	Sem comentários	Sem comentários

Anexo C – Resultados do Framework Semiótico

Tabela 6 : Resultados do Framework Semiótico

NÍVEL	DESCRIÇÃO	POSSÍVEIS SOLUÇÕES
<p>MUNDO SOCIAL Crenças, funções, compromissos, contratos, leis, cultura...</p>	<ul style="list-style-type: none"> • Melhoria da qualidade dos projetos • Melhoria do atendimento aos usuários • Melhor clareza nas informações • Diminuição do tempo de resposta • Melhoria da qualidade dos serviços/produtos de interesse à saúde 	<ul style="list-style-type: none"> • Organizar e implementar o manual de atendimento a usuário da portaria 16 (isso é uma portaria de lei) • Organizar e inferir um sistema de “cobrança” de resultados
<p>PRAGMÁTICO Intenções, comunicações, conversações, negociações...</p>	<ul style="list-style-type: none"> • Confiança e satisfação do usuário • Resposta ágil e conclusiva • Melhor aferição dos resultados dos técnicos, para efeito de prêmio de incentivo, visando melhorar atendimento, produtividade e resolubilidade. • Melhorar a avaliação dos projetos • Avaliação qualitativa das avaliações técnicas 	<ul style="list-style-type: none"> • Avaliação do técnico (produtividade, colaboração, prazo, resolubilidade) • Avaliação do Prêmio de incentivo • Adequação dos técnicos à nova realidade pós municipalização
<p>SEMÂNTICO Significados, proposições, validade, verdade, significação, denotação...</p>	<ul style="list-style-type: none"> • Padronizar a linguagem dos relatórios de parecer técnicos • Clareza e objetividade nos pareceres dos processos 	<ul style="list-style-type: none"> • Utilizar esses dados para o prêmio de incentivo • Padronizar as informações do Protocolo e dos técnicos para o público em geral

Anexo D – Resultados da análise Colateral

Tabela 7 : Resultados da Análise Colateral

Ciclo de Vida
Predecessor <ul style="list-style-type: none">- Sivisa- Siap- Planilhas Excel- Trabalho Manual
Sistema Focal <p>Sistemas de Banco de Dados Sistemas de Data Warehouse Interfaces</p>
Sucessor <p>Sistemas de Workflow Sistemas de acompanhamento de processos via Web.</p>
Funcionamento
Entrada <ul style="list-style-type: none">- Razão Social, Nome Fantasia, CNPJ, Endereço, Tipo de Estabelecimento/Atividade, CNAE fiscal, Responsável Legal, Responsável Técnico, Serviços Terceirizados.- Tipo de demanda em caso de produto: autorização de funcionamento, dados complementares, última licença emitida e data de validade.- Relatórios: Descrição da empresa, descrição da situação encontrada, Recomendações, providências, conclusão.- Relatórios sivisa- Recursos: Tipo do recurso, uso do recurso, quem, quando e para quê.- A demanda foi pra quem, quando foi recebida, quando retornou, quando foi feita a vistoria, quando foi feito o relatório.- Estabelecimento: Cadastro, ciclo de vida e de operação, Histórico de requisições, tipo de atividade.
Saída <ul style="list-style-type: none">- Facilidade de geração de relatórios com base em informações existentes de forma não linear.- Controle de solicitação: tempo de saída, tempo para relatório, tempo com o técnico (pediu viatura quando, fez a vistoria quando, qual era a equipe?)- Visualização de relatório: Técnico X Licença X Risco- Informações totalizadoras por estabelecimento: região, tipo de trabalho, porte, registro, complexidade.
Ambiente <ul style="list-style-type: none">- Centralizado, seguro, interface facilitada.- Software livre, sempre que possível.