

Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz



ESCOLA NACIONAL DE SAÚDE PÚBLICA  
SERGIO AROUCA  
ENSP

***“Uso de informações de saúde para suporte à decisão: uma metodologia focada no consumidor da informação”***

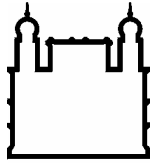
*por*

***Luiz Fernando de Aguiar Azevedo***

*Dissertação apresentada com vistas à obtenção do título de Mestre  
Modalidade Profissional em Saúde Pública.*

*Orientador: Prof. Dr. Álvaro César Nascimento*

*Rio de Janeiro, dezembro de 2009.*



Ministério da Saúde

**FIOCRUZ**  
**Fundação Oswaldo Cruz**



ESCOLA NACIONAL DE SAÚDE PÚBLICA  
SERGIO AROUCA  
ENSP

*Esta dissertação, intitulada*

***“Uso de informações de saúde para suporte à decisão: uma metodologia focada no consumidor da informação”***

*apresentada por*

***Luiz Fernando de Aguiar Azevedo***

*foi avaliada pela Banca Examinadora composta pelos seguintes membros:*

Prof. Dr. Sergio Miranda Freire  
Prof.<sup>a</sup> Dr.<sup>a</sup> Patricia Tavares Ribeiro  
Prof. Dr. Álvaro César Nascimento - Orientador

Catálogo na fonte  
Instituto de Comunicação e Informação Científica e Tecnológica  
Biblioteca de Saúde Pública

A994 Azevedo, Luiz Fernando de Aguiar  
Uso de informações de saúde para suporte à decisão: uma metodologia focada no consumidor da informação. / Luiz Fernando de Aguiar Azevedo. Rio de Janeiro: s.n., 2009.  
79 f.

Orientador: Nascimento, Álvaro César  
Dissertação (mestrado) – Escola Nacional de Saúde Pública Sergio Arouca, Rio de Janeiro, 2009

1. Sistemas de Informação - utilização. 2. Tecnologia da Informação. 3. Disseminação de Informação. 4. Software. 5. Validação de Programas de Computador. I. Título.

CDD - 22.ed. – 303.4833

*Para meu filho,  
Fernando*

## **Agradecimentos**

Ao Prof. Dr. Miguel Murat Vasconcellos (in memorian) e à Prof<sup>ª</sup> Dr<sup>ª</sup> Ilara Hämmerli Sozzi Moraes, pela idealização e concretização deste curso.

Ao Prof. Dr. Álvaro César Nascimento, por aceitar o desafio de orientar um projeto já iniciado.

A Alex Príncipe, pela orientação no curso de especialização, que levou a esta dissertação.

Aos colegas, pelo auxílio mútuo e por tornarem este tempo de curso bem mais leve.

Aos professores, pela generosidade e interesse em compartilhar seus conhecimentos.

A meus pais, Gilson (in memorian) e Eunice, sempre presentes, por absolutamente tudo!

À minha esposa, Mônica, pelo amor e compreensão, e por abrir mão do nosso tempo durante o curso.

A meu filho, Fernando, por me provar que o dia tem bem mais que 24 horas.

A Deus (mais uma vez!)

*“O homem precisa de utopias,  
e direi mesmo:  
- são umas quatro ou cinco utopias  
que ainda nos salvam.”  
(Nélson Rodrigues)*

*“O homem que luta pelos seus ideais  
é o homem que está vivo”  
(Cervantes)*

## Resumo

O objetivo deste trabalho é apresentar os passos necessários para a construção de *data marts / data warehouses* como uma solução para um ambiente de suporte à decisão. Seu foco não é, entretanto, o aprofundamento de cada uma das etapas deste método (uma vez que existe uma ampla literatura sobre o assunto), mas realçar a importância do envolvimento do usuário – aqui denominado consumidor da informação – durante o processo de criação e manutenção destes *data marts / data warehouses*. O consumidor da informação é chamado a participar de um plano integrado com os membros das áreas detentoras do conhecimento necessário para a construção destas soluções de suporte à decisão, incluindo a área de tecnologia da informação. A disseminação das informações contidas nestes bancos de dados para os diversos tipos de consumidores da informação (com diferentes recursos de *hardware*, *software* e humanos disponíveis), e sua aplicação no controle social, também são discutidas aqui.

**Palavras-chave:** Ministério da Saúde; suporte à decisão; usuário; data mart; data warehouse; armazém de dados; tecnologia da informação.

## **Abstract**

The purpose of this work is to present the required steps for the construction of data marts / data warehouses as a solution for a decision support environment. However, its focus is not to go deeper in each of the steps of this method (since there is a broad literature about the subject), but to highlight the importance of the user – here called information consumer – engagement during the process of creating and maintaining these data marts / data warehouses. The information consumer is called to take part in an integrated plan together with the members of the areas who own the necessary knowledge to build these decision support solutions, including the information technology (IT) area. The spread of the informations contained in these databases to the different types of information consumers (with different hardware, software and human resources available), and its application in social control, is also discussed here.

**Keywords:** Ministry of Health; decision support; user; data mart; data warehouse; information technology.



## Sumário

<b>1. Introdução .....</b>	<b>10</b>
1.1 <i>Objetivo geral e objetivos específicos</i> .....	13
1.2 <i>Metodologia</i> .....	14
1.3 <i>Contextualização</i> .....	16
<b>2. Informação, saúde e conhecimento .....</b>	<b>24</b>
2.1 <i>A informação</i> .....	24
2.2 <i>Ciência da informação e informação em saúde</i> .....	25
2.3 <i>Interdisciplinaridade e transdisciplinaridade</i> .....	26
2.4 <i>Inteligência competitiva e o metaconhecedor</i> .....	30
<b>3. Os sistemas de informação .....</b>	<b>32</b>
3.1 <i>Informações em saúde e os sistemas de informação</i> .....	32
3.2 <i>Problemas: integridade, fragmentação, integração</i> .....	33
<b>4. O usuário como consumidor da informação .....</b>	<b>39</b>
<b>5. A informação e o controle social .....</b>	<b>44</b>
<b>6. Um método para construção de data warehouses – etapas principais .....</b>	<b>50</b>
6.1 <i>Etapa 1: A orientação temática dos data marts e do data warehouse</i> .....	51
6.2 <i>Etapa 2: A classificação dos consumidores da informação</i> .....	52
6.4 <i>Etapa 4: A elaboração do modelo conceitual dimensional</i> .....	58
6.5 <i>Etapa 5: O projeto físico do banco de dados</i> .....	61
6.6 <i>Etapa 6: Os processos de extração-transformação-carga</i> .....	62
6.7 <i>Etapa 7: A construção das aplicações</i> .....	63
6.8 <i>Etapa 8: Validação e testes</i> .....	64
<b>7. A disseminação das informações .....</b>	<b>66</b>
<b>8. Considerações Finais .....</b>	<b>74</b>
<b>Referências bibliográficas .....</b>	<b>77</b>

## 1. Introdução

O título deste trabalho procura realçar dois pontos na produção de sistemas de informação: a tomada de decisões gerenciais sobre dados estatísticos em saúde e o foco no usuário, ou melhor, no que este autor prefere denominar consumidor da informação.

O termo “consumidor da informação” está relacionado, segundo Gómez e Chicanel<sup>1</sup>, ao conceito de “regime de informação”, cunhado por Bernd Frohmann. Segundo ele, regime de informação é um “sistema ou rede mais ou menos estável na qual a informação flui através de canais determináveis – de produtores específicos, via estruturas organizacionais específicas, a consumidores ou usuários específicos”<sup>2</sup>.

Os conceitos de produtores e consumidores da informação, de Frohmann, são muito úteis quando nos propomos a descrever os sistemas e ambientes de suporte à decisão, pois a variedade de fontes de informação pode ser generalizada sob o rótulo de produtores, e a diversidade de usuários, em seus mais variados níveis, também fica bem caracterizada sob o conceito de consumidores de informação.

A importância dos dois pontos realçados fica maior ainda quando a área à qual se aplicam tais sistemas de informação é a área de saúde pública.

Como mostra Moraes em seu livro “Informações em Saúde – da Prática Fragmentada ao Exercício da Cidadania”<sup>3</sup>, o conceito de sistema de informação vem da associação do conceito de Sistema com o da Informática. Sistema, por sua vez, pode ser definido como “um todo complexo e organizado, uma reunião de coisas ou partes formando um todo unitário e complexo”<sup>4</sup>.

Os sistemas de informação em saúde, por sua vez, como mostra Moraes<sup>3</sup>, segundo o Escritório Regional da OMS para a Europa, são aqueles cujo propósito é selecionar os dados pertinentes aos serviços de saúde e transformá-los na informação necessária para o processo de decisões, próprio das organizações e indivíduos que planejam, financiam, administram, provêm, medem e avaliam os serviços de saúde.

Não é demais lembrar o que determina a Constituição da República Federativa do Brasil, de 1988, em seu Título VIII - Da Ordem Social, Capítulo II, Seção II – Da Saúde, Artigo 196:

“A saúde é direito de todos e dever do Estado, garantido mediante políticas sociais e econômicas que visem à redução do risco de doença e de outros agravos e ao acesso universal e igualitário às ações e serviços para sua promoção, proteção e recuperação.”<sup>5</sup>

É possível observar que esta determinação da Constituição Federal de 1988, aliada ao conceito de sistemas de informação em saúde, visto anteriormente, mostra a importância da precisão, da acurácia dos dados produzidos por estes, que levarão a informações que irão auxiliar os gestores e demais atores do SUS a alcançarem o que preconiza o artigo 196 da Constituição.

Pela ótica do autor deste trabalho, alguns **problemas** podem ser observados nos sistemas de informação em saúde em atividade no Ministério da Saúde.

Em primeiro lugar, como será visto de forma detalhada mais à frente, estão os problemas de integridade, inconsistência e falta de padronização dos dados produzidos por alguns destes sistemas de informação.

Além disso, os produtores de informação, segundo constatado na vivência do próprio autor, raramente se falam e as informações produzidas são difundidas atreladas aos sistemas que as produzem, e não direcionadas à geração de informação para as instâncias de decisão, considerando as suas diversas necessidades e visões.

Diante deste quadro, este estudo parte dos seguintes pressupostos:

- Apesar dos esforços de padronização, como será visto adiante, os sistemas legados e os novos podem apresentar dificuldades de atenderem a padrões de representação de dados, levando a problemas de inconsistência, quando há uma observação sobre o mesmo dado sob a ótica de diferentes sistemas;
- Existem problemas de comunicação entre diferentes áreas produtoras de informação, além de tensionamentos por disputa de poder, que dificultam a resolução dos problemas de inconsistência no momento da produção pelos sistemas de informação;
- Os ambientes de suporte à decisão tendem a ser construídos, isoladamente, sobre as informações produzidas por um único sistema de informação, sob o olhar dos gestores destes sistemas, para atender às demandas informacionais destes, sem a preocupação com outros consumidores interessados nestas informações, e sem incluí-las no contexto das grandes atividades da saúde. Ou seja, os ambientes de suporte à decisão com informações produzidas por um sistema de informação

tendem a reproduzir a lógica, o pensamento e as representações de dados dos gestores e produtores da informação daquele sistema.

Pode-se, então, observar a **necessidade** de um ambiente informacional onde novos sistemas possam ter seus dados difundidos de forma integrada a dados já existentes, oriundos de outros sistemas correlatos, permitindo a personalização das necessidades de informação dos diversos tipos de consumidores.

A motivação para este trabalho e a identificação dos problemas citados não ocorreram de forma súbita ao autor, mas da experiência adquirida ao longo dos dezessete anos de trabalho no Departamento de Informação e Informática do SUS – DATASUS, desde a sua criação a partir da Diretoria de Saúde da DATAPREV, hoje responsável pela área de Internet e pelo sítio do DATASUS.

Neste contexto, foi possível perceber dificuldades dos usuários finais, especialmente na construção de um sistema de apoio à decisão para o Cadastro Nacional de Estabelecimentos de Saúde. A vivência com os desenvolvedores de sistemas de informação de várias áreas evidenciou problemas de falta de padronização dos dados, que dificultam muito o processo de integração dos diversos sistemas de informação em saúde.

Portanto, o **objeto** deste estudo é a elaboração de um método de trabalho para a construção de bases de dados, íntegras e padronizadas, destinadas à difusão de informações de saúde para suporte à decisão, considerando as necessidades e experiências dos consumidores destas informações (gestores, profissionais de saúde, pesquisadores e membros da academia, cidadãos, membros dos Conselhos de Saúde, etc), quando estas se fizerem necessárias ou quando um novo sistema de informação em saúde for desenvolvido.

A **relevância** deste trabalho está no fato dele contribuir para a padronização e integração de dados, orientando a construção de bases de dados e sistemas de apoio à decisão que forneçam informações direcionadas às necessidades dos consumidores de informação, dentro dos diferentes temas ligados ao setor saúde como promoção da saúde, gestão da saúde, regulação das ações de saúde e ações de vigilância em saúde.

Como pode ser observado em alguns trechos da Política Nacional de Informação e Informática em Saúde, em sua versão 2.0, disponível para *download* no *site* do DATASUS:

“Existe uma grande necessidade em avançar, tanto no caminho da integração entre os sistemas de informação, quanto no uso da informática para a melhoria da produtividade e qualidade dos processos de trabalho em saúde, da gestão e do controle social.”<sup>6</sup>

“Embora os sistemas de informação de abrangência nacional tenham sido criados por mecanismos diferentes, ora pelo Ministério da Saúde ora pelo Ministério da Previdência, como, por exemplo, o Sistema de Informações de Mortalidade e o Sistema de Pagamento das Internações Hospitalares, todos tiveram por objetivo atender a necessidade dos órgãos centrais em obter os dados para efeitos imediatos de gestão ou geração de conhecimento em saúde.”<sup>6</sup>

A construção destes sistemas de apoio à decisão pode contribuir para o enfrentamento de um desafio que a Política Nacional de Informação e Informática em Saúde coloca, ou seja, promover “informação de maior confiabilidade para gestão, geração de conhecimento e controle social”<sup>6</sup>.

O **produto** deste trabalho será um método para a produção de bancos de dados de saúde para apoio à decisão, comumente chamados de armazéns de dados de saúde, cuja produção seja norteada de acordo com as características dos consumidores destas informações, com um canal de comunicação para retorno de suas sugestões, dúvidas e avaliações, estreitando o nível de interatividade destes com os gestores dos armazéns de dados.

### 1.1 Objetivo geral e objetivos específicos

Partindo da hipótese de que a produção de dados de saúde pelo Datasus precisa ser aprimorada, conforme as necessidades e experiências dos diversos tipos de consumidores da informação, e do fato de que a adoção de armazéns de dados (*Data marts/Data warehouses*) pode subsidiar um melhor atendimento das demandas, o **objetivo geral** deste estudo pode ser definido como: **promover a adoção da tecnologia de armazéns de dados (*data marts/data warehouses*) para difundir informações de saúde para as instâncias de decisão, conforme as necessidades e experiências dos consumidores da informação.**

Como objetivos **específicos**, podemos citar:

- a) Orientar a criação das bases de dados não segundo os sistemas de informação, mas segundo as grandes atividades do setor saúde demandantes de informações estruturadas: (promoção da saúde, gestão da saúde, regulação das ações de saúde, ações de vigilância em saúde)
- b) Orientar a classificação dos consumidores de informação, bem como suas necessidades de visualização dos dados agregados, e identificação dos níveis de segurança a serem respeitados.
- c) Definir um método de trabalho para a construção e carga das bases de dados.
- d) Definir as diversas formas de extração de dados de modo a atender às diferentes necessidades dos consumidores das informações contidas nestes repositórios de dados.

Como **não-escopo** deste projeto, podemos dizer que:

- a) Ele não se propõe a implementar o método de trabalho.
- b) Ele não se propõe a dimensionar os recursos tecnológicos e humanos de nenhuma das fases do método de trabalho.

## 1.2 Metodologia

Pela formulação do objeto deste trabalho - a elaboração de um método de trabalho para a construção de bases de dados, íntegras e padronizadas, destinadas à difusão de informações de saúde para suporte à decisão, considerando as necessidades e experiências dos consumidores destas informações – é possível identificar dois grandes grupos passíveis de serem beneficiados com este estudo.

O primeiro engloba os consumidores da informação, os quais, normalmente, não possuem especialização nos campos de conhecimento da tecnologia da informação, e ficam reféns das deliberações técnicas dos especialistas da área. Conforme nos mostra Branco<sup>18</sup>, “o deslumbramento com as opções oferecidas pela tecnologia fez com que, frequentemente, fosse relegada a preocupação com o conteúdo dos dados e suas limitações. Confundi-se ferramenta com produto. As formas de geração e acesso à informação foram tratadas como se fossem, em si, a informação”.

O segundo grupo está no extremo oposto desta balança, e é composto pelos profissionais de TI responsáveis por construir as bases de dados citadas no objeto deste trabalho. São aqueles que Branco<sup>18</sup> cita como técnicos e especialistas que tornaram-se intérpretes das necessidades dos usuários, os quais acabaram colocando-se incompetentes para participar do processo de geração da informação. São técnicos da informação e informática que sentiram-se ‘gênios’, quase deuses. Como ela diz, são os “senhores do novo mundo, com muito a ensinar e a prescrever”.

O presente trabalho se propõe, então, a trazer novas ideias, ou preencher lacunas em ambos os mundos. Ou seja, para o primeiro grupo, o dos consumidores da informação, ele pretende mostrar como é o processo de construção das bases de dados para suporte à decisão, sem, no entanto, deter-se demoradamente nos aspectos técnicos mais profundos, considerando a existência de vasta bibliografia especializada neste processo, e que, normalmente, é de conhecimento dos profissionais deste campo na área de TI.

Para o segundo grupo, o dos profissionais de TI, ele pretende sensibilizá-los para a importância da parceria com os consumidores da informação, durante todo o processo, de modo a garantir a eficácia, e não apenas a eficiência, de todo o seu esforço de construção.

Assim sendo, a metodologia utilizada neste trabalho será baseada em uma estratégia convencional, focada em pesquisas bibliográficas, que permitam visitar o que for mais relevante, nos dois aspectos considerados críticos, na construção destas bases de dados para apoio à decisão.

Na parte ligada ao relacionamento com o usuário, ou consumidor da informação, buscou-se uma literatura voltada para as áreas de sistemas de informação, de informação em saúde, de filosofia e de informação propriamente dita.

Optou-se, então, como literatura básica, pelos livros “Informação e saúde”, de Branco<sup>18</sup>; “Política, tecnologia e informação em saúde – a utopia da emancipação”, de Moraes<sup>7</sup>; “O sonho transdisciplinar”, de Japiassú<sup>19</sup>; e “Ecologia da informação”, de Davenport<sup>20</sup>; além de outros citados nas referências bibliográficas.

No que diz respeito a artigos, utilizamos “Política de saúde e potencialidades de uso das tecnologias de informação”, de Vasconcellos, Moraes e Cavalcante<sup>12</sup>.

Além deste, podem ser citados os artigos “Informação e informática em saúde: caleidoscópio contemporâneo da saúde”, de Moraes e Gómez<sup>21</sup>; “Gestão governamental e sociedade: informação, tecnologia e produção científica”, de Ribeiro, Sophia e

Grigório<sup>22</sup>; e “Para uma reflexão epistemológica acerca da Ciência da Informação”, de Gómez<sup>23</sup>.

No que se refere ao aspecto mais focado na tecnologia da informação, mais especificamente nas técnicas de construção dos armazéns de dados e nas técnicas de extração de dados para carga nas bases de dados, foram pesquisados os livros “*The data warehouse lifecycle toolkit*”<sup>14</sup> e “*The data warehouse toolkit – second edition*”<sup>24</sup>, de Kimball; e “*Data warehouse: conceitos, tecnologias, implementação e gerenciamento*”, de Singh<sup>13</sup>, por já oferecerem uma boa cobertura destes assuntos, sendo referências no tema. Além destes, um bom subsídio para o trabalho foi encontrado em “*BI – business intelligence: modelagem & tecnologia*”, de Barbieri<sup>16</sup>; e “*Business intelligence no Microsoft Excel*”, de Leme Filho<sup>15</sup>.

Acreditamos que esta literatura permite uma boa revisão sobre o assunto, considerando suas diferentes nuances.

### 1.3 Contextualização

Antes do detalhamento efetivo deste trabalho, cabe um esclarecimento em relação a alguns aspectos conceituais, e uma contextualização do que será abordado mais à frente.

#### *a) Informação em Saúde*

Segundo o objetivo do presente trabalho, de elaborar um método para a produção de dados de saúde, íntegros e padronizados, que subsidiem informações de saúde personalizadas, de acordo com as necessidades e experiências dos consumidores de informação, é mister entender prioritariamente o conceito de Informação em Saúde.

Segundo Moraes<sup>3</sup>, enquanto dado “é uma descrição limitada do real, desvinculada de um referencial explicativo e difícil de ser utilizada como informação por ser ininteligível”, informação “é uma descrição mais completa do real associada a um referencial explicativo sistemático. Pode-se dizer que é a representação de fatos da realidade com base em determinada visão de mundo, mediante regras de simbologia. É



portanto, a ‘ponte’ entre fatos da realidade ou as ideias de algumas pessoas e as ideias ou conhecimentos de outras.”

Assim sendo, os dados produzidos pelos sistemas de informações, a partir do processamento dos dados de entrada, coletados de forma manual (via formulários analógicos) ou eletrônica (via formulários digitais ou gerados a partir de transações eletrônicas) serão a fonte a partir da qual se produzirão as informações em saúde.

Ainda no mesmo livro, a autora nos esclarece que “na prática, fica muito difícil delimitar o campo de abrangência das informações em saúde. Adotando o conceito ampliado de saúde, também devem ser entendidas e consideradas como informações em saúde, aquelas acerca de alimentação, moradia, saneamento, etc., ou seja, acerca das condições gerais da vida e trabalho”<sup>3</sup>. O que nos leva a entender que não são apenas os sistemas de informações em saúde que produzem dados úteis para a geração de informações em saúde.

Embora não seja objeto central deste estudo, vale a pena mover o olhar deste aspecto tecnológico para uma rápida varredura das informações em saúde sob uma ótica mais abrangente.

Moraes<sup>7</sup> nos mostra que as informações relacionadas às doenças, às condições da morte e do nascimento dos indivíduos e das populações, que formam a essência das Informações em Saúde, coincidem, segundo Foucault<sup>8</sup>, com o nascimento da Clínica, e de um novo tipo de prática médica (fins do século XVIII e início do século XIX), que rompeu com a medicina clássica, que considerava a doença como portadora de uma essência abstrata, adotando uma medicina moderna, fundada no pensamento anátomo-clínica e nos princípios da patologia, em que a doença passa a ter sua significação corporificada no indivíduo, que é o portador do corpo doente.

Há, então, a constituição de saberes específicos, classificatórios, que formarão a base para a classificação de doenças e para a formação dos acervos estatísticos, passíveis de comparação no tempo e espaço, o que, num caminho inverso, alimentava o próprio desenvolvimento dos saberes médicos.

Ainda segundo a autora, a constituição destas informações em saúde faz com que a saúde passe a ingressar na agenda política das sociedades, passando a ser uma questão de Estado, demandando políticas específicas, refletindo a maneira como os conflitos e tensões dos diferentes interesses envolvidos se confrontam.

É o poder. Ou, como Foucault<sup>9</sup> coloca, não existe uma realidade global e unitária chamada poder. O que existem são práticas de poder exercidas de forma complexa, múltipla e estratégica por toda a estrutura social. Segundo Foucault<sup>9</sup>, “o poder (...) é o ícone que se dá a uma situação estratégica complexa numa determinada sociedade. (...) Aí onde há poder, há resistência e no entanto, ou até por isso mesmo, ela já não está em posição de exterioridade em relação ao poder”.

Devemos, então, observar não apenas o aspecto das informações em saúde ligados à tecnologia da informação e comunicação, e suas diferentes ramificações, mas também à importância social e política, aos conflitos de interesse, e às tensões de poder causadas pela posse e controle destas informações em saúde.

Cabe, neste momento, uma reflexão sobre um Estado informacionalmente democrático, em que a informação é direito de todos e dever do Estado. Ou seja, o Estado tem o dever de responder às demandas informacionais da sociedade e o cidadão tem o direito de perguntar e de participar do processo de gestão da informação, ao contrário da prática atual, em que o Estado tem o direito de perguntar aos indivíduos e o indivíduo tem o dever de responder.

#### *b) Tecnologia da Informação (TI)*

O segundo conceito a ser abordado é o de tecnologia da informação. Uma definição deste assunto pode ser obtida em Batista<sup>10</sup>, onde a tecnologia da informação (TI) é caracterizada como o conjunto de *hardware* (equipamentos e acessórios), *softwares* (programas, utilitários, etc) e *firmware* (circuitos integrados de alguns equipamentos que possuem programas internos para determinadas atividades, como um torno CNC ou mesmo um telefone celular).

É interessante observar, neste momento, o contexto de saúde, e mais, o contexto sócio-político no qual as tecnologias de informação se enquadram.

Segundo Moraes<sup>7</sup>, é preciso acumular forças políticas para que haja uma participação maior dos interesses populares em mecanismos de gestão social sobre a política informacional em saúde. É preciso avançar tecnologicamente dentro das novas perspectivas da área, que surgem em velocidade exponencial, mas deve-se democratizar os conteúdos informacionais úteis e relevantes. A importância das conquistas na área de informática e das comunicações são inegáveis, não somente para a área de saúde, mas para todas as áreas. Deve-se pensar numa estratégia que se aproxime, ainda segundo a

autora, “do que Levy<sup>11</sup> chama de tecnodemocracia, onde uma reaproximação mental dos processos tecnológicos por parte dos cidadãos é um pré-requisito para a instauração progressiva da democracia”.

### *c) Sistema de Informação para Apoio à Decisão em Saúde*

O terceiro conceito a ser visto é o de Sistema de Informação para Apoio à Decisão em Saúde.

Segundo Vasconcellos, Moraes e Cavalcante<sup>12</sup>, a proposta de um Ambiente de Informações para Apoio à Decisão em Saúde parte da hipótese de que “ampliar as potencialidades de uso das TI na gestão de Saúde implica (...) em uma nova concepção de organização das informações em saúde, onde sejam estruturados mecanismos e condições que criem um ambiente propício para estabelecimento de uma sinergia de competências, recursos e memórias, frutos e matrizes do conhecimento sanitário, coletivamente produzido por seus sujeitos históricos”.

“Este Ambiente de Informações tem por objetivo apoiar decisões referidas a ações/intervenções no campo da saúde, ao monitoramento e avaliação dessas ações, estabelecendo mecanismos para acesso a informações úteis e relevantes para o gestor de problemas de saúde a serem enfrentados no cotidiano institucional, no desenho de cenários futuros e análises prospectivas”.

“O produto desse Ambiente é um Sistema de Informação de Apoio à Decisão (...) constituído de uma infraestrutura com uso intensivo de tecnologias de informação, composta de ferramentas voltadas para a aquisição e integração de dados originados de Registros Administrativos, Pesquisas/Inquéritos e Dados Transacionais, articulados a métodos de modelagem estatística, modelagem matemática e à inteligência computacional, incorporando a perspectiva espacial e temporal em novos patamares”.

Ainda com relação aos sistemas de informação de apoio à decisão, convém apresentar algumas definições, pertinentes à área dentro do campo da tecnologia da informação que estuda estes tipos de sistemas, as quais serão úteis na compreensão de detalhes técnicos mais à frente.

#### d) *Data Warehouse*

Talvez o melhor conceito de *data warehouse*, cuja tradução literal é Armazém de Dados, esteja em Singh<sup>13</sup> : “*Data warehouse* é o processo de integração dos dados corporativos de uma empresa em um único repositório a partir do qual os usuários finais podem facilmente executar consultas, gerar relatórios e fazer análises. Um *data warehouse* é um ambiente de suporte à decisão que alavanca dados armazenados em diferentes fontes e os organiza e entrega aos tomadores de decisões da empresa, independente da plataforma que utilizam ou de seu nível de qualificação técnica. Resumindo, *data warehouse* é uma tecnologia de gestão e análise de dados”.

Alguns autores defendem a ideia de que *data warehouse* é o repositório de dados, o armazém de dados em si; e o seu processo de construção é chamado de *data warehousing*. Isto costuma causar confusão, especialmente por serem termos em língua inglesa, na tradução dos livros originais para a língua portuguesa.

De qualquer forma, pode-se extrair conceitos importantes a partir da definição acima, como o fato do *data warehouse* ser um repositório de dados corporativos integrados, obtidos de diversas fontes, independente da plataforma tecnológica utilizada, e cujo propósito é permitir que os usuários finais realizem consultas, extraiam relatórios e façam análises, tornando-o um importante instrumento de gestão e análise de dados.

#### e) *Data Mart*

Segundo Kimball<sup>14</sup>, um *data mart* é um “subconjunto lógico de um *data warehouse* completo”. E, “um *data mart* representa um projeto que pode ser completado ao invés de ser um projeto impossível de dimensões galácticas”.

Da definição acima podemos constatar que um *data mart* é um subconjunto do *data warehouse*, mas, funcionalmente, equivale a este, servindo aos mesmos propósitos. Isto significa que o processo de construção do *data warehouse* completo pode passar pela construção primeira de seus subconjuntos, ou *data marts*. Daí a explicação de

Kimball<sup>14</sup>, onde ele diz que o *data mart* é um projeto com possibilidades reais de término, e, por isso, a abordagem deve ser a da divisão do problema do *data warehouse* em problemas menores, os *data marts*, os quais, uma vez integrados, darão vida ao *data warehouse*.

#### f) *Data Mining*

Leme Filho<sup>15</sup> nos mostra que “*data mining* é um conjunto estruturado de processos de exploração e análise de grandes volumes de dados, por meio de métodos automáticos e semi-automáticos, com o intuito de descobrir e determinar padrões, tendências, relacionamentos e regras, que permitam converter tais dados em conhecimento de alto valor estratégico”.

#### g) *Ferramentas OLAP*

Singh<sup>13</sup> nos diz que OLAP significa (em língua portuguesa) Processamento Analítico On-Line, que descreve uma classe de tecnologias destinadas a acesso a dados e consultas não previstas ao vivo. Segundo ele, as “ferramentas OLAP fornecem meios tecnológicos para análise complexa do negócio, possibilitando aos usuários analisar e navegar os dados para detectar tendências, exceções e obter detalhes para entender melhor os altos e baixos das atividades de seu negócio”.

Barbieri<sup>16</sup> é feliz ao comparar as duas técnicas citadas acima, realçando as diferenças existentes entre ambas: “Enquanto as técnicas OLAP objetivam trabalhar os dados existentes, buscando consolidações em vários níveis, trabalhando fatos em dimensões variadas, a técnica de *mining* busca algo mais que a interpretação dos dados existentes. Visa fundamentalmente realizar inferências, tentando como que ‘adivinhar’ possíveis fatos e correlações não explicitadas nas montanhas de dados de um *data warehouse* / *data mart*”.

#### *h) Granularidade*

Segundo Leme Filho<sup>15</sup>, “granularidade é uma técnica que determina o nível de detalhe das informações disponíveis para consulta. Quanto mais detalhe for possível obter, menor é a granularidade. Quanto menos detalhe existir, maior será a granularidade. (...) Uma das razões que tornam a granularidade uma das principais questões de projetos de *data warehouse* reside no fato de que esta afeta diretamente o volume de dados que serão armazenados e, ao mesmo tempo, nos tipos de consulta que se pode obter. O volume de dados é diretamente proporcional ao nível de detalhe de uma consulta”.

A granularidade também influi no nível de agregamento possível no *data warehouse*. Por exemplo, se definirmos que o nível de granularidade irá descer até os dados do paciente individual (como no caso do prontuário do paciente), podemos facilmente fazer totalizações “subindo” ao nível do município, da unidade da federação e ao nível federal. Porém, se definirmos que o nível de granularidade irá no máximo até os dados do município, só poderemos agregar valores daí para cima; não nos será possível obter nenhuma informação em níveis mais detalhados que do município.

#### *i) Modelo de Dados*

De acordo com Singh<sup>13</sup>, um modelo de dados é “uma representação gráfica dos dados de uma área de interesse específica. Essa área de interesse pode ser tão abrangente quanto todos os requisitos de dados integrados da empresa (modelo de dados da empresa) ou uma única área do negócio ou aplicação (modelo de dados de um assunto). (...) O objetivo do modelo de dados é transmitir claramente o significado dos dados, os relacionamentos entre eles, os atributos e registrar as definições precisas desses dados”.

Nos *data warehouses/data marts* normalmente se adotam os modelos de dados dimensionais, em que se identificam os fatos (uma internação de paciente, um óbito, um nascimento são exemplos de fatos), e também se identificam as dimensões pelas quais

se quer visualizar estes fatos (dimensões temporais, dimensões geográficas, dimensões de faixa etária, sexo, escolaridade, etc).

#### *j) Software Livre*

Saindo dos conceitos de suporte à decisão, entramos em outro que vale a pena ser citado, que é o conceito de *software* livre.

A melhor definição para *software* livre pode ser encontrada no *site* da Fundação para *Software* Livre (*Free Software Foundation*)<sup>17</sup>, mas podemos resumi-la aqui:

“*Software* livre se refere à liberdade dos usuários executarem, copiarem, distribuírem, estudarem, modificarem e aperfeiçoarem o *software*. Mais precisamente, ele se refere a quatro tipos de liberdade, para os usuários de *software*:

- A liberdade de executar o programa, para qualquer propósito (liberdade nº 0).
- A liberdade de estudar como o programa funciona, e adaptá-lo para as suas necessidades (liberdade nº1). Acesso ao código-fonte é um pré-requisito para esta liberdade.
- A liberdade de redistribuir cópias de modo que você possa ajudar ao seu próximo (liberdade nº 2).
- A liberdade de aperfeiçoar o programa, e liberar os seus aperfeiçoamentos, de modo que toda a comunidade se beneficie (liberdade nº 3). Acesso ao código-fonte é um pré-requisito para esta liberdade.

Um programa é *software* livre se os usuários têm todas essas liberdades”.

## 2. Informação, saúde e conhecimento

### 2.1 A informação

Analisando o título deste trabalho, “Uso de informações de saúde para suporte à decisão: uma metodologia focada no consumidor da informação”, pode-se ver por duas vezes a ocorrência da palavra “informação”. Embora tenha sido já contextualizada na introdução desta dissertação, vale a pena abordar um pouco melhor o campo de conhecimento da ciência da informação.

Uma variação da comparação “dado *versus* informação”, descrita na contextualização deste trabalho, pode ser obtida em Branco<sup>18</sup>:

“Entende-se por DADO um valor qualitativo ou quantitativo obtido para caracterizar um fato ou circunstância. O termo INFORMAÇÃO é utilizado no sentido de conhecimento obtido a partir dos dados e implica em interpretação por parte do usuário.”

Ainda segundo a autora, a informação pode ser definida segundo três aspectos:

- a) Como objeto físico, no que se refere aos documentos, papéis, etc, que tenham potencial informativo.
- b) Como conhecimento, no que diz respeito a algo intangível, a algo que é comunicado sobre determinado fato, tema ou evento, permitindo agregar algum ganho de conhecimento.
- c) Como processo, no que toca ao ato de informar e ser informado, de produzir informação como forma de alterar o estado de conhecimento.

Estas três ramificações são tão importantes que, juntas, formam o corpo central do estudo da disciplina que hoje se chama Gestão do Conhecimento.

Os três olhares sobre a informação também nos permitem observar o caráter transformador da informação e sua importância para quem a recebe, quando se diz que ela tem potencial informativo, que permite agregar algum ganho de conhecimento, e que é uma forma de alterar o estado do conhecimento.



Ou seja, a finalidade principal da informação é produzir conhecimento, em nada sendo útil se não relacionada a quem a produziu ou a quem é direcionada, e que terá seu estado de conhecimento modificado por ela.

Como diz Bonsack<sup>25</sup>, citado em Branco<sup>18</sup>,

“Uma informação, em si, não é nada, não tem nada de informação se ninguém é informado ou se a informação recebida não é utilizável para aquele que a recebeu”.

## 2.2 Ciência da informação e informação em saúde

Segundo Gómez<sup>23</sup>, “a constituição de um campo científico da ciência da informação sempre foi uma questão em aberto. Difícil, para muitos; não relevante, para outros; desafiante, para alguns.”. Isto, talvez, por causa da característica estratificada, difusa, da informação.

Estas várias dimensões do objeto informacional, a amplitude de sua abrangência, talvez justifique a orientação interdisciplinar e transdisciplinar deste campo do conhecimento, assim como as diversas tentativas de “paternidade” de sua conceituação e definição.

Como nos mostra a autora, espaços discursivos e institucionais da epistemologia são cada vez mais reduzidos, com pequeno avanço, sendo preteridos pelo “incremento em escala geométrica dos recursos e atividades dedicados à pesquisa e desenvolvimento de tecnologias de informação”.

É fácil entender esta diferença de tratamento quando se pensa, principalmente, no retorno financeiro propiciado por este esforço. No entanto, a “tecnologia pela tecnologia” pode trazer sérias consequências em várias áreas, especialmente, como será visto, no campo da informação em saúde pública e nos sistemas de informação afetados diretamente por esta corrida tecnológica.

Só para se ter noção de como o campo da ciência da informação é muito novo, vale transcrever um trecho do livro de Branco<sup>18</sup>, em que McGarry<sup>26</sup> compara o tempo da evolução do homem na Terra com um período de 24 horas:

“Tomando como ponto inicial arbitrário trinta mil anos atrás, quando o homem começava sua fase de pinturas e esculturas nas

cavernas, tem-se o início do período de 24 horas, a partir da meia-noite. Cada hora do relógio corresponde a 1200 anos e, conseqüentemente, cada cinco minutos a cem anos. Por milhares de anos, o progresso se deu numa velocidade glacial. Todo o período paleolítico transcorreu entre dez horas da manhã e duas horas da tarde. Às 20h, surgiu a escrita cuneiforme nos vales da Mesopotâmia e os hieróglifos nos templos egípcios. Às 22h, tem início a civilização grega, com sua contribuição para a arte, a filosofia e a matemática. Às 22h30min, surgem os primórdios da cultura judaico-cristã na Europa. Às 23h33min, inventa-se a tipografia, que impulsiona o conhecimento científico. Às 23h48min, começa a Revolução Industrial na Inglaterra. A revolução das comunicações tem lugar nos últimos cinco minutos. O computador tem atravessado gerações no último minuto e, no caso do microcomputador, estamos falando dos últimos segundos das 24 horas decorridas”.

### 2.3 Interdisciplinaridade e transdisciplinaridade

A amplitude da abrangência do objeto informacional, citada acima, reflete-se diretamente no tema deste trabalho, quando fala de difusão de informações de saúde para suporte à decisão; ou, como descrevem Vasconcellos, Moraes e Cavalcante<sup>12</sup>, na construção de um Ambiente de Informações para Apoio à Decisão em Saúde.

Segundo os autores, a abordagem metodológica adotada para a estruturação deste ambiente está “assentada na busca de um aprofundamento da circulação inter/transdisciplinar do conhecimento entre campos de saberes”.

Cabe aqui visitar os conceitos de interdisciplinar e transdisciplinar.

Em primeiro lugar, segundo Japiassú<sup>19</sup>, vem o conceito de disciplina:

“Trata-se de um conjunto específico de conhecimentos possuindo suas características próprias no campo do ensino, da formação, dos métodos, dos mecanismos e dos materiais, numa palavra, do monodisciplinar, no saber científico, ela institui a divisão e a

especialização do trabalho; suas fronteiras, sua linguagem e seus conceitos próprios tendem a isolá-la das demais disciplinas”.

A partir dessa conceituação, define-se multidisciplinaridade, interdisciplinaridade e transdisciplinaridade. Segundo Naomar de Almeida Filho<sup>27</sup>,

“Multidisciplinaridade – conjunto de disciplinas que simultaneamente tratam de uma dada questão, problema ou assunto, sem que os profissionais implicados estabeleçam entre si efetivas relações no campo técnico ou científico; é um sistema que funciona através da justaposição de disciplinas em um único nível, estando ausente uma cooperação sistemática entre os diversos campos disciplinares;

Interdisciplinaridade – para além da disciplinaridade, a noção de interdisciplinaridade implica uma axiomática comum a um grupo de disciplinas científicas conexas, cujas relações são definidas a partir de um nível hierárquico superior, ocupado por uma delas; esta última, geralmente determinada por referência à sua proximidade a uma temática unificada, atua não somente como integradora e mediadora da circulação dos discursos disciplinares, mas principalmente como coordenadora do campo disciplinar;

Transdisciplinaridade – indica uma integração das disciplinas de um campo particular sobre a base de uma axiomática geral compartilhada;...sua coordenação é assegurada por referência a uma base de conhecimento comum, com tendência à horizontalização das relações interdisciplinares.”

Segundo o autor, ecologia, energia, alimentação, demografia, comunicação internacional são exemplos de problemas que, exigindo um percurso através de várias disciplinas, constituem exemplos de transdisciplinaridade. Diz ainda que, na ausência deste espírito transdisciplinar, há que se contentar com aproximações multidisciplinares por vezes conduzindo a colaborações interdisciplinares, com a utilização, numa disciplina, de um resultado ou método de outra disciplina. Segundo ele, a interdisciplinaridade desloca o centro em direção às fronteiras, onde está o problema; mas o transdisciplinar instaura uma visão globalizante: neutraliza a oposição centro-fronteira.

Portanto, conclui-se que é errado acreditar que a especialização em Tecnologia da Informação, nas mais modernas técnicas e ferramentas para construção de sistemas de informação para suporte à decisão, possa, por si só, satisfazer às necessidades de informação em saúde, contando com o “apoio” dos profissionais dos demais campos do conhecimento envolvidos na estruturação deste ambiente.

Aliás, segundo Japiassú<sup>19</sup>,

“O modo de pensamento ou de conhecimento fragmentado, monodisciplinar e simplesmente quantificador, tomando como critério de construção o ponto de vista (o paradigma) de um ramo do saber autodeterminado ou disciplina, com todos os seus interesses subjacentes, é responsável pela prevalência de uma inteligência bastante míope ou cega na medida em que é sacrificada a aptidão humana normal de religar os conhecimentos em proveito da capacidade (também normal) de separar ou desconectar.”

Pela experiência do próprio autor, este é um modo de pensar típico dos profissionais de informática, e, a situação pode se tornar mais preocupante quando estes profissionais não têm um comprometimento direto com a área de Saúde, e, em nome de uma suposta qualidade aliada à rapidez na execução dos projetos, o que muitas vezes é interessante sob o ponto de vista político, apresentam soluções “fabricadas”, “instantâneas”, mas que estão longe da real necessidade informacional da área de Saúde.

Mas, por trás da busca pelo lucro rápido, existe uma questão conceitual que envolve o próprio processo de especialização, aliado às disputas de poder causadas por este “esfacelamento do saber”, como diz Japiassú.

Segundo o autor,

“A especialização sem limites culminou numa fragmentação crescente do horizonte epistemológico. Chegamos a um ponto em que o especialista se reduz ao indivíduo que, à custa de saber cada vez mais sobre cada vez menos, termina por saber tudo (ou quase tudo) sobre o nada, em reação ao generalista que sabe quase nada sobre tudo.”<sup>19</sup>

Ainda segundo ele,

“Até Kant, praticamente todos os filósofos dominavam a física. Tratava-se de ver um problemas sob todos os seus ângulos e em

todas as suas dimensões. A especialização *stricto sensu* nasce apenas no século XIX da aceleração galopante dos conhecimentos e da sofisticação crescente das novas tecnologias. Na segunda metade do XX, surge e rapidamente se impõe a hiperespecialização, provocando a multiplicação indefinida de disciplinas e subdisciplinas cada vez mais focadas em reduzidos objetos de estudo.”<sup>19</sup>

Assim, o território do saber é retalhado e dividido, pelo contínuo desenvolvimento da especialização, objetivando ocupá-lo, dominá-lo e conhecê-lo, como forma de sobre ele exercer um poder. As disciplinas, fechadas e estanques, viram fonte de ciúme, glória, poder e atitudes dogmáticas.

Moraes<sup>7</sup> nos fala, como será visto no próximo capítulo, dos interesses e das “disputas pela manutenção dos micropoderes presentes nos feudos técnicos, que lutam pela preservação de seus espaços de atuação, mantendo-os estanques, com receio de que o compartilhamento represente perda de autonomia e liberdade de ação”.

Segundo Almeida Filho<sup>27</sup>, o próprio paradigma cartesiano, que decompunha a coisa a conhecer, reduzindo-a às suas partes mais simples, valorizava a especialização dentro do processo de produção do capitalismo industrial. “Então, o princípio da parcimônia, no sentido da simplificação reducionista, validaria os modelos explicativos do novo modo de produção de conhecimento – pois o conhecer reduzia o agora objeto aos seus componentes elementares”.

Então, que campos de saberes são estes aos quais se referem Vasconcellos, Moraes e Cavalcante? Eles próprios destacam as áreas de Tecnologias de Informação, Gestão em Saúde, Planejamento e Epidemiologia.

Ainda segundo os autores, “um dos desafios desta proposta tem sido o desenvolvimento desta metodologia integradora, incorporando métodos e técnicas de diferentes disciplinas que se complementam”

É preciso, para isso, como diz Japiassú, substituir “a concepção do poder mesquinho e ciumento do especialista pela concepção de um *poder partilhado* e socializado”.

Ainda segundo o autor, “o homem inteligente opõe a compreensão das situações complexas; sabe que o todo não é obtido pela soma das partes; sabe articular, convergir e concertar, pois compreende que os elementos do saber só possuem sentido pelas relações mútuas que mantêm entre si”.

#### 2.4 Inteligência competitiva e o metaconhecedor

Com esta característica difusa, de amplo espectro, a informação, como visto, mesmo assim, é responsável por agregar conhecimento, e é necessário organizar este conhecimento.

Segundo Gómez<sup>23</sup>,

É no cenário contemporâneo que o conhecimento, numa gama ampliada que inclui a ciência e a tecnologia junto a outros conhecimentos políticos e econômicos, considerados estratégicos, passa a ser objeto preferencial de estudos econômicos, políticos e administrativos, reunidos sob as denominações de *knowledge management* (inteligência organizacional), *inteligência competitiva*, dentre outras.

A construção de bases informacionais, como *data warehouses*, constituem um acervo que permite subsidiar a construção de um processo de gestão do conhecimento na própria área da saúde.

O que permitiu este avanço na área de gestão do conhecimento foram as novas estratégias, emergentes, segundo Gómez<sup>23</sup>, na década de 90, “da ordem da construção, da modelagem e da intervenção, e que se orientam à formulação de modelos e algoritmos (conheço aquilo acerca do qual conheço o algoritmo de sua produção)”.

Segundo a autora, nesta perspectiva, se fala do conhecimento de duas formas contrapostas:

- a) A primeira é o conhecimento que é “visto como objeto de um observador, que pode conhecê-lo e observá-lo com independência de seu produtor, sendo então o conhecimento da ordem do objetivado e objetivo”. Não é o tipo de conhecimento que interesse à construção de um sistema de suporte à decisão, mas é o conhecimento obtido por quem vai utilizar seus produtos.
- b) Na segunda, “o conhecimento é considerado como a construção de um metaconhecedor intencional e ativo que pode objetivar, modelar, gerir um processo de conhecimento de outros conhecedores, assim como intervir sobre o observado e a observação”.

É este papel de metaconhecedor que poderia caber ao profissional de Tecnologia de Informação, ao “costurar”, ao modelar e objetivar o processo de conhecimento de outros conhecedores na elaboração de um ambiente de suporte à decisão, mas numa posição hierarquicamente igual aos demais, na característica transdisciplinar da horizontalização das relações interdisciplinares.

É interessante observar que este papel de metaconhecedor do profissional de Tecnologia da Informação, se por um lado o tira da posição centralizadora de tomada de decisões sobre todo o processo de construção dos sistemas de informação, por outro lado o livra das exigências inerentes a esta centralização, especialmente sobre áreas do conhecimento que não são sua especialidade.

Para isso, todos terão que fazer uso do espírito inter- ou transdisciplinar, o qual, segundo Japiassú<sup>19</sup>, “coloca à nossa disposição a fascinante capacidade do espírito humano de congregar, concertar, compreender, inventar e elaborar projetos suscetíveis de nos permitir uma melhor compreensão da realidade e de nossas ações”.

### 3. Os sistemas de informação

#### 3.1 Informações em saúde e os sistemas de informação

Na introdução deste trabalho foi mostrado o conceito de sistema de informação, bem como foi contextualizado o conceito de informação em saúde, mostrando como a constituição das informações relacionadas às doenças, às condições da morte e nascimento dos indivíduos e das populações, que formam a essência das Informações em Saúde coincidiram com o nascimento da Clínica e de um novo tipo de prática médica.

Foi mostrado, também, como a constituição destas informações em saúde faz com que a saúde passe a ser uma questão de Estado, e uma fonte de poder.

Porém, o Estado não é, segundo Moraes<sup>7</sup>, o órgão central e único de poder, que estende sua rede de poderes sobre a sociedade, mas existem outras fontes de poder para além dos aparelhos do Estado.

Na verdade, de acordo com a autora, existem relações de poder, que vão desde as decisões de políticas de dados até as relações interpessoais e locais, que circulam o poder pelas veias da sociedade, e inserem-se em instituições como as da saúde, com técnicas de dominação mais ou menos sutis.

E estas relações de poder atuam de forma direta sobre a produção dos saberes ligados ao processo informacional em saúde.

Segundo a autora<sup>7</sup>,

“O processo de produção e disseminação de informação pressupõe decisões políticas e econômicas eivadas de relações de poder e produção de saber. Decisões que, em sua maioria, são tomadas pelos “produtores de informação”, sem “ouvir” as reais demandas da sociedade. Decisões que trazem consequências para o próprio processo de democracia de um povo, na medida que incidem na produção do conhecimento: científico e tecnológico e/ou social, político, econômico, ideológico, cultural, ou seja, na consciência do mundo que cerca os indivíduos.”



O fato de que estes sistemas de informação, e as informações em saúde que eles produzem, ainda que utilizados para proveito próprio, ou como diz Moraes, “utilizada pela elite política, econômica e tecnológica, de forma a instrumentalizar, escamotear e/ou justificar decisões ou não-decisões voltadas para o avanço ou a manutenção dos interesses hegemônicos em uma dada conjuntura”, não deve esquecer de seus princípios básicos, “de que a informação em saúde deve ser trabalhada no sentido de reforçar os direitos humanos, de contribuir para a eliminação da miséria e das desigualdades sociais e ao mesmo tempo subsidiar o processo decisório na área da saúde, em prol de uma atenção com efetividade, qualidade e respeito à singularidade de cada indivíduo e ao contexto de cada população”.<sup>7</sup>

Mesmo que os dados contidos em um *data warehouse* ou *data mart*, cuja construção será tratada no capítulo 6, estejam em um nível de suporte à decisão, e, portanto, longe da “linha de fogo” do dia-a-dia dos sistemas de informação de produção, ligados às interações do cidadão com o SUS em suas várias formas (internações, atendimentos ambulatoriais, vacinações, etc), é bom termos em mente que são os dados produzidos por estes últimos, conforme visto na introdução desta dissertação, que irão alimentar os primeiros. Deve-se considerar que a razão da existência de ambos é o cidadão, e que estes sistemas de informação devem ser olhados da “perspectiva dos pobres, da população que se vê reduzida a números sem rostos, que dizem muito pouco de seus sofrimentos singulares, de sua bravura cotidiana ao enfrentar as filas para serem atendidos pelo aparato de saúde – que não consegue atender com equidade ao dever constitucional de uma oferta de serviços de saúde compatíveis com as necessidades da população.”<sup>7</sup>

### 3.2 Problemas: integridade, fragmentação, integração

Já foi analisada anteriormente a importância da precisão e da acurácia dos dados produzidos pelos sistemas de informação, que levarão as informações que irão subsidiar as ações e responder às necessidades dos gestores e demais órgãos do SUS.

Do ponto de vista dos sistemas de apoio à decisão, mais especificamente dos bancos de dados para suporte à decisão que serão construídos, o grande problema dos sistemas de

informação são os dados que eles produzem, e que serão as fontes de alimentação dos bancos de dados para suporte à decisão.

Os problemas principais envolvendo os dados produzidos pelos sistemas de informação são a sua integridade, a sua integração e a sua padronização.

O problema da integridade ocorre, normalmente, de forma isolada. Ou seja, a sua causa está associada ao sistema de informação que o produz. Não necessariamente ao sistema propriamente dito, enquanto conjunto de programas computacionais, embora estes possam, por uma falha na sua construção, causar a produção de dados incorretos ou incoerentes na sua comparação. Mas essas falhas de integridade podem ser causadas por preenchimento incorreto de formulários, outras irregularidades na entrada de dados, ou mesmo corrupção dos bancos de dados que os armazenam.

Muitas vezes essas falhas são de fácil detecção e correção, através de ferramentas que vão desde a manutenção corretiva do sistema de informação em questão à mudança nos processos operacionais relacionados ao sistema ou mesmo à capacitação dos usuários finais.

Os problemas de integração e padronização são mais complicados e estão intimamente relacionados.

A integração de dados é tema de um estudo separado, tamanha a sua complexidade, mas pode ser resumida neste trabalho como a forma com que os dados de um sistema de informação específico se relaciona com os dados produzidos por outro sistema, levando à integração dos próprios sistemas de informação.

Moraes<sup>7</sup> nos mostra que um alicerce importante para a paulatina organização de bases informacionais em saúde é o “dossiê individual”, ou prontuário clínico. Segundo a autora, as bases genealógicas dos atuais sistemas de informação em saúde situam-se na “ligação definitiva da informação em saúde com a Clínica”, e “pode-se, talvez, considerar como uma das bases de sustentação instrumental desse biopoder o que é chamado como Prontuário do Paciente, ou, com os avanços tecnológicos contemporâneos, como Prontuário Eletrônico do Paciente”.

O esforço de integração de dados em bases distintas passa necessariamente pela questão da padronização de dados. Ou seja, a padronização dos dados produzidos por várias fontes é uma etapa fundamental para a integração dessas fontes. É comum encontrarmos bases de dados com codificações completamente diferentes para os mesmos dados. Assim, encontramos o atributo sexo sendo codificado de forma alfabética ('M', 'F', 'I' –

masculino, feminino e ignorado, respectivamente) em alguns bancos e os mesmos atributos sendo representados por 0, 1 e 2 em outros. O mesmo acontece com a representação de faixas etárias, escolaridade, estado civil, entre outros.

Embora não seja objeto deste estudo, vale citar aqui os esforços de movimentos objetivando a padronização e integração dos dados, como os Comitês de Compatibilização de Base de Dados e o de Padronização de Registros Clínicos, ambos Comitês Temáticos Interdisciplinares da Oficina de Trabalho Interagencial da Rede Interagencial de Informações para a Saúde (RIPSA). A RIPSA é uma iniciativa da Organização Pan Americana de Saúde, cujo propósito é a disponibilização adequada, oportuna e abrangente de dados básicos, indicadores e análises de situação sobre as condições de saúde e suas tendências no país. Vale citar também a iniciativa de Padronização de Sistemas de Informação na Área de Saúde, iniciativa do setor privado, envolvendo a Associação Brasileira de Medicina de Grupo, Federação Brasileira de Hospitais, Associação Brasileira dos Serviços Assistenciais de Saúde Próprios de Empresas, Comitê de Integração de Entidades Fechadas de Assistência à Saúde, Federação Nacional das Empresas de Seguros Privados e de Capitalização e Confederação das Misericórdias do Brasil<sup>12</sup>.

São importantes, também, os esforços de padronização da ABNT.

Como bem nos mostra Moraes e Gómez<sup>21</sup>, a gestão de saúde é inserida no modo como o Estado brasileiro implementa suas políticas sociais, ou seja, de forma tópica, atomizadora, fragmentadora em suas respostas às demandas sociais e reativa a situações de crise. (...) Daí resulta um modo de tomada de decisão política que fragmenta e retraduz a interdependência dos problemas sociais.

(...) As informações em saúde constituem um dos dispositivos desse Estado fragmentado(...). Essa análise reforça a hipótese que a fragmentação existente nas informações em saúde é expressão da fragmentação e da forma do Estado brasileiro atuar.

Esta fragmentação se expressa nas soluções “focadas no problema” dos sistemas de informações, que procuram resolver a situação específica de um determinado usuário, muitas vezes atrelada a prazos quase impraticáveis para entrega do produto final.

Esta atitude pode ser observada no próprio Departamento de Informática do SUS, da Secretaria Executiva do Ministério da Saúde. Segundo Branco<sup>18</sup>,

“...propunha-se que distintas bases de dados de interesse para a saúde, provenientes de fontes intra e extra-setoriais, fossem compatibilizadas, de forma que suas variáveis pudessem estar disponíveis num mesmo sistema. Não só os fatores que inviabilizavam essa integração, aqui já discutidos, se mantiveram, como o próprio Datasus, talvez até mesmo por esses fatores, atuou de forma oposta ao que estava propondo. Responsável pelo desenvolvimento de novos sistemas e pela manutenção dos já existentes, terminou por atender à lógica das ‘necessidades específicas’, desenvolvendo *softwares* cujas bases de dados continham variáveis afins, entretanto, em sua maioria, sem qualquer compatibilidade entre si.”

Outro agravante para este problema, também citado por Moraes e Gómez<sup>21</sup>, é o processo de sucateamento das instâncias públicas de gestão da informação e informática em saúde, com a terceirização, enfatizada pela forte pressão das empresas privadas de informática na disputa pela fatia de mercado do setor público da saúde. Esta visão da solução “focada no problema”, ainda mais forte na terceirização, foge do ideal de integração entre os conhecedores dos problemas e necessidades na Saúde Pública e os possuidores do “saber tecnológico” na busca de soluções para produção de informações em saúde.

A pluralidade de sistemas de informação oriundos de diversas fontes acaba produzindo efeitos colaterais. De acordo com Branco<sup>18</sup>,

“Não é pouco frequente encontrar profissionais e gestores que desconhecem o acervo de dados proporcionado por esses sistemas, que não sabem como acessá-los, ou não entendem o que significam certos termos e variáveis utilizados e mesmo o conteúdo de alguns relatórios.”

Uma maneira de tentar resolver esta questão da padronização e da integração, ao mesmo tempo em que tira dos ombros do desenvolvedor a tarefa de buscar a fonte gestora de um determinado dado, sua representação, e mesmo a disponibilização da definição do dado para outros desenvolvedores, é a criação de uma área de Administração de Dados, para centralizar e disponibilizar definições de dados, identificando gestores destes dados e dirimindo conflitos que possam surgir entre as áreas de desenvolvimento.

Segundo a experiência do próprio autor, e de entrevistas realizadas com outros administradores de dados de outras empresas do setor público e privado, quando da implantação desta área na então Coordenação Geral de Informações de Saúde do DATASUS, embora esta iniciativa seja buscada tanto na área pública quanto na área privada, ela é muito difícil de implementar. Fatores políticos e de disputa de poder tornam quase impossível a intervenção em sistemas já existentes, e, a menos que haja um efetivo patrocínio por parte da alta administração da instituição, qualquer tentativa de padronização e integração fica atrelada à boa vontade dos desenvolvedores.

Porém, pelo levantamento feito junto a estas outras empresas, ficou evidenciada a opção delas pelas tecnologias de armazéns de dados (*data warehouse/data mart*), que, além de propiciarem bancos de dados para apoio à decisão, conseguiam, também, como benefício extra, a integração de dados e a padronização, por serem pré-requisitos para a construção destes armazéns de dados, e por eles não estarem lidando com os próprios bancos de produção dos sistemas de informação, comumente chamados de bancos de dados transacionais.

Estas considerações, aliadas à experiência do autor com a elaboração de *data marts* para a Secretaria de Atenção à Saúde, levaram à escolha desta solução tecnológica para inclusão no método de trabalho para a construção de um ambiente de informações para apoio à decisão em saúde.

Tendo caracterizado todos os problemas relativos à integridade, à fragmentação e à integração dos dados nos sistemas de informação em saúde, vale a pena, por outro lado, registrar a “pouca idade” destes sistemas de informação.

O próprio uso da informática na área de saúde é relativamente recente. Segundo Branco<sup>18</sup>, em 1971 foi criado o Núcleo de Informática, vinculado à Assessoria de Coordenação Geral da Secretaria Geral do Ministério da Saúde, tendo por incumbência “manter atualizado um sistema de dados estatísticos, relatórios e informações outras, necessárias aos estudos e trabalhos do Ministério”. Porém, este núcleo só iniciou suas atividades em 1975, como Núcleo de Informática da Secretaria de Planejamento, da Secretaria Geral do Ministério da Saúde, com a incumbência de “coordenar o planejamento e desenvolvimento do Sistema de Informação de Saúde – SIS”

Após a criação do Sistema Único de Saúde, o próprio Departamento de Informática do SUS – DATASUS foi criado apenas em 1991, a quem cabia desenvolver, implantar e operar sistemas de informações relativos às atividades finalísticas do SUS.

Existem outros eventos muito recentes, se comparados com outras tecnologias de informação e comunicação, como é o caso da própria explosão comercial da *Internet*, que colocou vários *sites* no ar, e que foi acontecendo gradativamente, mas de forma acelerada, nos primeiros anos da década de 90.

O próprio *site* do DATASUS entrou no ar em 1995, com a disponibilização de informações em saúde através de um produto de tabulação desenvolvido no DATASUS, o TABNET.

Somente a partir daí é que o desenvolvimento dos sistemas de informação e seu uso foram amadurecendo, bem como foi amadurecendo a relação com os usuários.

No que diz respeito ao DATASUS, várias mudanças de direção a partir do final dos anos 90 trouxeram diversas orientações políticas e tecnológicas, causando uma falta de padronização dos sistemas e uma grande diferença no perfil tecnológicos dos mesmos.

Quanto aos usuários, principalmente após o processo de descentralização, fica fácil imaginar o quanto ficaram “refêns” dos donos da tecnologia.

Como nos mostra Alice Branco<sup>18</sup>,

“Onde a lógica do deslumbramento predominou, os sistemas de informação tornaram-se um *mix* de *know-how* tecnológico e saber técnico que intimidou sujeitos, tornando-os cúmplices do papel de ‘usuários’ a eles destinado.

Como ser sujeito de um processo que técnicos e tecnologia mostravam estar tão claramente distante para ‘pobres mortais’? A incorporação dessa lógica, pelos ‘usuários’, revelou-se, por exemplo, por meio das autodesignações anedóticas do tipo ‘sou daqueles usuários que acham que computador morde’ ou ‘esse universo todo da informática é um mistério para mim’.”

A autora nos mostra, também, que o passar do tempo fez crescer a insatisfação entre os usuários. A hegemonia tecnicista, com sua ‘interpretação das necessidades’, criou sistemas que eram abandonados pelos potenciais usuários depois de implantados.

#### 4. O usuário como consumidor da informação

Conforme visto na introdução deste trabalho, o termo “consumidor da informação”, cunhado por Bernd Frohmann, relacionado ao conceito de “regime de informação”, define melhor o papel do usuário em relação às informações em saúde.

Como mostrado, também, o rompimento com a hegemonia “tecnicista” para implantação de uma metodologia integradora, que incorpore métodos e técnicas de diferentes disciplinas que se complementam, dos diversos campos de saberes que irão construir um ambiente de suporte à decisão, é a melhor forma de trabalhar a questão das informações em saúde (e, porque não, dos próprios sistemas de informações).

Conhecer o usuário, ou consumidor da informação, é passo fundamental para o sucesso desta empreitada, e por isso, é tema de discussão como uma das etapas na construção do método de trabalho a que se propõe esta dissertação.

Segundo Branco<sup>18</sup>,

De acordo com a abordagem técnica e tecnológica, a concretização do potencial informativo é tanto mais favorecida quanto mais se procura conhecer as características, preferências, interesses e necessidades dos prováveis interessados. Portanto, os que se encontram envolvidos no processo de produção de informação devem otimizar os elementos que podem ter repercussão positiva na atribuição de relevância, tais como: atualidade da informação, oportunidade, confiabilidade, qualidade, custo e acessibilidade (Horne<sup>28</sup>, McGarry<sup>29</sup>, Menou<sup>30</sup>, Platt<sup>31</sup>, Saracevic<sup>32</sup>, Stroetmann<sup>33</sup>, Taylor<sup>34</sup>).

Estes elementos (atualidade, oportunidade, confiabilidade, qualidade, custo e acessibilidade) estão ligados diretamente ao conceito de utilidade para o usuário, o qual influencia a atribuição de significado, a relevância e o valor que o usuário dá aos dados.

A acessibilidade será discutida no capítulo 7 desta dissertação, mas vale a pena fazer algumas considerações sobre os outros, especialmente tendo em mente a realidade de disseminação de informações hoje existente.

Assim, a atualidade e a oportunidade são especialmente importantes, pois estão relacionadas à capacidade da informação de reduzir incertezas e de atender à finalidade

a que se propõe, de agregar conhecimento e mudar o estado do conhecimento. Como diz a autora, falhas na oportunidade e na atualidade depreciam o valor da informação.

Estes dois conceitos são interessantes de serem avaliados, especialmente se considerarmos a produção da informação nos três diferentes níveis de governo.

Assim, um dado é produzido no município e está pronto para uso com uma velocidade, em geral, muito maior do que um dado produzido no nível federal.

Este ambiente de suporte à decisão deve estar preparado para isto, ou seja, em disponibilizar o dado municipal à medida que for tendo acesso a ele, ao invés de esperar que ele tenha que “subir” ao nível federal para ser processado junto com os outros estados e municípios, para só então disponibilizá-lo nos bancos de dados.

Ter que aguardar por um dado que já está disponível em mãos, faz com que cresça no município o sentimento de descrédito em relação a estes bancos de dados para suporte à decisão, ou à instituição, ou instituições, incumbidas de disponibilizá-los.

Esta diferença de velocidade na produção dos dados “nas pontas” deve ser tratada com cuidado, ainda que certas comparações e avaliações intermunicipais ou interestaduais fiquem prejudicadas por um tempo, necessário para que todos os municípios ou estados tenham processados os mesmos dados.

É a questão da oportunidade, que faz com que o município passe a usar diretamente o dado que produziu, sem acessar o ambiente de suporte à decisão, e com isso, deixe de obter outras informações interessantes que ele não possui; e a questão da atualidade, pois demorar muito tempo para atualizar no banco de dados para suporte à decisão um dado que o município já dispõe, também leva ao descrédito e ao abandono da solução, como acontece, segundo foi mostrado, com alguns sistemas de informação após implantados.

A confiabilidade e a qualidade também estão relacionadas. A confiabilidade diz respeito ao grau de precisão do dado informado, pois não adianta ter o dado, atual e em tempo oportuno, se não há precisão nele, se não é possível confiar na informação que ele possibilita. Da mesma forma que os anteriores, este critério da confiabilidade poderá fazer todo o ambiente de suporte à decisão cair em descrédito.

O custo, segundo a autora, segue a lógica de custo-benefício, “entendendo-se que os benefícios derivados do uso da informação é que determinarão se o custo para obtê-la foi satisfatório”<sup>18</sup>.



Finalmente, a qualidade da informação é uma mistura, ou está relacionada, a todos os outros atributos. Assim, a qualidade é uma resultante do processo onde o dado confiável e atual deve ser disponibilizado oportunamente, a um custo razoável e ser acessível (ser de fácil acesso) ao usuário, ou consumidor da informação.

O envolvimento do usuário está diretamente relacionado à questão da relevância do dado, ou seja, como diz Branco<sup>18</sup>,

a noção de significado está associada ao conceito de relevância, que se refere ao grau de adequação entre o que se espera de uma informação e o que ela parece ser capaz de proporcionar, segundo um julgamento fundamentalmente subjetivo e situacional. Assim, a possibilidade de um dado ser informativo dependerá das circunstâncias que envolvem o emissor e o receptor a cada momento, determinando o grau de relevância que lhe será atribuído naquela situação específica.

Então, dentro do esquema emissor-receptor, a relevância é considerada dos dois pontos de vista, mas seu valor será dado por quem a irá assimilar e utilizar.

Há que se ter cuidado, quando não se envolve o usuário como parte do problema e da solução, quando não se busca o que Vasconcellos, Moraes e Cavalcante<sup>12</sup> chamaram de “aprofundamento da circulação inter/transdisciplinar do conhecimento entre campos de saberes”, com a exclusão de “variáveis políticas, econômicas, sociais e culturais envolvidas no processo de produção de informação”<sup>18</sup>, agarrando-se ao processo técnico de organização, tratamento e análise dos dados, numa lógica linear em que dado gera informação, que produz conhecimento, que leva à ação (dado -> informação -> conhecimento -> ação).

É o paradigma onde “usuário é aquele que usa o que foi produzido por aquele que detém o ‘saber’ – o técnico –, que por sua vez é aquele que sabe o que o outro necessita.”<sup>18</sup>

Para ressaltar este fato de que o usuário é aquele que usa o que já foi produzido, Gómez<sup>35</sup> nos mostra que “o contato da fonte com o usuário, para partir da fórmula já clássica, acontece, de fato, *a posteriori* de um julgamento antecipado de relevância, que acontece quando o usuário escolhe ou aceita um item de informação (...) a partir de uma ‘oferta’ do sistema igualmente organizada pelas expectativas sistêmicas de pertinência e relevância”.

Em suma, dentro do atual paradigma de construção de sistemas de informação, sejam eles sistemas de informações transacionais, de ponta, ou sistemas de apoio à decisão, os técnicos e especialistas se colocam como intérpretes das necessidades dos usuários, os quais, por sua vez, aceitam este papel, julgando-se incompetentes para atuar como sujeitos do processo de geração de informação.

O capítulo 6 desta dissertação irá discutir a classificação dos usuários, ou consumidores de informação, para um ambiente de suporte à decisão, bem como uma sugestão de como obter suas preferências dentro de um conjunto de informações em saúde oferecidas a ele.

No entanto, é bom caracterizar o fato de que uma dada informação certamente terá importância para outros usuários além do gestor, ou produtor desta informação. A constante preocupação no desenvolvimento de soluções de informática para gestores frequentemente esquece as necessidades informacionais de outros usuários.

Esta atenção a um espectro mais amplo de consumidores da informação está registrada na diretriz de nº 6 da Política Nacional de Informações e Informática em Saúde, que pode ser baixada do *site* do DATASUS, na *Internet*<sup>6</sup>:

“Estabelecer mecanismos de compartilhamento de dados de interesse para a saúde e ampliar a produção e disseminação de informações de saúde de forma a atender tanto às necessidades de usuários, profissionais, gestores, prestadores de serviço e controle social, quanto ao intercâmbio com instituições de ensino e pesquisa, outros setores governamentais e da sociedade e instituições internacionais”

Uma outra classificação dos usuários que seria interessante abordar, pode ser encontrada em Leme Filho<sup>15</sup>, que os separa da seguinte forma:

a) Usuários exploradores

São aqueles que sabem usar as ferramentas OLAP e têm bom conhecimento técnico em informática. Com isso, conseguem montar suas próprias consultas ao banco de dados de suporte à decisão, baseados nos seus bons conhecimentos do metadados. Ou seja, sabendo o que existe disponível no *data warehouse / data mart*, e sabendo quais são os recursos que a ferramenta de consulta lhes oferece, estes usuários são capazes de montar suas próprias consultas, fazendo os cruzamentos que julgam pertinentes, sem o auxílio de nenhum suporte adicional.

Na verdade, estes poderão ser os grandes “marqueteiros” deste ambiente, se conseguirem obter os resultados esperados para suas consultas. Por terem acesso às ferramentas OLAP, em geral têm também uma ótima infraestrutura de rede e *Internet*, e de equipamentos, que lhes permite aproveitar todos os recursos destas ferramentas.

#### b) Usuários casuais

Estes usuários, em geral, têm um bom conhecimento dos metadados, mas podem não conhecer todos os recursos de uma ferramenta OLAP, ou não têm acesso a ela.

Com isso, recebem modelos de consultas previamente elaborados pela área de administração do *data warehouse/data mart* (ou uma área de *Business Intelligence*), e são capazes de realizar pequenos ajustes para adequar estas consultas às suas necessidades.

Certamente irão precisar de um suporte técnico para realizar os ajustes em suas consultas ou mesmo para operar as ferramentas de consulta de que dispõem.

#### c) Usuários de Consultas Prontas

Estes são os usuários que menos conhecimento têm do uso das ferramentas para elaboração de consultas, sejam elas ferramentas OLAP ou não.

É talvez o maior grupo de usuários do ambiente de suporte à decisão, principalmente se for levado em conta que os *data warehouses/data marts* estarão disponibilizados na *Internet*, como dados de acesso público.

Considerando o diferente leque de pessoas que irão fazer essas consultas, pode-se afirmar que a grande maioria não tem conhecimentos de informática necessários para montar suas próprias consultas ou alterar as já existentes.

Assim sendo, é essencial que se disponibilize um cardápio de consultas prontas, dentro das mais usadas, de uma forma geral, para tentar suprir as necessidades informacionais deste imenso grupo.

Mas, o importante é ter em mente que, quando se fala em montar um ambiente de suporte à decisão em saúde, não existe apenas o usuário gestor das informações, mas vários grupos de usuários, ou consumidores de informação, que irão se beneficiar com o que for disponibilizado através deste ambiente de suporte à decisão.

Uma sugestão de como classificá-los será feita no capítulo 6 deste trabalho.

## 5. A informação e o controle social

O objetivo deste capítulo não é entrar na discussão sobre o controle social, mas mostrar a importância de um ambiente de suporte à decisão e, em última instância, das informações em saúde, como mecanismos de apoio ao controle social e do exercício da cidadania.

Segundo Silva, Cruz e Melo<sup>36</sup>, “a expressão controle social tanto é empregada para designar o controle do Estado sobre a sociedade, quanto para designar o controle da sociedade sobre as ações do Estado.”

Embora os ambientes de suporte à decisão, em especial seus *data warehouses e data marts*, possam servir como instrumento de controle do Estado sobre a sociedade, é a segunda parte da definição que se quer colocar em foco, ou seja, o seu uso para o controle da sociedade sobre as ações do Estado.

Como visto no capítulo 4, a importância dos elementos que podem influenciar a atribuição de relevância pelo usuário/consumidor da informação (atualidade, oportunidade, confiabilidade, qualidade, custo e acessibilidade), se dão, também, pelo fato - como afirma Branco<sup>18</sup> e pelo que também foi dito por Moraes<sup>7</sup> - de que “a informação é um instrumento insubstituível de controle de poder, ou um contrapoder”.

Mas este poder precisa estar acessível. Ainda segundo Silva, Cruz e Melo<sup>36</sup>, o aparato que envolve a informação em saúde deve estar a serviço dos sujeitos políticos, o acesso às informações precisa ser democratizado e as barreiras existentes, superadas. Ou seja, a acessibilidade da informação passa a ser condição necessária para o exercício pleno da cidadania.

Já no relatório final da 9ª Conferência Nacional de Saúde estava registrado que “o acesso à informação é fundamental para garantia da democratização e viabilização do SUS. O controle das atividades de saúde por parte da sociedade só é possível se houver acesso livre à informação de boa qualidade, apresentada de forma clara e acessível a todos.”<sup>37</sup>

Mas, uma vez garantida a acessibilidade, como atribuir relevância à informação que está nestes repositórios de dados? Ou, melhor dizendo, o que a população pode fazer com este acervo, do ponto de vista do controle social? Ela sabe a sua real utilidade, o seu

significado? Sabe como cruzar os dados para obter informações precisas, relevantes? Quais são os variados interesses sobre estes acervos?

Uma das formas de democratizar a informação em saúde e torná-la acessível e útil para esta população é implantar e manter em funcionamento os Conselhos de Saúde, os quais são instituídos nas três esferas de governo.

Segundo Silva, Cruz e Melo<sup>36</sup>, estes órgãos devem tratar da “formulação de estratégias e do controle da execução das políticas de saúde, com pleno acesso aos aspectos econômico-financeiros voltados para o setor, propondo critérios para a programação, a execução, o acompanhamento, a avaliação e, também, a elaboração e a aprovação do plano de saúde, bem como o estabelecimento de estratégias para a sua execução”.

O relatório final da 9ª Conferência Nacional de Saúde já fazia referência aos Conselhos e à garantia de acesso à informação:

-“garantia de acesso, aos Conselhos, a todas as fontes de dados e informações da administração sobre a questão financeira da seguridade social e vigilância epidemiológica.”

-“obrigatoriedade dos órgãos governamentais e setor privado contratado ou conveniado em prestar toda e qualquer informação necessária ao adequado cumprimento das competências dos Conselhos, fornecendo, inclusive, estudos e assessoria técnica”.

-“treinamento, através de Oficinas de Trabalho periódicas, para a capacitação dos conselheiros municipais para exercerem corretamente suas atribuições, dando informações específicas e necessárias sobre como funciona o sistema”.

-“garantia de que as informações sobre Saúde sejam veiculadas para o público em geral através dos meios de comunicação, instituições e movimentos sociais”.

-“que os Conselhos de Saúde constituam, quando necessário, assessorias técnicas de informação, em todos os níveis de governo, com uniformidade, transparência, amplitude, qualidade e fidelidade”<sup>37</sup>.

O que pode ser observado a partir deste relatório é que os Conselhos deverão ter suporte, inclusive em capacitação, dos órgãos governamentais, para que possam exercer suas atribuições legais. Isto tem um impacto imediato sobre as soluções de *data warehouse/data mart* e das ferramentas OLAP, para que destes possa ser extraído material relevante para o apoio ao trabalho a ser exercido pelos Conselhos.

O outro ponto a ser observado é a preocupação do que Branco<sup>18</sup> coloca como a assimilabilidade da informação. A assimilabilidade refere-se à capacidade de tornar a informação assimilável para o cidadão comum, ou seja, indo além da simples acessibilidade. Isto deve ser alcançado ainda que os Conselhos de Saúde tenham que recorrer a assessorias técnicas. Esta é a preocupação da 9ª Conferência com a informação que, embora esteja disponível, não permite que seja usada como instrumento para a ação social.

Como será discutido no capítulo 7, os instrumentos deste ambiente de suporte à decisão, seus *data warehouses / data marts* e as ferramentas OLAP com suas consultas disponibilizadas deverão ter em mente esta assimilabilidade, e, como sua disseminação será feita principalmente usando a *Internet* como meio, com um público-alvo disperso, ela deverá prover recursos de auto-instrução e de ensino à distância para tentar reduzir ao mínimo as dúvidas com sua utilização e com o seu conteúdo.

É preciso entender que, uma vez concluindo-se que o controle social é exercício da cidadania, não se pode deixar unicamente nas mãos do Estado este controle, exigindo “a presença de organizações legítimas de representação de interesses dos diversos segmentos sociais”.<sup>18</sup>

De acordo com Moraes<sup>7</sup>,

“Não se pode deixar tal poder “nas mãos” unicamente dos produtores/gestores de informações. É preciso democratizar tanto poder, incluindo no processo novos interlocutores que podem contribuir com preocupações diferentes, para além da busca da eficiência do “olhar do Estado” sobre a sociedade, defendendo, por exemplo, exatamente, a necessidade de impor limites a esses dispositivos.”

Podemos, por exemplo, encontrar no site Descentralização On Line (DOL)<sup>39</sup>, no menu principal, em “Reuniões da CIT”, na reunião de 17-11-2005, que trata de “Considerações sobre a Política de Informação e Informática”, entre discussões acerca de itens específicos sobre alguns sistemas de informação do Ministério da Saúde, detalhes importantes como atos deliberativos do Ministério da Saúde tomados de forma unilateral, como a descontinuidade do desenvolvimento do Sistema de Regulação e de Identificação do Usuário (SISREG). Encontramos, também, uma proposição do CONASS (dentre várias): “Diante dos problemas apresentados, o CONASS propôs: que

todos os assuntos relativos à informação e informática, de interesse dos gestores das três esferas de governo, fossem submetidos à aprovação da CIT”.

Estes instrumentos de *data warehouse/data mart* servirão, então, como mecanismos de transparência, conferindo visibilidade aos atos dos gestores, e permitirão o acesso às informações necessárias à ação de controle.

É a expansão da democracia representativa para algo além do direito de votar, para a democracia participativa, como propõe o documento do Ministério da Saúde, de 1993, intitulado “Descentralização das ações e serviços de saúde: a ousadia de cumprir e fazer cumprir a lei”<sup>38</sup>:

“Trata-se, portanto, de consolidar instrumentos que viabilizem a prática da democracia participativa, criando condições para os cidadãos, através de órgãos representativos da sociedade civil, participarem das decisões do Estado que afetam suas vidas nas questões de saúde.”

É evidente, como mostra Moraes<sup>7</sup>, que o gestor eleito tem responsabilidade na gestão da República (assim como na produção dos dados para o ambiente de suporte à decisão em saúde). Mas é direito do cidadão tomar parte neste processo.

Ainda, segundo Moraes<sup>7</sup>,

“A renovação da práxis democrática começa pela construção de mecanismos democráticos de participação política que não restrinjam esta ao ato de votar. Implica, pois, uma articulação entre democracia representativa e democracia participativa, em um processo complexo de ações criativas desencadeadas por uma sociedade, redefinindo e ampliando o campo político de atuação, como, por exemplo, com a constituição de mecanismos de gestão colegiada da informação em saúde.”

O conceito de “mecanismos de gestão colegiada da informação” se encaixa perfeitamente (como será visto no próximo capítulo) na abordagem metodológica proposta por Vasconcellos, Moraes e Cavalcante<sup>12</sup>, quando falam em “aprofundamento da circulação inter/transdisciplinar do conhecimento entre campos de saberes”. Embora corra o risco de ser repetitivo, o autor desta dissertação citou novamente este trecho do artigo de Vasconcellos, Moraes e Cavalcante porque ele tira o centro do ambiente de informações para apoio à decisão da monodisciplinaridade da Tecnologia da

Informação. Uma boa proposta de se conseguir isto é através dos mecanismos de gestão colegiada da informação.

Como diz Moraes<sup>7</sup>,

“Daí a essencialidade de propostas que impliquem a ampliação do leque de interlocutores participando, de forma ativa, não só de todo o processo informacional, conforme exposto anteriormente, mas, das questões que se referem à própria Política de Disseminação de Informação, integrante da política da verdade, e às políticas públicas, em um projeto de democracia participativa e emancipatória.”

Estas instâncias colegiadas de gestão da informação permitiriam, inclusive, a criação e a manutenção de critérios explícitos para a alteração dos bancos de dados de *data warehouse/data mart*.

Ou seja, dada a complexidade das bases de dados nacionais, estaduais e municipais, é inconcebível que um gestor isolado decida sobre dados e questões que afetam um conjunto inteiro de informações de saúde e seus consumidores, que repercutem sobre séries históricas e podem, até mesmo, dismantelar um acervo de dados construído. É inadmissível que este gestor, sem consultar ninguém, altere o conceito de um campo dentro de um sistema de informação que irá repercutir num banco de dados de suporte à decisão, mudando completamente seu significado neste mesmo banco de dados, tornando a informação sem relevância ou valor algum para os demais consumidores da informação.

Segundo Moraes<sup>7</sup>, isto indica uma profunda irresponsabilidade com a gestão da “coisa pública” e traz à tona o conceito de “informação pública”, ou de informação como um bem comum.

“O fundamental é a conquista, por parte da população, do princípio de que bases de informações não são “do governo”, mas, sim, “da sociedade” e, portanto, seus representantes devem participar das decisões sobre sua gestão.”<sup>7</sup>

Ou seja, a informação deixa de ser “do governo” e passa a ser pública, enfatizando o princípio de que ela é dever do Estado e direito do cidadão.



Este é o ponto de vista que o autor deste trabalho deseja enfatizar na relação analista de sistemas-usuário, produtor-receptor. Ou seja, que os produtores de informação deixem de tratar a informação como uma “informação governamental”, apropriada para que o Estado controle o cidadão e para interesse próprio dos gestores, para uma “informação pública”, que tem valor não só para o gestor, mas para outros consumidores de informação, e que é chave para o controle social e o exercício da cidadania.

Estas instâncias coletivas de gestão da informação atuariam, também, como agentes para superar a pouca comunicabilidade entre os sistemas de informação e os agentes produtores de informação nacionais, estaduais e municipais.

Esta proposta, encontrada em Moraes<sup>7</sup> é inovadora, e como tantas outras coisas, deverá encontrar resistência entre aqueles que desfrutam a situação estrutural atual.

No entanto, opor-se a esta inovação é abraçar um modelo que já se mostrou obsoleto, em que os interesses políticos, calçados no arsenal tecnológico, suplantam os interesses de informação da saúde, onde o poder de decisão é alicerçado nas instâncias de Tecnologia da Informação, que muitas vezes justificam suas decisões com base na evolução tecnológica de ferramentas e técnicas, mais uma vez impondo ao usuário a condição de espectador passivo da construção de soluções, as quais, futuramente, serão rejeitadas por ele. Para além disso, há o risco de que estas instâncias de Tecnologia da Informação, esquecendo o seu compromisso com a Saúde Pública, acabem se transfigurando em Centros de Processamento de Dados das instâncias superiores às quais estão subordinadas.

## 6. Um método para construção de data warehouses – etapas principais

Antes que sejam discutidas as fases de construção de um *data warehouse* e seus *data marts*, é preciso especificar o que este trabalho não pretende abordar, mas que está diretamente relacionado, até mesmo faz parte, do processo de construção do *data warehouse* e dos *data marts*. O fato deste trabalho não entrar nesses assuntos não significa que não são importantes, mas eles cabem em uma discussão a parte, e sua inclusão aqui introduziria um ponto de fuga ao eixo central desta dissertação.

Parte-se do pressuposto, por exemplo, de que já está definido que será desenvolvido um *data warehouse* destinado à disseminação de informações em saúde, e que toda a discussão política e técnica que resultou nesta decisão chegou ao fim. Não se pretende aqui entrar no mérito de qual é o escopo deste *data warehouse*, se de todos os campos de atuação do Ministério da Saúde, ou se de uma pequena parte destes. Ou seja, já se fez um mapeamento de que áreas serão cobertas pelo *data warehouse*.

Também se parte do princípio de que toda a negociação em torno dos recursos financeiros e humanos para o andamento do projeto já foi realizada e chegou-se a um consenso a este respeito.

Ainda, já se sabe onde o *data warehouse* e os *data marts* serão instalados, e toda a infraestrutura necessária à fase de implantação já foi negociada. Todos os atores a serem envolvidos já foram levantados e estão comprometidos com o projeto.

Partindo destes pontos básicos, é possível dar continuidade à construção do *data warehouse* e dos *data marts*.

Como foi visto na contextualização deste trabalho, um *data warehouse* é composto de seus vários *data marts* temáticos.

A construção dos *data marts*, e, em última instância, do *data warehouse*, é, na verdade, a construção dos bancos de dados físicos que o compõem, desde a sua modelagem – sua representação do mundo real – até suas estruturas físicas, considerando os aspectos de armazenamento, desempenho e outros.

Este capítulo pretende abordar brevemente as principais etapas de construção de um *data mart* / *data warehouse*, sem se aprofundar em cada uma delas, tendo em vista que isto iria carregar o documento de aspectos técnicos que já estão amplamente cobertos na literatura sobre o assunto e que já são de domínio de quem se dispõe a construir um *data mart* / *data warehouse*. Além disso, existem diferentes métodos e ferramentas para a

realização de cada uma das etapas. Com isso, o autor julga mais importante mostrar os passos necessários e estabelecer a relação deles com o gestor e os demais consumidores da informação produzida por este(s) *data mart(s)* e pelo *data warehouse*.

Assim, este capítulo dirige-se, principalmente, aos profissionais de ciências sociais interessados na construção destes armazéns de dados, e não aos profissionais de TI, uma vez que não se deve esperar poder utilizar este método de forma estrita, como um manual, na produção de um *data warehouse* ou seus *data marts*.

### 6.1 Etapa 1: A orientação temática dos *data marts* e do *data warehouse*

Em primeiro lugar, considerando que o *data warehouse* é composto pelos *data marts* temáticos, é preciso saber como escolher o “motivo” ou tema dos *data marts*.

Pela experiência do autor e pela literatura corrente, é comum ver *data marts* associados aos sistemas de informação que fornecem os dados que os alimentam. Por esta ótica, existiria o *data mart* do CNES, o *data mart* do SIA, o do SIH, o do PNI, e depois seria buscada uma forma de integrá-los.

Outra forma comum de organizar os *data marts* é por departamentos. São os *data marts* departamentais. Assim, associado a cada departamento, ou subdivisão deste, na estrutura do Ministério da Saúde, haveria um *data mart* correspondente. Desta forma, seria possível existir um *data mart* para a SAS, outro para a SVS, etc.

Uma das primeiras coisas aprendidas em análise de sistemas é que os dados são mais estáveis que os processos que os geram ou manipulam. Assim sendo, os dados “imposto de renda retido” e “valor pago por SADT” são muito mais difíceis de serem mudados, ou deixarem de existir, do que a forma como são gerados, o processo que os manipula.

Analogamente, podemos inferir que a criação de novos sistemas de informação, ou a extinção de outros, tendem a ser mais frequentes do que a alteração dos campos de atuação do Ministério da Saúde. Da mesma forma que a estrutura hierárquica do Ministério da Saúde, alterada quase que a cada mudança de governo, é mais mutável do que os campos de atuação deste Ministério.

Assim sendo, esta metodologia recomenda que se busque critérios mais estáveis para a escolha dos temas dos *data marts*, como, por exemplo, os campos de atuação do Ministério da Saúde, como os levantados na tese de Machado<sup>40</sup>, a saber:

- a) Assistência à saúde (incluindo a regulação do segmento de assistência médica suplementar);
- b) Vigilância epidemiológica (e prevenção e controle de doenças sob vigilância);
- c) Vigilância sanitária;
- d) Desenvolvimento, produção e regulação de insumos para a saúde;
- e) Formação e gestão de recursos humanos para a saúde.

Esta escolha dos temas deve, na verdade, obedecer à orientação do próprio *data warehouse*. Ou seja, se os *data marts* forem escolhidos com base no campo de atuação do Ministério da Saúde, não tem sentido a inclusão de um novo *data mart* com base na estrutura hierárquica departamental.

Assim sendo, a orientação temática dos *data marts* deve ser uma decisão inicial do projeto do *data warehouse*. Ou seja, ao se decidir pela criação de um *data warehouse*, deve-se identificar os campos de atuação ou áreas a serem abordadas, com a sua respectiva divisão em *data marts*, os quais deverão seguir uma orientação temática, e também deve-se resolver qual será o primeiro *data mart* a ser confeccionado.

Vale lembrar que esta é apenas uma recomendação, e que, qualquer forma de orientação temática dos *data marts* deve ser uma atividade executada pelos representantes de todos os campos do conhecimento envolvidos em sua construção.

## 6.2 Etapa 2: A classificação dos consumidores da informação

A etapa seguinte é a classificação dos usuários ou consumidores da informação (o termo ‘usuário’ ou ‘consumidor de informação’ pode ser considerado de igual significado, para efeito deste trabalho).

Pode-se, em primeiro lugar, tentar estabelecer quais são os grandes grupos de consumidores de informação e suas necessidades de informação.

Uma classificação possível é obtida no livro “A experiência brasileira em sistemas de informação em saúde - volume 2”<sup>41</sup>, que “apresenta o trabalho de investigação qualitativa baseado em entrevistas com gestores qualificados, procurando identificar as experiências exitosas, os pontos problemáticos existentes e as sugestões propostas pelos

entrevistados em prol do aprimoramento dos sistemas de informação em saúde, da Ripsa e do DATASUS”.

Dos grupos entrevistados, alguns podem formar uma boa tentativa de se estabelecer o conjunto dos principais consumidores de informação em saúde:

- a) Gestores – profissionais responsáveis por fazer cumprir a lei no que se refere a receber, analisar, enviar e divulgar os dados de saúde no país, nos três níveis: federal, estadual e municipal.
- b) Pesquisadores – pessoas que desempenham atividades de ensino e pesquisa e cujas linhas de investigação exigem uso constante de dados dos sistemas de informação em saúde para a produção de artigos, teses, consultorias e avaliação.
- c) Conselheiros de Saúde – integrantes do Conselho Nacional de Saúde e dos conselhos estaduais e municipais.

Além destes, pode ser considerado um quarto grupo:

- d) Usuários comuns – composto pelas pessoas sem nenhum vínculo com o SUS, mas que, considerando que as informações disponibilizadas são públicas, estão fazendo alguma investigação de cunho jornalístico, ou por interesse pessoal ou político.

É com este quarto grupo que deve, talvez, haver a maior preocupação no que diz respeito à acessibilidade e à facilidade de uso das interfaces entre o consumidor da informação e o(s) *data mart(s)*, uma vez que este grupo tende a não dominar nem o jargão do campo de conhecimento da saúde pública, nem a linguagem própria e os comandos comuns às ferramentas OLAP. Para tal, como será visto no capítulo dedicado à disseminação das informações contidas nos *data marts / data warehouses*, é fundamental que estas ferramentas sejam acompanhadas de um sistema de ajuda bem estruturado, que permita a navegação por parte destes consumidores sem a necessidade de nenhum suporte adicional.

Em seguida, deve ser provida alguma forma de tentar identificar o consumidor da informação, individualmente. Não se trata, na verdade, de manter um bom relacionamento com o consumidor da informação, embora isto não seja excluído, mas de buscar a sua satisfação, ou seja, de observar a sua real necessidade de informação, de procurar mostrar uma visualização da informação que corresponda ao seu conhecimento e expectativas, de manter um canal permanente de comunicação e atendimento, de estabelecer uma relação de confiança com ele.

Para tal, a pergunta que primeiro deve ser respondida é: “Quais são os consumidores de informação deste *data mart*?”

A resposta a esta pergunta é o primeiro objetivo específico que este trabalho irá procurar alcançar. A identificação dos consumidores de informação e sua classificação é fundamental para a definição das formas de visualização e disponibilização das informações em saúde. Sem ela, não há personalização.

Dentro desta pergunta principal, estão embutidas outras perguntas secundárias, porém igualmente importantes, como: onde estão localizados? Que nível de importância têm nos processos decisórios ligados à gestão do SUS? E nos processos executivos? Quais são as suas necessidades em relação às informações de saúde a serem disponibilizadas? São gestores daquelas informações específicas? Que conhecimento possuem sobre o significado daquelas informações e que conhecimento possuem para poderem analisá-las? Conseguem identificar padrões e tendências que possam transformar aquelas informações em conhecimento? São afetados diretamente por aquelas informações? Qual é o seu nível de comprometimento com os objetivos da Política Nacional de Informações e Informática?

Em tese, qualquer cidadão brasileiro pode ser um consumidor da informação, estando incluído em algum dos grupos estabelecidos. Porém, deve-se ter em mente que o data warehouse e seus data marts não irão aumentar a quantidade de consumidores que iriam buscar informações de saúde, independente da tecnologia utilizada para sua disseminação. Ou seja, a adoção das tecnologias de data mart / data warehouse deve ser pensada como um elemento facilitador para aqueles que necessitam das informações.

Além disso, considerando os aspectos tecnológicos das possíveis soluções de implementação, é preciso saber: que recursos de equipamentos computacionais, de telecomunicações e humanos possui o consumidor daquelas informações? Qual é o seu nível de conhecimento nas tecnologias da solução apresentada?

É importante que a classificação dos consumidores de informação seja feita não de acordo com uma estrutura hierárquica e organizacional, mas em consonância com os campos de atuação do Ministério da Saúde. Este rol de perguntas não pretende, de nenhuma forma, ser um conjunto fechado, e, certamente, novas perguntas surgirão com o uso do próprio método.

Para que esta identificação se concretize, e as perguntas possam ser respondidas, é necessário que as ferramentas utilizadas pelos consumidores de informação para acesso aos bancos implementem um mecanismo de autenticação, do tipo *login/senha*.

Evidentemente que esta autenticação não será obrigatória, e o consumidor da informação poderá navegar pelos instrumentos de acesso ao banco de forma anônima. Porém, a identificação por *login/senha* permitirá que o consumidor se beneficie de alguns recursos, como por exemplo, a possibilidade de guardar suas consultas mais utilizadas, de criar uma “lista de favoritos”, contendo as estruturas do *data mart* que acessa mais frequentemente, dentre outras facilidades; em suma, permitirá que ele crie um “ambiente de trabalho” que lhe seja familiar.

Em contrapartida, no cadastramento de seu *login/senha*, lhe será apresentado um questionário contendo perguntas do tipo das enumeradas anteriormente, que permitirá à equipe de administração dos *data marts/data warehouses* traçar um perfil de cada consumidor de informação, e, eventualmente, manter um sistema de estatísticas e perfis de acesso.

É de se esperar que tais perguntas, aliadas à elaboração dos próprios sistemas de apoio à decisão, gerem expectativas aos próprios consumidores de informação, como é comum no desenvolvimento de qualquer sistema de informação. Logo, é preciso criar um mecanismo de retorno destas expectativas, uma forma pela qual os construtores destes sistemas de apoio à decisão possam avaliar a qualidade do serviço que está sendo prestado.

Uma proposta é que esta avaliação seja feita de três formas: a primeira delas através de um *e-mail* exclusivo para contato dos consumidores; a segunda através de um número telefônico, de preferência com discagem gratuita; e a terceira através de um formulário *on-line*, cujas dúvidas, sugestões e respostas fiquem armazenadas em um banco de dados, permitindo a criação de uma lista de perguntas frequentes (FAQ) a ser disponibilizada em um *site* específico de uma área de *Business Intelligence* (BI) responsável por estes *data marts/data warehouses*.

Estes canais de comunicação permitirão checar constantemente os pontos críticos do processo de trabalho e a validade do método de forma integral.

Além disso, alguma forma de retribuição ao trabalho de resposta às perguntas de identificação deve ser pensada, além da possibilidade da personalização do ambiente de trabalho, como por exemplo, a possibilidade de manter o consumidor da informação

atualizado, via *e-mail*, sobre a criação de novos *data marts* ou as atualizações estruturais, com inclusões, alterações e exclusões de tabelas em *data marts* já existentes.

Assim, a maneira de estabelecer quais grupos de consumidores da informação, com seus elementos identificados individualmente, têm interesse em um novo *data mart*, pode ser resumida da seguinte forma:

É preciso que na fase inicial de criação do *data warehouse*, esteja definida a classificação dos grupos de consumidores da informação.

Quando o primeiro *data mart* for criado, e a cada criação de um novo *data mart*, deverá ser feita uma varredura nestes grupos para verificar que necessidades de informação eles têm sobre o novo *data mart*. Para ser mais preciso, a cada nova mudança estrutural de um *data mart*, o impacto sobre os grupos deverá ser avaliado.

Cada consumidor da informação individual, identificado, será associado a um grupo de consumidores da informação quando se autenticar na ferramenta de acesso aos dados, e herdará as preferências daquele grupo ao qual foi adicionado. A exceção ficará por conta do usuário anônimo, que será associado ao grupo de usuários comuns, e terá acesso a todos os dados públicos, sem preferências.

Vale lembrar que as preferências de dados, exibidas como uma espécie de “Meus Favoritos” aos consumidores de informação quando estes se autenticarem na ferramenta de acesso aos dados, não excluem a sua possibilidade de ver todos os dados, uma vez que são públicos, nem impedem que eles montem a sua própria lista de “Meus Favoritos”.

É preciso observar que os cadastros de consumidores individuais serão bem dinâmicos, pela própria característica de uso do ambiente de suporte à decisão, com a entrada e saída frequente de usuários.

Mais uma vez, é bom lembrar que este trabalho não tem a pretensão de criar uma lista de grupos de consumidores da informação, tampouco de tentar identificar que tipos de dados são importantes para determinados grupos.

A proposta é que, conforme visto no capítulo 5, sejam criados mecanismos de gestão colegiada da informação, os quais, segundo Vasconcellos, Moraes e Cavalcante<sup>12</sup>, implementem um Ambiente de Informações para Apoio à Decisão, cuja abordagem metodológica esteja assentada na “busca de um aprofundamento da circulação inter/transdisciplinar do conhecimento entre campos de saberes, com destaque para as áreas de: Tecnologias de Informação, Gestão em Saúde, Planejamento e



Epidemiologia”, em busca de uma “metodologia integradora, incorporando métodos e técnicas de diferentes disciplinas que se complementam”.

Estes grupos de gestão colegiada da informação, segundo defende o autor, teriam como uma de suas responsabilidades a definição dos grupos de consumidores de informação, bem como a tarefa de identificar, junto com os gestores do *data mart* específico que está sendo criado, as necessidades de informação dos demais grupos. Também seria objeto de trabalho deste grupo, a conscientização dos gestores para a publicação das informações sob sua responsabilidade, nos *data marts* / *data warehouses*.

Isto é fundamental para que estas definições não fiquem exclusivamente a cargo de um grupo de gestores que está voltado para suas próprias atribuições e necessidades, nem sejam unicamente de responsabilidade de uma equipe de especialistas em TI, cujo campo de conhecimento, mesmo não sendo monodisciplinar, não possui, obrigatoriamente, a transversalidade necessária para perceber as nuances de importância que uma determinada informação em saúde possa ter em diferentes campos de atuação do Ministério da Saúde. A estes especialistas em Tecnologia da Informação cabe mais, como foi visto, o papel de metaconhecedores ativos e intencionais, que podem objetivar, modelar, gerir um processo de conhecimento de outros conhecedores.

### 6.3 Etapa 3: A escolha do *data mart* a ser criado

Outra questão de fundamental importância é a ordem de criação dos *data marts*. Qual será o primeiro *data mart* a ser criado? Que critérios irão nortear esta escolha? Que critério irá determinar o próximo *data mart* a ser criado? Quem tomará esta decisão? Que campos de atuação serão afetados?

É fácil imaginar as disputas de poder envolvidas na escolha de um determinado campo de atuação em detrimento a outros. Todas estas questões nos remetem novamente aos mecanismos de gestão colegiada da informação.

De acordo com Vasconcellos, Moraes e Cavalcante<sup>12</sup>:

“A noção de que cabe a uma única instituição coordenar de forma isolada práticas informacionais de diferentes instâncias produtoras é politicamente inviável. Assim, tornam-se relevantes iniciativas de descentralização das

informações, com o estabelecimento de parcerias com as entidades estaduais e municipais, e a busca por alternativas de coordenação de forma colegiada, com ampla participação, tanto de produtores de informações e das sociedades científicas, quanto de representantes da sociedade civil organizada, nos moldes em andamento em vários países.”

Feita a escolha do *data mart* a ser desenvolvido, os próximos passos estão ligados à sua construção.

#### 6.4 Etapa 4: A elaboração do modelo conceitual dimensional

A etapa seguinte é o levantamento das necessidades e a elaboração do modelo conceitual dimensional do *data mart*.

Embora estes conceitos já tenham sido explorados na contextualização deste trabalho, em sua introdução, é necessário reforçá-los aqui.

O conceito de modelo dimensional usa, primeiramente, o conceito de modelo. Conforme dito anteriormente, um modelo é uma representação gráfica dos dados de uma área de interesse específica, ou seja, é a tradução do mundo real em um gráfico cuja linguagem simbólica permite ao usuário validar se o que foi passado por ele está ali corretamente representado.

Dois modelos conceituais são comumente usados na construção de bancos de dados:

- a) O modelo relacional, muito usado na modelagem conceitual de bancos de dados transacionais, que são bancos que devem fornecer uma boa resposta às inclusões, alterações e exclusões frequentes que o atualizam, e cujo estudo está fora do escopo deste trabalho.
- b) O modelo conceitual dimensional, usado em *data warehouses* e *data marts*, aonde se identificam os fatos (uma internação de paciente, um óbito, um nascimento são exemplos de fatos), e também se identificam as dimensões pelas quais se quer visualizar estes fatos (dimensões temporais, dimensões geográficas, dimensões de faixa etária, sexo, escolaridade, etc).

O levantamento de necessidades consiste em estudar os modelos de dados relacionais correspondentes aos bancos de dados transacionais existentes, e outras fontes de dados.

Ou seja, é dos bancos de dados transacionais, que servem aos diferentes sistemas de informação existentes, que serão extraídos os dados para a alimentação do *data mart*.

Além dos bancos de dados transacionais, poderão existir outras fontes de dados, incluindo fontes externas ao Ministério da Saúde. Como bem lembram Vasconcellos, Moraes e Cavalcante<sup>12</sup>:

“As informações de interesse para a saúde não se restringem àquelas produzidas pelo Ministério da Saúde. Destacam-se, por exemplo, as informações estatísticas oriundas do IBGE, fundamentais para o entendimento da situação de saúde em um determinado tempo e espaço geográfico.”

Algumas linhas de desenvolvimento defendem a criação de um pré-modelo resultante deste esforço de levantamento. O pré-modelo conterá a integração das fontes de dados existentes, após a limpeza dos dados indesejáveis, ou seja, dos dados que não interessam ao *data mart* em questão. No pré-modelo serão identificados os fatos de interesse para o *data mart* (por exemplo, internações, nascimentos, faturamentos, etc) e suas dimensões (data, unidade da federação, sexo, faixa etária, etc). Cada fato identificado no pré-modelo e suas dimensões gerará um modelo dimensional. A junção destes modelos dimensionais constituirá o modelo conceitual dimensional do *data mart*.

Posteriormente, a integração do modelo conceitual deste *data mart* com os modelos conceituais dos outros *data marts* irá formar o modelo conceitual dimensional do *data warehouse*.

Nesta etapa é fundamental a participação de representantes de todos os campos de saberes envolvidos, para que o modelo possa expressar o conjunto mais completo possível de necessidades de informação dos diferentes atores. Idealmente, o trabalho de levantamento deve ser feito com todos os interessados ao mesmo tempo, com os profissionais de TI atuando como mediadores dos conflitos de poder que irão surgir.

É nesta fase de modelagem conceitual dimensional que muitos problemas de integração de dados irão vir à tona, com a disputa entre áreas que utilizam o mesmo dado pela hegemonia de significado e representação do mesmo.

Uma definição importante de ser citada é a de métricas, ou medidas, ou medições. Segundo Singh<sup>13</sup>, “medições fornecem os dados numéricos coletados que ficam armazenados na tabela de fatos central”. As métricas são atributos (campos) das tabelas fatos sobre as quais normalmente são feitas as operações de agregação com soma, média e outras.

Exemplos de métricas podem ser: quantidade de leitos de UTI, valor pago SADT, valor líquido recebido pelo hospital, quantidade de nascidos vivos, etc, que poderão ser trabalhados pelas dimensões, como por exemplo, quantidade de nascidos vivos (métrica), no período de 1994 a 2004 (dimensão temporal) em São Paulo (dimensão geográfica); quantidade de leitos de UTI (métrica) no município do Rio de Janeiro (dimensão geográfica); número de casos de gripe suína (métrica) em mulheres (dimensão sexo), na faixa de 20 a 29 anos (dimensão faixa etária), no Rio Grande do Sul (dimensão geográfica), no ano de 2009 (dimensão temporal) .

Juntamente com as tabelas fato e as dimensões, devem ser definidas as métricas que serão utilizadas para cada fato.

É válido lembrar que, pela proposta deste trabalho, o importante é caracterizar a etapa da modelagem conceitual dimensional como um passo da elaboração de um *data mart*, mas não mostrar os detalhes da disciplina de modelagem de dados. É pressuposto do autor que o profissional de TI que irá conduzir a construção do modelo conceitual dimensional já possua os conhecimentos específicos para tal.

Além disso, já existe vasta literatura específica sobre este assunto, como os livros de Ralph Kimball citados nas referências bibliográficas deste trabalho.

Outro ponto de grande importância para o *data warehouse* é o seu catálogo, também conhecido como “metadados”.

Segundo Leme Filho<sup>15</sup>, metadados pode ser definido como “dados sobre dados”, ou ainda como “um mecanismo que possibilita que os usuários do *data warehouse* saibam quais são os assuntos disponíveis, quais as informações que estes contém, que sistema de informação ou fonte externa originou aquele dado, que regras foram aplicadas dos sistema de informação original para o dado atual (por exemplo, o valor masculino para o campo sexo, representado pelo código ‘1’ no sistema de informação que gerou o dado, e que virou ‘masculino’, descrito textualmente, no mesmo campo no *data warehouse*), uma linha do tempo informando que mudanças ocorreram com aquele dado ao longo do tempo (por exemplo, uma cotação de um medicamento que era feita em UFIR e que,

a partir de uma determinada data, passou a ser feita em Real), e outras informações que forem julgadas necessárias.

Por ser um dicionário de dados voltado para compreensão do *data warehouse* por parte do consumidor da informação, é interessante que a atualização do metadados seja feita em paralelo com a construção dos modelos dimensionais.

O metadados é de fundamental importância, como será visto no capítulo de disseminação de informações, para dar ao consumidor da informação uma ideia clara do que é aquele dado, de modo a evitar que mesmo o consumidor de informação mais leigo não compare elementos incomparáveis.

Com o modelo conceitual dimensional do *data mart* pronto, é possível agora integrá-lo ao modelo conceitual dimensional do *data warehouse*.

É um trabalho que requer extremo cuidado, pois é preciso comparar as dimensões do *data mart* com as dimensões existentes no *data warehouse*, verificando se a dimensão do *data mart* ainda não existe no *data warehouse*, ou se já existe com outros atributos no *data warehouse*, se é possível combinar as duas, etc. Também é preciso fazer a mesma coisa com as tabelas fato.

## 6.5 Etapa 5: O projeto físico do banco de dados

A próxima etapa consiste no projeto físico do banco de dados correspondente ao modelo conceitual dimensional do *data mart*.

Por ser uma atividade do campo de conhecimento da informática, ela fica quase que totalmente restrita aos profissionais de TI, mais especificamente aos administradores de bancos de dados.

Do modelo conceitual dimensional serão derivadas as estruturas que irão compor o banco de dados físico. Aliadas a estas surgirão outras estruturas, destinadas a apoiar o banco de dados nas questões de desempenho, de registros de auditoria, etc. Isto é inerente ao próprio *software* que gerencia os bancos de dados, e varia de acordo com o fabricante do *software*. É, no entanto, de conhecimento dos administradores de bancos de dados.

Neste momento, são discutidas questões relevantes a respeito do volume do banco de dados, dos processos de carga, se o banco de dados estará fisicamente num único local ou se será distribuído por vários locais, dos melhores índices (estruturas auxiliares para agilizar o acesso às tabelas de fatos e dimensões) a serem utilizados, dentre outras decisões que dizem respeito à equipe de administração de bancos de dados.

#### 6.6 Etapa 6: Os processos de extração-transformação-carga

Após o projeto físico do banco de dados, a etapa seguinte é o projeto de ETL (*extract, transformation, load*) ou, extração-transformação-carga. É neste momento que serão pensadas as estratégias de alimentação dos dados do data mart a partir de suas fontes. Como o nome sugere, extração-transformação-carga, os dados serão extraídos dos bancos de dados, arquivos convencionais, planilhas, e outros acervos que servem de fonte de dados, serão transformados, ou seja, serão adaptados para uma exibição mais amigável e serão carregados no banco de dados físico do *data mart*. Na verdade, o processo ETL pode ser dividido em tantas fases intermediárias quanto forem necessárias para seu tratamento, antes da inserção nas tabelas do banco de dados físico final.

Por exemplo, pode-se extrair dados de mais de uma fonte para alimentar uma única dimensão ou fato no *data mart*, como, por exemplo, dados de mortalidade oriundos do Sistema de Mortalidade (SIM) e do IBGE. Isto é uma extração; pode-se, também, alterar a forma como o dado é representado numa dimensão ou numa tabela fato, como, por exemplo, numa tabela fato de internação, pode-se mudar a codificação do campo sexo de 0,1,2 para ‘masculino’, ‘feminino’ e ‘ignorado’, respectivamente. Isto é feito normalmente para ganhar tempo na exibição dos dados, por não ter que fazer a conversão naquele momento. Isto é um exemplo de transformação. E a carga é a inserção do registro, já transformado, no *data mart*.

Aqui, nesta fase de ETL devem ser pensados os procedimentos para extração de dados de sistemas de suporte à decisão já existentes.

Embora esta seja uma atividade inerente à equipe de TI, estes processos de extração-transformação-carga podem ser desenvolvidos pela própria equipe de TI usando uma linguagem de programação compatível com o ambiente onde os dados estão armazenados, ou podem ser usados produtos prontos para este fim, disponíveis comercialmente ou sob licença de *software* livre.

A vantagem do uso de produtos específicos para este fim é que eles permitem que todo o processo de extração-transformação-carga pode ser definido utilizando uma interface gráfica, geralmente bastante poderosa, e que facilita e agiliza a sua construção. A interface, por trás dos panos, constrói os programas na linguagem apropriada e permite o agendamento da execução dos programas na ordem necessária e no horário que se quiser determinar.

Contudo, independente do modo de desenvolvimento dos processos de extração-transformação-carga, as regras para transformações dos dados devem ser estabelecidas em conjunto com as demais áreas de conhecimento envolvidas na construção do *data mart*.

O processo de extração-transformação-carga, segundo a literatura, e pela experiência do próprio autor, pode-se transformar num gigantesco entrave ao cumprimento de qualquer cronograma, uma vez que é nesta fase que os dados oriundos dos sistemas de informação e das fontes externas serão depurados para a inclusão no *data mart/data warehouse*. Situações não previstas, problemas inesperados de integridade, diferenças no padrão de representação de dados, dentre outros, podem surgir nesta fase. Algumas situações são de fácil resolução; outras, irão requerer intervenção de instâncias superiores, requisitos técnicos serão preteridos por questões políticas, etc.

Em compensação, é o resultado do bom trabalho realizado nesta fase que irá garantir a qualidade da informação disseminada e, em última instância, a confiabilidade e o sucesso do *data warehouse*.

### 6.7 Etapa 7: A construção das aplicações

Com o banco de dados físico correspondente ao *data mart* criado e carregado, a próxima etapa é a construção das aplicações que irão consultar os dados.

As aplicações para consulta de dados serão discutidas no próximo capítulo, mas, resumidamente, podem ser divididas em três tipos:

- a) As aplicações OLTP, de processamento analítico on-line, que permitem que tanto consultas prontas quanto consultas não previstas (*ad hoc*) sejam executadas sobre o *data mart*;

- b) As aplicações *web*, disponibilizadas em *site* próprio, que permitem que o consumidor de informações acesse um menu de consultas prontas sobre o *data mart*;
- c) As aplicações para extração agendada de dados, que permitem que o consumidor da informação informe os dados que deseja consultar e deixe agendada a extração. A aplicação irá, na data/hora programada, efetuar a extração dos dados e disponibilizará, em uma página em *site* próprio, o *link internet* para que o consumidor da informação faça o *download* do arquivo resultante da extração.

Uma vez mais, estas são apenas recomendações de aplicações, mas a gestão colegiada da informação é que deverá indicar as necessidades de disponibilização da informação.

#### 6.8 Etapa 8: Validação e testes

Como em qualquer desenvolvimento de *software*, é fundamental que o *data mart* criado seja submetido a testes e seus resultados sejam validados pelos gestores e demais representantes do grupo de gestão colegiada da informação, antes da sua liberação para uso.

Como dito anteriormente, as etapas que compõem este método de trabalho são comuns a toda a literatura sobre o assunto, com algumas pequenas variações, e já foram vivenciadas pelo autor, juntamente com outros colegas, em trabalhos anteriores no DATASUS.

Na verdade, o que é importante registrar é, como dizem Vasconcellos, Moraes e Cavalcante<sup>12</sup>, “a circulação inter/transdisciplinar do conhecimento entre campos de saberes” que caracterizam o processo de construção deste “ambiente de informações para apoio à decisão”.

As técnicas e ferramentas para construção de *data marts* e *data warehouses* são amplamente conhecidas no meio da Tecnologia da Informação e o uso destes bancos de dados também não é novidade. A abordagem que se propõe com este método de trabalho, e que se baseia no que já havia sido proposto por Vasconcellos, Moraes e Cavalcante<sup>12</sup> é que o autor não tem conhecimento, em especial no Ministério da Saúde.

Como mostra um trecho do artigo citado acima, dos três autores:



“O desafio que se coloca é pensar mecanismos que favoreçam sua apropriação [dos ambientes de informações para apoio à decisão em saúde] tanto pelos gestores, no exercício diário de sua função de tomadores de decisão, quanto pelos Conselheiros de Saúde e demais representantes da sociedade civil organizada para o pleno exercício de suas funções de sujeitos históricos, comprometidos com o avanço da democracia e da melhoria da saúde da população brasileira.”

## 7. A disseminação das informações

Uma vez criado e carregado o *data mart* ou o *data warehouse*, é preciso tecer uma estratégia de disseminação das informações ali existentes.

Interesses políticos e econômicos são os personagens que disputam o poder neste último ato da montagem do *data mart / data warehouse*. Isto porque uma vez que os dados estão presentes no banco de dados é preciso utilizar alguma ferramenta para ler e exibir estes dados.

Vendedores importantes no mercado e defensores da utilização de ferramentas de *software* livre exercem sua influência, “pressionando” os construtores do *data mart/data warehouse* na expectativa de verem seus produtos presentes na solução final adotada. Na verdade, pelos *data marts* e *data warehouses* serem bancos de dados como outros tantos, os quais são acessados das mais diversas formas, o autor deste trabalho defende que não haja uma única solução de disseminação, mas sim políticas de disseminação que levem em conta as necessidades, os recursos tecnológicos e humanos existentes nos consumidores das informações, e que várias opções de solução possam ser implementadas para vários tipos de consumidores ou usuários.

No entanto, como vem sendo mostrado neste trabalho, é mais um tipo de decisão que justifica e fortalece uma instância coletiva de gestão da informação.

Assim, como visto no capítulo 6, pode-se, a princípio, identificar três tipos de soluções possíveis para a extração e exibição dos dados:

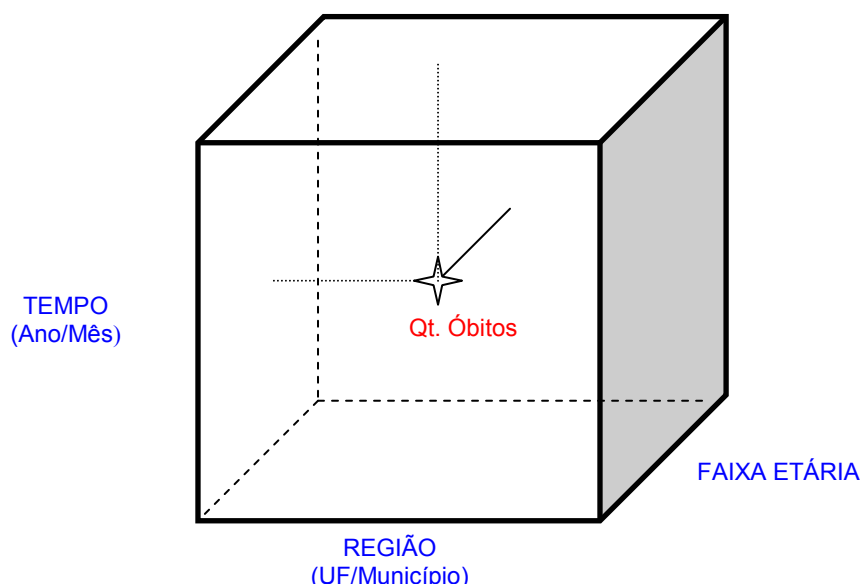
A) As aplicações *OLAP*, de processamento analítico *on-line*, que permitem que tanto consultas prontas quanto consultas não previstas (*ad hoc*) sejam executadas sobre o *data mart*. Este tipo de aplicação requer uma ferramenta específica que permita a tabulação dos dados através de estruturas especiais chamadas “cubos multidimensionais”.

Existem várias formas de implementar estes cubos, o que diferencia as soluções *OLAP*, mas isto foge ao escopo deste trabalho. O importante é entender como funciona um cubo.

Na modelagem dimensional foi visto que FATO é o evento que se quer medir (por exemplo, uma quantidade de órbitos, numa tabela fato “ÓBITO”). Da mesma forma, DIMENSÃO foi conceituada como uma maneira, ou um atributo pelo qual se quer

observar o fato (como por exemplo, as dimensões tempo, região geográfica, faixa etária, sexo, etc).

O cubo nada mais é do que a representação gráfica desta situação acima, onde no centro do cubo está o fato que se quer contabilizar, como por exemplo, a quantidade de óbitos, e cada face do cubo é uma dimensão:



No exemplo acima tem-se um cubo onde o fato a ser medido é a quantidade de óbitos, visto pela ótica das dimensões tempo (ano/mês), região (UF/município) e faixa etária.

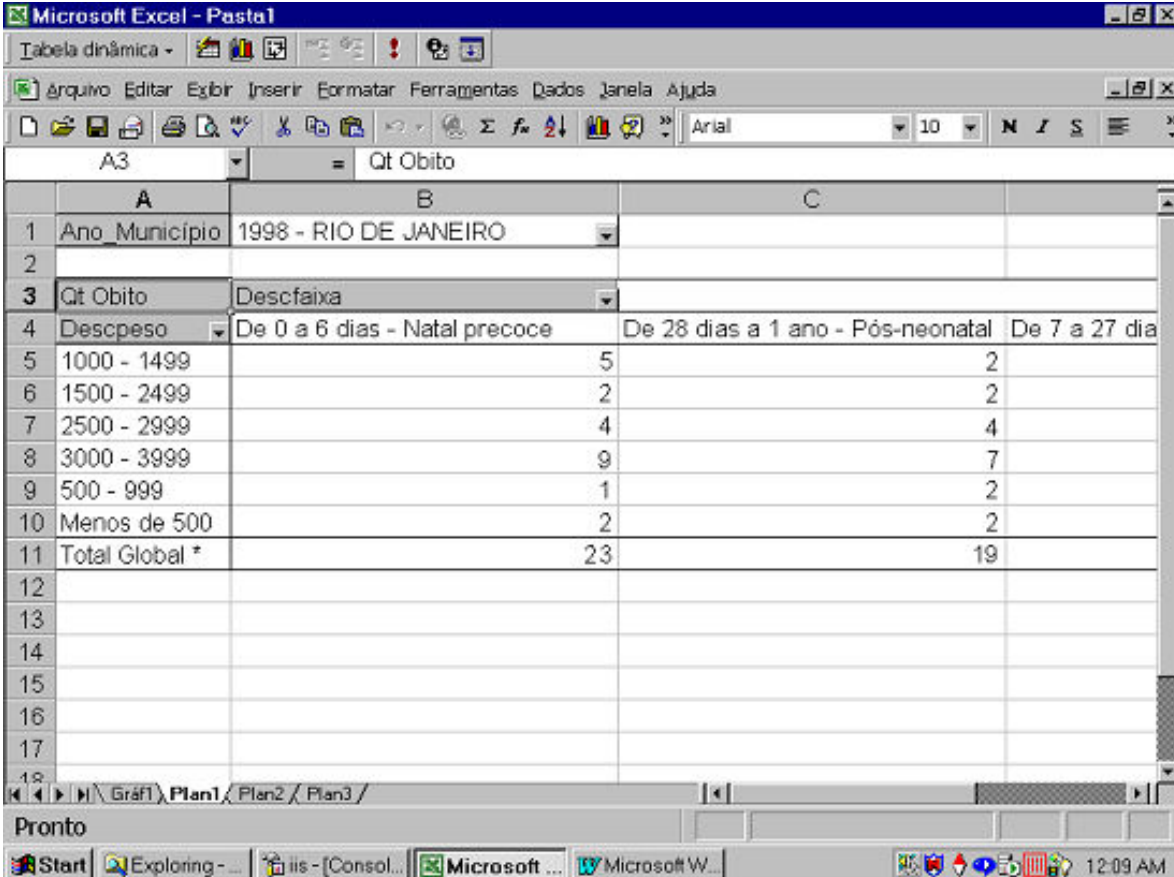
O “visto pela ótica das dimensões” pode ser uma das seguintes coisas:

- Observar a quantidade de óbitos dada uma UF e um município, uma faixa etária específica e um tempo (ano/mês) específico. Isto é observar um ponto do cubo.
- Fixar uma das dimensões (por exemplo, faixa etária) e variar as outras duas, observando a quantidade de óbitos. Isto é uma operação de fatiamento, pois é como se estivesse sendo tirada uma “fatia de bolo” do cubo.

Porém, é difícil ter esta noção sem ver uma representação numa tela de computador, até porque a tela do computador é plana (duas dimensões) e o cubo tem três dimensões.

Como tentativa de exemplo, a figura abaixo mostra uma ferramenta OLAP (na verdade, o Microsoft Excel, que também serve como ferramenta OLAP).

Os dados exibidos são fictícios, servindo apenas como modelo para apresentação da tela da ferramenta OLAP.



The screenshot shows a Microsoft Excel window titled "Microsoft Excel - Pasta1". The active cell is A3, containing the formula "= Qt Obito". The pivot table is structured as follows:

	A	B	C	
1	Ano_Município	1998 - RIO DE JANEIRO		
2				
3	Qt Obito	Descfaixa		
4	Descpeso	De 0 a 6 dias - Natal precoce	De 28 dias a 1 ano - Pós-neonatal	De 7 a 27 dia
5	1000 - 1499	5		2
6	1500 - 2499	2		2
7	2500 - 2999	4		4
8	3000 - 3999	9		7
9	500 - 999	1		2
10	Menos de 500	2		2
11	Total Global *	23		19
12				
13				
14				
15				
16				
17				
18				

Pode-se verificar que o fato a ser observado é a QUANTIDADE DE ÓBITOS (Qt Óbito), que são os dados numéricos nas células centrais da planilha (coluna B, linha 5 até coluna C, linha10)

As dimensões pelas quais se observa este fato são três: a dimensão FAIXA DE PESO (Descpeso), a dimensão FAIXA ETÁRIA (Descfaixa), e a dimensão COMBINADA DE ANO COM MUNICÍPIO (Ano\_Município).

Como é impossível representar três dimensões numa planilha com apenas duas dimensões (linhas e colunas), o Excel (e várias outras ferramentas OLAP) resolvem o problema da seguinte maneira: uma dimensão (faixa de peso) fica nas linhas, outra dimensão fica com as colunas (faixa etária) e a dimensão combinada de ano com município (Ano\_município) é apresentada numa caixa de seleção alterável no topo da

planilha. Isto significa que mudando o valor do Ano\_Município (por exemplo, de 1998 – RIO DE JANEIRO para 2005 – CABO FRIO) os valores do fato Quantidade de Óbitos (Qt Obitos) mudarão automaticamente nas células centrais).

Outras operações são possíveis com estas ferramentas OLAP. Por exemplo é possível trocar de lugar a dimensão Ano\_Município com a dimensão Faixa de Peso (Descpeso) ou com a dimensão Faixa Etária (Descfaixa). Assim, o Ano\_Município passa a ocupar a coluna ou a linha, e a dimensão correspondente (Faixa de Peso ou Faixa Etária) vai para a caixa de seleção no topo da planilha. São as rotações, exatamente como se fossem rodadas as faces de um cubo de verdade.

É possível fazer, também, o que na literatura se chama de “*drill*”, no sentido exato de uma perfuração. Tendo-se, por exemplo, uma dimensão hierarquizada de unidades da federação, e dentro destas, os municípios, é possível visualizar a quantidade de óbitos por unidade da federação. Ao se encontrar uma anomalia em alguma unidade da federação, pode-se desejar fazer um aprofundamento (uma perfuração), o que a literatura chama de “*drill-down*”, visualizando os quantidades de óbitos, detalhadas agora, pelos municípios daquela unidade da federação específica. Da mesma forma, é possível num quadro muito confuso do ponto de vista estético, em que as quantidades de óbitos são apresentadas totalizadas por municípios, fazer o que a literatura chama de um “*drill-up*”, subindo o nível de detalhe para as unidades da federação, que apresentarão, então, os totais dos municípios correspondentes a cada uma delas.

No exemplo acima, da planilha Excel, os dados das faixas etárias são totalizados na última linha, e (embora não esteja visível na figura) os dados de cada faixa de peso são apresentados totalizados na última coluna da planilha, do lado direito.

O interessante é que o cubo, figura geométrica, tem três dimensões, mas, embora este seja o termo utilizado, os cubos multidimensionais reais das ferramentas OLAP frequentemente têm várias dimensões, apenas fica difícil imaginá-los do ponto de vista geométrico.

Isto é só um exemplo, longe de querer ser uma “aula sobre ferramentas OLAP”, para mostrar do que estes *softwares* são capazes de fazer. Tudo isto é feito de forma dinâmica, sem que os dados precisem ser recarregados do banco de dados.

No entanto, são ferramentas que requerem um nível de conhecimento razoável, por parte do consumidor da informação, na área de informática.

Por conta disso, elas dão a opção de que um profissional conhecedor da ferramenta construa consultas “programadas”, e o consumidor da informação pode apenas rodar estas consultas prontas na hora que precisar das informações.

Estas ferramentas irão, também, requerer uma infraestrutura de *hardware* robusta e uma rede de comunicação de dados com uma boa banda disponível para que as transações com os bancos de dados possam ser realizadas em tempo aceitável.

B) As aplicações *web*, disponibilizadas em *site* próprio, que permitem que o consumidor de informações acesse um *menu* de consultas prontas sobre o *data mart*. Mais ou menos como a solução de consultas prontas das ferramentas OLAP. No entanto, neste tipo de solução, as páginas contendo as respostas das consultas são recarregadas a cada solicitação, o que não atrapalha em nada este tipo de solução.

Existe uma grande vantagem no custo de construção destas consultas, que pode ser feita por uma equipe de desenvolvedores de uma das áreas de Tecnologia de Informação do Ministério da Saúde.

A grande desvantagem é que só podem ser rodadas consultas “prontas”, ou seja, qualquer pergunta diferente das propostas ficará sem resposta até que seja desenvolvido um programa que realize aquela consulta.

Por exemplo, suponha-se que tenha sido desenvolvido uma consulta que responde à questão: “Quantos óbitos aconteceram no município do Rio de Janeiro em 2000?”. Mas a questão “Quantos óbitos aconteceram no município do Rio de Janeiro em 2000, por faixa etária?” ficará sem resposta até que seja desenvolvida um programa específico para esta consulta. Na solução OLAP, caso o consumidor da informação tenha o conhecimento específico para tal, ele pode rapidamente implementar esta consulta, caso ela não exista.

C) As aplicações para extração agendada de dados, que permitem que o consumidor da informação informe os dados que deseja consultar e deixe agendada a extração. A aplicação irá, na data/hora programada, efetuar a extração dos dados e disponibilizará, em uma página em *site* próprio, o *link internet* para que o consumidor da informação faça o *download* do arquivo resultante da extração.

Ou seja, neste último caso, o consumidor da informação irá interagir com uma aplicação em *site* na *Internet*, onde irá informar, através de um *menu*, por exemplo, que dados ele precisa, em que tabelas estão, os períodos de tempo e demais atributos que deseja de cada uma delas. Uma aplicação rodando em *background* (ou seja, rodando como um

processo transparente, que não interage com o usuário e sobre o qual o usuário não pode intervir) irá coletar estas informações e gerar, não necessariamente naquele momento, mas obedecendo a uma fila de requisições, uma página de resultados contendo os *links* para que o consumidor faça o *download* dos arquivos criados.

De posse dos arquivos, ele irá fazer suas tabulações localmente em seu computador, com a ferramenta que ele preferir, como, por exemplo, o TABWIN.

Como se pode ver, são três soluções que não excluem, de modo algum, outra solução que possa ser pensada, e que não são excludentes, devendo, todas, serem oferecidas aos consumidores da informação.

Na verdade, o perfil de cada consumidor/usuário é que irá determinar qual a melhor solução a ser adotada. Para os que possuem uma boa infraestrutura, com uma boa rede de comunicação de dados, com recursos humanos capacitados em tecnologia da informação, a solução OLAP pode ser a melhor, até por apresentar maior flexibilidade para as consultas não previstas (*ad hoc*). Para os que não têm estes recursos, porém têm acesso à *Internet* sem grandes problemas de excesso de uso da rede, e com uma boa conectividade, a segunda ou a terceira opção são indicadas. E, para aqueles que possuem acesso restrito à *Internet*, com poucos computadores conectados, com muita concorrência no acesso, a terceira opção é a mais indicada.

Porém, a questão de que ferramenta usar não é a questão central, mas é precedida, conforme foi dito no início deste capítulo, de um problema maior.

Um *data warehouse*, ou um *data mart*, criado, é fonte de informações úteis, mas revela, algumas vezes, mais do que se desejaria mostrar. Nasce aqui um ponto fundamental na construção destes instrumentos de suporte à decisão em saúde e de disseminação de informações. O que pode ser divulgado e o que não pode?

No entendimento do autor deste trabalho, as informações são públicas, e não “governamentais”. Portanto, não deve haver nenhuma restrição ao acesso dos dados contidos nestes bancos.

Se há algum dado que não deve ser mostrado publicamente (por exemplo, a identificação do paciente nos casos de internação hospitalar), então ele nem sequer deveria ter sido carregado no *data mart* / *data warehouse*.

É o que diz Moraes<sup>7</sup>,

"Não se trata mais, apenas, de denunciar o que não é dito, não é divulgado; o desafio agora é tentar desvendar as diferentes maneiras de não dizer, mesmo que aparentemente dizendo: como e onde estão distribuídos os que podem e os que não podem falar; como se decide o que vai ser disseminado; que tipo de discurso é autorizado. Não existe um só, mas, muitos silêncios e discursos de verdade em luta, em conflitos que estão presentes nas disputas em torno das estratégias de disseminação."

Cabe, aqui, também a afirmação de Silva, Cruz e Melo<sup>36</sup>,

"O grande desafio para o setor é o de tornar essa informação acessível aos usuários do SUS de maneira universal. Ou seja, não basta divulgar os dados pura e simplesmente, é necessário apresentar a informação numa linguagem adequada, de forma a permitir que qualquer sujeito possa se apropriar do conhecimento."

Ou seja, é preciso ter cuidado, cuidado no sentido de atenção, de respeito, na hora de divulgar os dados, porque a disseminação numa linguagem não adequada, desconhecida, tira do consumidor a possibilidade de se apropriar do conhecimento.

A disponibilização da informação numa linguagem adequada e a sua devida contextualização são uma necessidade para a própria existência de instâncias de gestão da informação no processo decisório, tirando esta atribuição das áreas de Tecnologia da Informação.

Silva, Cruz e Melo<sup>36</sup> nos falam da dificuldade no processo de apropriação da informação em saúde pelos conselhos, que são instâncias legítimas de onde poderiam ser obtidos subsídios para a reflexão de que informação a população desejaria ter acesso, e de como a informação sem significado, sentido e contexto empobrece o potencial de uso da informação nas discussões sobre as políticas de saúde.

Esta discussão remete, ainda, à questão dos metadados e da disseminação da informação sobre a informação.

O que os autores acima escreveram aponta, também, para a necessidade de um sistema de ajuda nos *data marts* / *data warehouses*. Ou seja, a documentação sobre o que o *data mart* / *data warehouse* possuem de acervo, sobre como acessá-los, deve estar muito clara para o consumidor da informação, especialmente, como foi dito, considerando-se o fato de que estes consumidores são muito variados, estão em pontos espalhados pelo



planeta, em horários de acesso que varrem as vinte e quatro horas do dia. Além disso, o próprio conteúdo dos bancos de dados, ou seja, seu metadados, deve estar muito claro, para que o usuário/consumidor da informação não tenha dúvidas sobre a informação desejada.

*Data warehouses* contêm, como era de se esperar, uma quantidade imensa de informação, até porque o custo dos componentes de *hardware* permitiu que maiores níveis de detalhe e mais períodos de histórico fossem acumulados nos *data warehouses*. Mesmo com todo o cuidado na hora de montar seus *data marts*, é preciso cautela com a “*overdose*” de informação sobre o consumidor da informação.

Segundo Branco<sup>18</sup>,

“Prover informação foi tornando-se sinônimo de proporcionar pilhas de dados, relevantes na perspectiva de quem os produzia, de forma fácil e ágil. O fato de a tecnologia ter propiciado coletar, armazenar e processar dados num volume jamais alcançável até então levou ao equívoco de acreditar que quanto mais dados e informações houver sobre um fenômeno ou problema, mais facilmente se poderia entendê-lo ou encontrar soluções.”

O que há de importante por trás destas mensagens é o que Moraes<sup>7</sup> retrata como o objetivo: “ter a informação necessária para tomar decisões menos empíricas e casuísticas, ou situar-se diante de cenários conflituosos para posicionar-se politicamente enquanto cidadão, ou inúmeras outras situações onde a informação seja necessária”.

## 8. Considerações Finais

Pelo título desta dissertação pode-se concluir que seu foco principal é a criação de um método de trabalho para construção de *data marts* e *data warehouses*.

No entanto, como foi visto, existe uma boa literatura que cobre com detalhes, maiores ou menores, este roteiro de construção de armazéns de dados (*data marts/data warehouses*).

O mais importante é caracterizar a necessidade do “estabelecimento de uma sinergia de competências, recursos e memórias”<sup>12</sup> com a “circulação inter/transdisciplinar do conhecimento”<sup>12</sup> e a “incorporação de métodos e técnicas de diferentes disciplinas que se complementam”<sup>12</sup> para a construção de um Ambiente de Informações para Apoio à Decisão<sup>12</sup>.

Pelo que foi apresentado neste trabalho, não há mais como sustentar que o centro das decisões esteja em uma área de Tecnologia de Informação, ficando os demais participantes em uma posição submissa, como se fossem os ignorantes do assunto, mais prontos a atrapalhar do que a serem co-participantes na criação da solução.

Por isso, o método de trabalho de construção de um armazém de dados apresentado aqui tem, exceto nas etapas cujos procedimentos envolvem saberes exclusivos do campo de conhecimento da Tecnologia da Informação, o foco permanente no usuário, ou consumidor da informação, desde o momento em que este começa a sua interação com o ambiente de suporte à decisão, até o momento em que ele extrai as informações que necessita.

Porém, é preciso ter em mente que implantar um método de trabalho com estes conceitos, e montar um grande *data warehouse* da saúde, com base nos campos de atuação do Ministério da Saúde, é a construção de uma catedral medieval. É uma obra de longa duração, que, embora não dure séculos como as catedrais, certamente irá atravessar governos e ministros.

Faz-se necessário, então, um planejamento de longa duração, republicano, que tenha como meta o bem comum da sociedade denominado informação em saúde.

Ao mesmo tempo em que os produtos resultantes deste trabalho poderão beneficiar atores nas diferentes esferas do SUS, ele poderá, se mal interpretado, parecer por em risco o *status* alcançado por outros atores nas diferentes áreas por ele abrangidas.

Como foi dito por Maquiavel<sup>42</sup>, em 1513, “não há nada mais difícil de administrar, ou de sucesso mais duvidoso, ou mais perigoso de se lidar do que tomar a dianteira na introdução de uma nova ordem de coisas. Porque o inovador tem inimigos em todos aqueles que estão se saindo bem sob a velha ordem, e tem apenas mornos defensores em todos aqueles que podem se sair bem sob a nova ordem”.

O próprio processo de construção destes armazéns de dados irá expor divergências na definição dos dados em suas diferentes fontes, que poderão variar de sua representação ao seu conteúdo, com chance de causar um tensionamento de poder entre as áreas responsáveis por estas fontes.

Do ponto de vista da Tecnologia da Informação, numa definição formal, *Business Intelligence* (BI) é “a reunião de diversos recursos para extrair, transformar, analisar grandes volumes de dados, produzindo conhecimento capaz de auxiliar a empresa a tomar decisões de negócios com mais garantia de sucesso”<sup>15</sup>. Evidentemente, o mesmo conceito pode ser aplicado às áreas governamentais, incluindo a área de saúde.

Muitas iniciativas têm sido feitas, com os nomes de BI (*Business Intelligence*) e *data warehouse*, que nada mais são do que utilizar as ferramentas de *data warehouse* para a tabulação dos dados de determinados sistemas. São iniciativas pontuais, que, embora dêem uma roupagem nova ao tratamento dos dados de um sistema, são limitantes na abrangência da informação.

Pulverizar *data marts*, correspondentes aos sistemas de informação que lhes servem de fonte de dados, é repetir a lógica fragmentadora analisada neste trabalho; é cometer os mesmos erros de falta de integração de dados e falta de padronização. É bom lembrar que por trás da facilidade de utilizar as ferramentas OLAP e de construir *data warehouses* / *data marts* para produzir soluções pontuais, pode existir a motivação política, ansiosa por soluções imediatistas.

Uma última pergunta pode ter ficado na mente do leitor: Por que o *data warehouse*? Por que esta tecnologia pode ser considerada melhor que outras para colocar o foco no consumidor? Não existem outras tecnologias que poderiam atender melhor à disseminação de informações com foco no consumidor da informação? Seria esta mais uma “onda tecnológica”?

A resposta é, no entender do autor deste trabalho, que não se deve excluir opções de disseminação, e quanto mais formas existirem, mais alternativas terá o consumidor da informação para obter o que lhe interessa.

A vantagem da tecnologia de data mart / data warehouse é permitir que os dados oriundos de diferentes fontes possam ser armazenados e consolidados em um repositório único, com uma estrutura que permite facilmente realizar agregações e detalhamentos, sob diferentes visões (dimensões). Assim, o data warehouse irá atuar como um “concentrador” de dados para disseminação de informações, utilizando uma linguagem única, facilmente assimilável pelos consumidores de informação.

Desta forma, pode-se argumentar que um determinado município, que tem acesso imediato às informações que produz, não teria necessidade de esperar que elas fossem carregadas no data warehouse para poder utilizá-las. Isso é verdade. Porém, desta forma, ele só teria acesso às suas informações, e se quisesse, por exemplo, comparar suas informações com as de outros municípios de seu estado, poderia, facilmente, fazer isto através do data warehouse.

Deve-se ter em mente que, além das pressões políticas e das ondas de tecnologia, existe uma meta a ser atingida: a informação de qualidade, útil, disponibilizada em tempo hábil, em linguagem clara e para todos.

Citando Davenport<sup>20</sup>: “Nosso fascínio pela tecnologia nos fez esquecer o objetivo principal da informação: informar. Todos os computadores do mundo de nada servirão se seus usuários não estiverem interessados na informação que esses computadores podem gerar”.

## Referências bibliográficas

1. Gómez MNG, Chicanel M. A mudança de regimes de informação e as variações tecnológicas.
2. Frohmann B. Taking policy beyond Information Science: applying the actor network theory for connectedness: information, systems, people, organizations. Annual Conference Canadian Association for Information Science (Anais). 23. Edmond, Alberta;1995.
3. Moraes IHS. Informações em saúde – da prática fragmentada ao exercício da cidadania. Rio de Janeiro: Editora Hucitec;1994.
4. Chaves M. Saúde e sistemas. 2ª ed. Rio de Janeiro: Fundação Getúlio Vargas; 1978.
5. Brasil. Constituição da República Federativa do Brasil; 1998.
6. Ministério da Saúde. Política nacional de informação e informática em saúde (proposta versão 2.0). Brasília;2004[acesso em 09-11-2009]. Disponível em [http://w3.datasus.gov.br/APRESENTACAO/PoliticaInformacaoSaude29\\_03\\_2004.pdf](http://w3.datasus.gov.br/APRESENTACAO/PoliticaInformacaoSaude29_03_2004.pdf).
7. Moraes IHS. Política, tecnologia e informação em saúde – a utopia da emancipação. Rio de Janeiro: Editora Casa da Qualidade; 2002.
8. Foucault M. O nascimento da clínica. Rio de Janeiro: Editora Forense – Universitária;1980.
9. Foucault M. Microfísica do poder. Rio de Janeiro: Editora Graal;1982.
10. Batista EO. Sistemas de informação: o uso consciente da tecnologia para o gerenciamento. São Paulo: Editora Saraiva; 2004.
11. Levy P. As tecnologias da inteligência: o futuro do pensamento na era da informática. Rio de Janeiro: Ed 34;1993.

12. Vasconcellos MM, Moraes IHS, Cavalcante, MT. Política de saúde e potencialidades de uso das tecnologias de informação. Saúde em Debate. set/dez-2001.
13. Singh H. Data warehouse: conceitos, tecnologias, implementação e gerenciamento. Rio de Janeiro: Editora Makron Books;2001.
14. Kimball R, Reeves L, Ross M, Thornthwaite W. The data warehouse lifecycle toolkit. New York: John Wiley & Sons, Inc.;1998.
15. Leme Filho T. Business intelligence no Microsoft Excel. Rio de Janeiro: Editora Axcel Books do Brasil;2004.
16. Barbieri C. BI – business intelligence: modelagem & tecnologia. Rio de Janeiro: Editora Axcel Books do Brasil;2001.
17. Fundação para o Software Livre - Free Software Foundation [homepage na Internet]. Boston:Free Software Foundation.c1996-2000.[atualizada em 05-05-2005; acesso em 09-11-2009]. Disponível em <http://www.gnu.org/>. O que é software livre?[1 tela]. Disponível em <http://www.gnu.org/philosophy/free-sw.pt.html>.
18. Branco MAF. Informação e saúde – uma ciência e suas políticas em uma nova era. Rio de Janeiro: Editora Fiocruz;2006.
19. Japiassu H. O sonho transdisciplinar – e as razões da filosofia. Rio de Janeiro: Imago Editora;2006.
20. Davenport TH. Ecologia da informação: porque só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura; 1998.
21. Moraes IHS, Gómez MNG. Informação e informática em saúde: caleidoscópio contemporâneo da saúde. Ciência e Saúde Coletiva.2007;12(3):553-565.
22. Ribeiro P, Sophia DC, Grigório DA. Gestão governamental e sociedade: informação, tecnologia e produção científica. Ciência e Saúde Coletiva.2007;12(3):623-631.

23. Gómez MNG. Para uma reflexão epistemológica acerca da Ciência da Informação. *Perspect. cienc. inf.* jan.-jun./2001;6(1):5-18.
24. Kimball R, Ross M. *The data warehouse toolkit – second edition*. New York: John Wiley & Sons, Inc.;2002.
25. Bonsack F. ¿Es objetivable y matematizable la información? In:Guérault M(Org.) *El Concepto de Información em la Ciencia Contemporánea – Coloquios de Royaumont*. México:Siglo Veintiuno Ed.,1966.p.223-242.
26. McGarry KJ. *The changing context of information: an introductory analysis*. Londres:Clive Bingley, 1981.
27. Almeida Filho N. Intersetorialidade, transdisciplinaridade e saúde coletiva: atualizando um debate em aberto. *Revista de administração pública*. Nov./Dez. 2000.
28. Horne EE. Information technology and human needs. In: Debons A. & Cameron WJ (Eds.) *Perspectives in Information Science*. NATO Advanced Study Institutes Series. Serie E: Applied Science, 10:31-40. Massachusetts: Noordehoff-Leyden; 1975.
29. McGarry KJ. *The changing context of information: an introductory analysis*. Londres:Clive Bingley; 1981.
30. Menou MJ. The impact of information – I.Toward a research agenda for its definition and measurement. *Information Processing & Management*, 31(4):455-477;1995.
31. Platt W. *A produção de informações estratégicas*. Rio de Janeiro:Agir;1967.
32. Saracevic T (Ed.). *Introduction to information science*.Nova York, Londres:R.R.Bowker Company;1970.
33. Stroetmann KA. Information management for the 90's: a conceptual framework. *International Forum on Information and Documentation*, 18(2):9-14; abr.1993.

34. Taylor RS. Value-added processes in information systems. Norwood:Ablex Publishing Corporation;1986.
35. Gómez MNG de. O objeto de estudo da Ciência da Informação:paradoxos e desafios. *Ciência da Informação*, 19(2):117-122;jul.-dez.1990.
36. Silva AX, Cruz EA, Melo V. A importância estratégica da informação em saúde para o exercício do controle social. *Ciência & Saúde Coletiva*, 12(3):683-688;2007.
- 37 Ministério da Saúde. 9ª Conferência Nacional de Saúde: relatório final. Brasília:MS; 1992.
- 38 Ministério da Saúde. Descentralização das ações e serviços de saúde: a ousadia de cumprir e fazer cumprir a lei. Brasília:MS/SUS;1993.
39. Descentralização On Line [homepage na Internet]. [atualizada em 19-10-2009; acesso em 19-10-2009]. Disponível em <http://www.ensp.fiocruz.br/descentralizar/>
40. Machado CV. Direito universal, política nacional: O papel do Ministério da Saúde na política de saúde brasileira de 1990 a 2002[tese de doutorado]. Rio de Janeiro:Universidade do Estado do Rio de Janeiro. Instituto de Medicina Social;2005.
41. Ministério da Saúde, Organização Pan-Americana da Saúde, Fundação Oswaldo Cruz. A experiência brasileira em sistemas de informação em saúde - volume 2. Brasília:Editora do Ministério da Saúde;2009.
42. Maquiavel N. O príncipe. Rio de Janeiro: Bertrand Brasil;1991.