

# PROPOSTA DE METODOLOGIA PARA DESENVOLVIMENTO DE DATA WAREHOUSE VISLUMBRANDO ATINGIR METAS PARCIAIS

Fábio Câmara Araújo de Carvalho, Mestrando  
João Ernesto Escosteguy Castro, M. Eng.

LABSAD – Laboratório de Sistemas de Apoio à Decisão – UFSC/CTC/EPS - Caixa Postal 476 – Campus  
Universitário – Trindade - CEP-88040-900 - Florianópolis/SC – fbcamara@eps.ufsc.br, castro@eps.ufsc.br

Cristiano Hugo Cagnin, Mestrando  
Aline França de Abreu, PhD

IGTI – Instituto para Gestão da Tecnologia de Informação – UFSC/CTC/EPS - Caixa Postal 476 – Campus  
Universitário – Trindade - CEP-88040-900 - Florianópolis/SC – cagnin@eps.ufsc.br, aline@eps.ufsc.br

## Abstract

The objective of this work is to propose a model for creation of integrated databases (Data Warehouse) that assists to the technical requirements and the needs in skilled time.

For that, a common theoretical model will be described among several authors and what happens in the practice of introduction of data warehouse in the companies will also be described. The new model will be proposed. It will also be made critics for both cases, and comments about strategical information systems and the importance of human capital in enterprises.

**keywords:** data warehouse, strategical information systems, human capital.

**Área Temática:** Gestão da Informação

## Resumo

O objetivo deste trabalho é propor um modelo de criação de base de dados integrada (*Data Warehouse*) que atenda aos requisitos técnicos e às necessidades de resposta à tomada de decisão em tempo hábil.

Para isso, será descrito um modelo teórico comum entre diversos autores e o que ocorre na prática de implementação de *data warehouse* nas empresas. Também serão feitas críticas a ambas as abordagens, além de comentários sobre alguns sistemas de informação estratégicos e a importância do capital humano nas empresas.

Em seguida, serão descritos problemas que normalmente ocorrem com terceirizações e, posteriormente, será proposto o modelo em si.

## 1. Introdução

No cenário atual, as empresas vêm buscando novos mercados e vantagem competitiva através da adaptação de seus sistemas de informação para captação, tratamento, distribuição e disseminação de informação dentro e fora da mesma. As organizações também aproveitam de novas tecnologias para compartilhar o conhecimento entre todos que participam do processo de dinamizar o negócio, o capital humano da empresa.

A busca de novos mercados faz com que se utilize de sistemas de informação (SIs), que, na quase totalidade dos casos, são implementados de forma independente, para cada setor da empresa, para obter apoio na tomada de decisão e, conseqüentemente, vantagem competitiva.

A utilização de SIs independentes cria um aumento do volume de dados desagregados, que geram informações não confiáveis.

O problema é ainda maior quando empresas, no ambiente de mercado globalizado, incorporam outras, adquirindo, dessa forma, sistemas legados diferentes daqueles utilizados por elas.

Verifica-se então a crescente necessidade de busca de dados e informações integradas. Nesse contexto, é proposta uma base de dados trabalhando sobre sistemas operacionais heterogêneos. Essa base é conhecida como um grande armazém de dados chamado *data warehouse*.

Segundo Inmon (1997), "*data warehouse é um conjunto de dados baseado em assuntos, integrado, não-volátil, e variável em relação ao tempo, de apoio às decisões gerenciais*".

O *data warehouse*, segundo a maioria dos modelos teóricos propostos pelos autores (Inmon, 1997, Brackett, 1996, Harrison, 1998, Meyer, 1998, dentre outros), utiliza as bases de dados do nível operacional para construir um sistema de dados históricos em forma bruta ou razoavelmente resumidos.

Essa nova filosofia de armazenamento de dados vem sendo proposta para reunir as informações contidas nas diversas bases de dados não-integradas existentes nas organizações que geram relatórios imprecisos, dados redundantes e desconexos, dentre outros, que proporcionam uma ineficiência global das ferramentas de suporte à tomada de decisão.

Os principais clientes de um *data warehouse* são aqueles que tomam ou que auxiliam na tomada de decisão. Dentre esses, pode-se citar: gerentes, de um modo geral, e analistas de SAD (Sistemas de Apoio à Decisão).

Atualmente sistemas como EIS (Sistemas de Informações Executivas) e ferramentas como OLAP (Processamento Analítico On-Line), além de *data marts* - que surgiram com o advento da filosofia de *data warehouse* -, estão buscando dados diretamente do ambiente operacional; gerando assim uma visão limitada à abrangência das bases nas quais se buscam os dados. Dessa forma, o *data warehouse* é proposto para proporcionar uma abrangência global e integrada para esses sistemas, ferramentas e bases específicas departamentais.

Nesse cenário, o *data warehouse* é, literalmente, um grande armazém organizado que fornece dados e informações aos sistemas gerenciais. Além disso, há um vislumbamento de democratização das informações gerenciais departamentais, e um achatamento e mistura dos sistemas classificados como de apoio à decisões hierarquicamente estruturados (ESS, MIS, DSS, KWS, OAS, e TPS), segundo Laudon e Laudon (1996, p. 19).

## 2. Descrição do Modelo Teórico

No modelo teórico, entendido como comum entre os diversos autores citados, o *data warehouse* é construído tendo como origem um nível chamado operacional (OLTP – *On-Line Transaction Processing*), onde se encontram as fontes de dados (planilhas eletrônicas, documentos, dados transacionais de produção, marketing, recursos humanos, finanças, contabilidade e bases externas).

Meyer (1998) apresenta uma figura que mostra uma solução de arquitetura de dados proposta pela IBM.

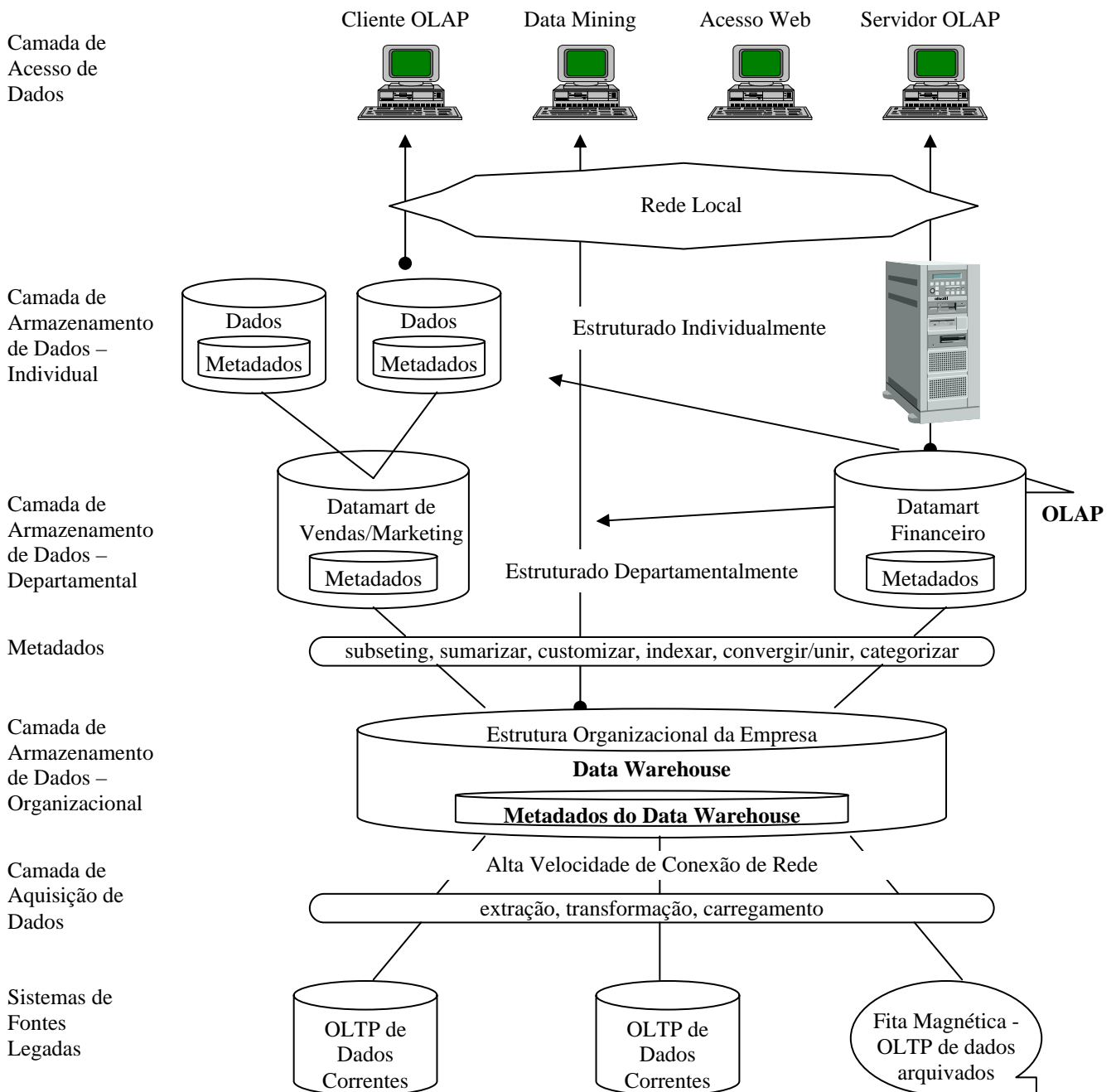


Figura 1: solução de arquitetura de dados proposta pela IBM, Meyer (1998)

No nível dos sistemas de fontes legadas, que se denominou também de operacional, encontra-se tudo o que se produz de dados, tanto históricos desconexos (arquivos em papel ou eletrônicos), quanto dados que suportam as operações diárias das empresas encontradas nas diversas bases de dados heterogêneas da organização.

Na camada de aquisição de dados, os *softwares* gerenciadores de *data warehousing* realizam a extração, a limpeza, a agregação, a sumarização, a consolidação e o carregamento dos dados constantes nas diversas bases operacionais para uma única base integrada.

Na camada de armazenamento de dados organizacionais, encontra-se a dita base integrada: o *data warehouse*. Também se encontra aqui a ferramenta chamada de Metadados, que pode ser comparado a um dicionário e um mapa de dados, pois através dele se pode saber quais os dados existentes, o que eles significam, onde estão alocados e como acessá-los. Assim, Metadados pode ser visto como dados sobre dados.

A partir dessa camada, as bases são derivadas da anterior. Na camada de armazenamento de dados departamentais, nos chamados *data marts*, encontram-se bases específicas e com informações úteis para a tomada de decisão gerencial de interesse dos departamentos.

Entram em cena ferramentas denominadas OLAP (*On-Line Analytical Processing*). Segundo Brackett (1996), OLAP é “o processamento que suporta a análise da tendência e projeções do negócio. É também conhecido como processamento de suporte a decisão”.

Inmon (1997) aborda que a noção de tempo de resposta no acesso a informações no ambiente OLAP é bem diferente do OLTP. O tempo de resposta para o OLTP geralmente é medido em segundos e minutos, por isso é considerado um fator crítico nesse ambiente, pois, quando este fica ruim, os negócios são afetados de forma imediata. Como no OLAP não ocorre um relacionamento tão direto, esse fator deixa de ser crítico, no momento em que o tempo de resposta para esse ambiente pode ficar em torno de horas, e até em dias.

Enquanto que o OLTP trabalha com dados que movimentam o negócio em tempo real, o OLAP trabalha com dados históricos no sentido de gerar informações para analisar o negócio. O OLTP, então, passa a ter a função de alimentar uma base de dados, a qual o OLAP a utilizará para a “transformação do conteúdo em uma forma útil de informações que possam ser entregues a um grande número de usuários. Os documentos OLAP – representação de dados em relatórios e gráficos – são criados ‘dinamicamente’ – o aspecto ‘on-line’ do OLAP – para atender às necessidades de informação do usuário” (Harrison, 1998).

O OLAP, em geral, executa cinco funções básicas, que são: de interface, de consulta, de processo, formato e exibição. (Harrison, 1998).

A camada de armazenamento de dados individual, pode ser considerada como um extrato do *data mart* departamental para atender à necessidade de setores departamentais. Figuram nessa camada ferramentas OLAP de Metadados derivadas do *data warehouse*.

A camada de acesso de dados pode constar de inúmeras ferramentas de apoio a decisão, entre elas: Cliente/Servidor OLAP, *Data Mining*, Cliente/Servidor Web, EIS (Sistemas de Informações Executivas) e ferramentas SAD (Sistemas de Apoio à Decisão) de um modo geral.

O acesso via OLAP pode ser realizado através das mesmas aplicações que são utilizadas para os *data marts*. O acesso via Web pode ser feito por meio de navegadores de hipertexto (*browsers*) ou aplicações específicas integradas ao *Browser*. A vantagem dessa última é a interface e a rapidez, quando se possui um *data warehouse* ou *data mart* para Internet/Intranet.

Enquanto se projetam armazéns de dados históricos (*Data Warehouse*) para fornecer a memória da empresa, o *Data Mining* (mineração de dados) explora e analisa essa memória para produzir conhecimento, ou seja, a inteligência artificial se aproximando da necessidade empresarial. Isso é feito através de meios automáticos ou semi-automáticos.

Empregam-se técnicas de estatística, da ciência da computação e de Inteligência Artificial para realizar tarefas como classificação, estimativas, previsões, agrupamento por afinidade, reunião e descrição.

Alguns fatores para se implementar o *data mining* são: 1. Os dados estão sendo produzidos; 2. Os dados estão sendo armazenados; 3. O poder da computação está disponível a preços acessíveis; 4. A pressão competitiva é forte e 5. Os *softwares* de *data mining* comerciais tornaram-se disponíveis. (Harrison, 1998).

Dentre as técnicas utilizadas, cita-se: 1. Análise de seleção estatística; 2. CBR (raciocínio baseado em casos); 3. Algoritmos genéticos; 4. Detecção de agrupamentos; 5. Análise de vínculos; 6. Árvores de decisão e indução de regras; 7. Redes neurais artificiais e 8. Visualização.

As ferramentas de EIS, que não estão contempladas na Figura 1, são as mais antigas existentes, mesmo antes de se pensar em base de dados integrada. Antes disso a obtenção de informações se dava diretamente do ambiente operacional. Com o advento do *data warehouse* essa ferramenta ganhou o poder de proporcionar informações mais abrangentes e confiáveis.

Tanto o acesso via Web quanto o EIS podem utilizar de *data marts* e OLAP. Porém, o *data mining* extrai informações diretamente do *data warehouse* pela necessidade do mesmo fazer uma varredura (*scanning*) em toda base de dados, podendo assim obter informações inesperadas.

Dessa forma, a criação desse cenário vislumbrado na Figura 1 se dá de forma sequencial, teoricamente de baixo para cima, e a previsão de duração para conclusão é, segundo Kimball (1996), para grandes empresas, em média, de 3 anos.

### 3. Descrição do que Ocorre na Prática

Na realidade da grande maioria das organizações, torna-se imprescindível respostas e retorno a curto e médio prazos. Assim, poucas são as empresas que se aventuram num projeto de implementação de *data warehouse*, como é proposto teoricamente, para posteriormente poderem obter informações/resultados não apenas de apoio à tomada de decisões como também de descobertas de novas tendências.

O dinamismo do atual mercado junto com a necessidade de constantes informações para que as empresas possam, no mínimo, se manter nesta acirrada competição, fazem com que as mesmas busquem qualquer informação (desconexa ou não) que as permita tentar entrar na disputa por novos clientes e na manutenção daqueles considerados como tal.

A primeira aplicação utilizando da filosofia OLAP como forma de obtenção de informações de cunho gerencial (de apoio à tomada de decisão) foi desenvolvida para dar suporte a sistemas de informações executivas – EIS.

Um exemplo desse tipo de aplicação é o *software* da empresa Pilot (Meyer, 1998, Groth, 1998). O mesmo realiza consultas a partir das bases operacionais, utiliza a filosofia OLAP e possui uma *interface* e técnicas de EIS.

No mercado existem *softwares*, soluções como a da Figura 1, da IBM. Porém, o custo e o tempo de desenvolvimento, além do tempo de retorno, fazem com que apenas empresas de grande porte e/ou de visão futura realizem uma implementação deste nível.

As soluções alternativas buscadas se aproximam do modelo do *software Lightship* (da Pilot), desenvolvendo uma espécie de *data mart* a partir das bases operacionais utilizando técnicas OLAP.

A idéia é inverter a seqüência de implementação teórica da Figura 1, realizando no sentido de cima para baixo, ou seja, partir das bases setoriais, *data marts* individuais e/ou departamentais, para convergir para uma base única; ao invés de estudar as bases operacionais, projetar, desenvolver o *data warehouse*.

O importante é obter retornos em áreas específicas em um tempo hábil, utilizando o que existe no mercado. Não há grandes preocupações e/ou tempo para construir uma base única e integrada, dando uma visão global da empresa. No entanto, é de vital importância possuir dados de toda empresa para se obter informações mais confiáveis e, consequentemente, um melhor posicionamento e/ou sobrevivência no mercado.

#### 4. Modelo Teórico x Realidade

Segundo Brackett (1996) “o desafio das organizações é parar de criar dados desconexos (dados essencialmente diferentes e que não comportam qualquer comparação) para se criar e transformar estes em dados integrados que darão suporte e atenderão a demanda de informações presente e futura. Assim, deve-se desenvolver bons metadados para que as pessoas saibam quais os dados existentes, o que eles significam, aonde estão alocados e como acessá-los”.

No entanto, segundo este mesmo autor, “o real desafio, contudo, é aprender a sobreviver com dados desconexos enquanto se desenvolve dados (base de dados - *data warehouse*) integrados”.

Os modelos teóricos dos diversos autores têm a filosofia de que se deve construir um *data warehouse* para se obter informação integrada, assim as pessoas têm de “se virar” para conviver com o caos existente pela não-integração durante os anos em que se implementa essa filosofia. Além disso, não propõem alternativas para se extrair resultados/retornos parciais, de curto/médio prazo.

Por outro lado, do ponto de vista mercadológico, as empresas que se dizem provedoras de “soluções” corporativas, fornecem produtos que, mesmo funcionando, só servem para atender às necessidades mínimas e imediatas das organizações.

Confrontando a teoria e a prática, um personagem em comum, que é fator crítico de sucesso, é o chamado Metadados. Na teoria, todos os autores ressaltam a importância dos mesmos para que o *data warehouse* e os *data marts* funcionem de forma efetiva.

Na prática, os metadados ainda estão sendo deixados de lado ou mesmo sendo ignorados em sua importância, e não estão sendo implementados com as aplicações mais comercializadas.

Pelo que se observa, dificilmente as organizações que optam pelo modelo prático de construção de um *data warehouse*, a partir dos *data marts*, conseguirão sucesso sem um criterioso planejamento e uma boa ferramenta de metadados, para que no momento em que se vá fazer a união das bases de dados, não ocorra o problema inicial, de redundância, desconexão de dados, etc.

Ainda, nesse contexto, como se vai fazer *data mining* em *data marts*? Mesmo que a partir do *data marts* se desenvolva um *data warehouse*, haverá ou não necessidade de se realizar *data mining* nos *data marts*?

Da mesma forma que se trabalha com EIS baseado em algumas bases operacionais, com um relativo sucesso, pode-se também realizar um *data mining* baseado em diversas bases departamentais, com um possível sucesso. Porém, a filosofia do *data mining* está em procurar por relações impensadas e situações não previstas, necessitando, dessa maneira, do maior número de informações possível. Portanto, é imprescindível que se tenha um *data warehouse*.

Pode-se até fazer um *data mining* em bases departamentais, mas é muito mais complicado, tecnicamente, implementá-lo. Além disso, é menos confiável, menos efetivo e menos produtivo, do ponto de vista de não aproveitar o potencial existente em uma base única e integrada, com o maior número de informações possíveis.

## 5. Proposta de Modelo

Entendendo que o modelo teórico é o ideal para se buscar uma base de dados integrada, segura e confiável, a proposta que será feita tem como base a teoria comum dos autores já comentados.

No entanto, o fator chave que diferencia este modelo do teórico, é a possibilidade de obtenção de resultados parciais a curto prazo.

Na maioria das empresas, os gerentes da alta administração não costumam realizar investimentos em projetos que demandem grande quantidade de tempo e recursos e com Retorno Sobre Investimento (ROI) demorado.

Quando isso ocorre, essa demora generalizada de retorno e de resultados, surge um clima organizacional de desconfiança e descrença da tecnologia.

Além disso, pelo que já foi descrito do modelo proposto na prática, conclui-se que não há nada que comprove que o mesmo irá convergir para uma base única integrada. Então o empresário acaba investindo em tecnologias que não irão chegar onde as mesmas “prometem”. O que deveria ser solução, torna-se, então, mais um componente problemático e agravante das chamadas “ilhas de informação”, com dados e informações desencontradas, perdidas e que não chegam às pessoas certas, no lugar certo, no momento certo.

Assim, será proposto um modelo baseado na Figura 2 a seguir:

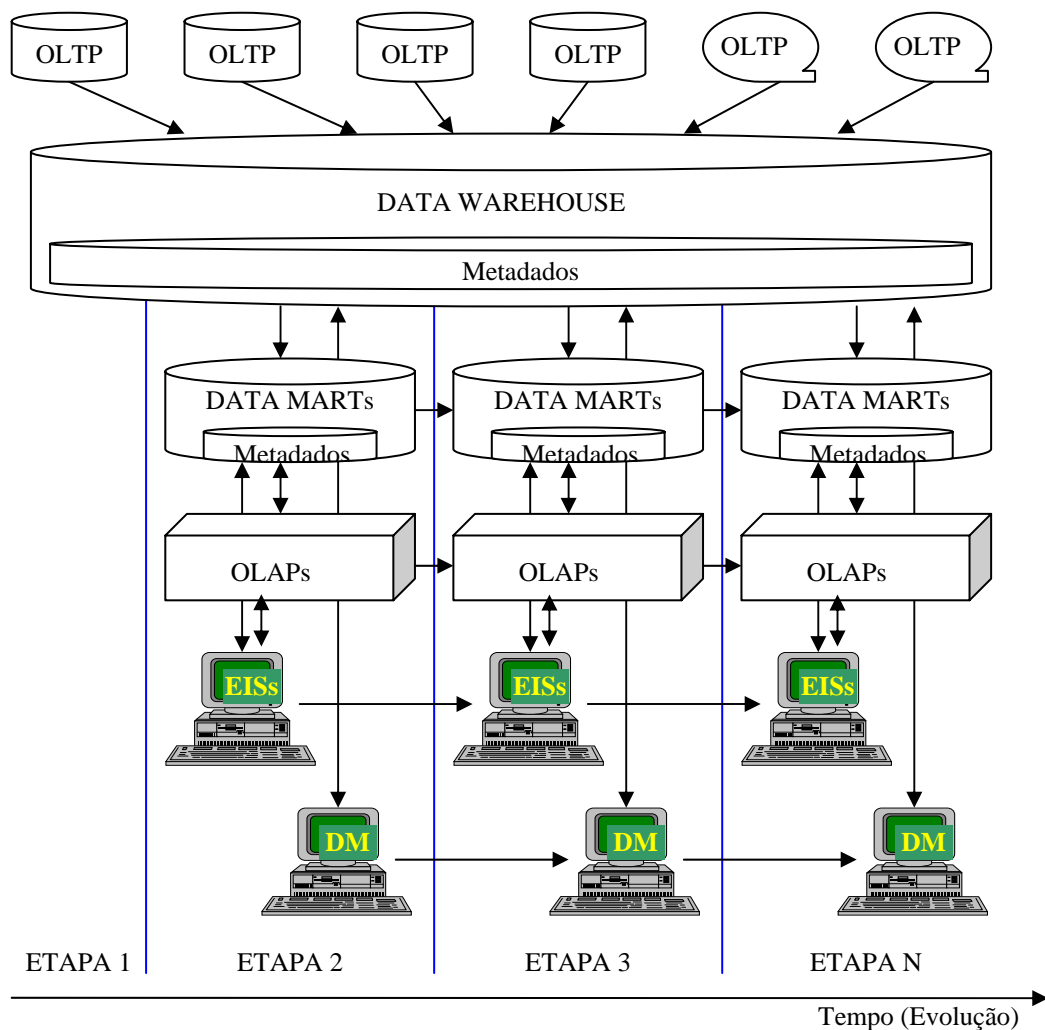


Figura 2: Modelo de Desenvolvimento Proposto

A principal idéia proposta por este modelo (Figura 2) é a existência de uma simultaneidade no desenvolvimento tanto do *data warehouse* como o de ferramentas OLAP, EIS e *data mining*. Isso se daria através de um planejamento minucioso objetivando focar em áreas específicas, desenvolvendo assim *data marts* gradativos, de modo que se possa, ao fim de cada etapa, evoluir em termos de complexidade e volume de informações as ferramentas descritas na Figura 2. Cada uma dessas etapas seriam estabelecidas num planejamento integrado, com metas parciais, que, após serem atingidas, seriam avaliadas, redimensionadas e realimentadas para redefinição da próxima etapa.

Assim, num primeiro momento seria realizado o planejamento integrado do *data warehouse*, juntamente com os *data marts* e as ferramentas OLAP, EIS e *data mining*. Nesse planejamento seriam estabelecidas metas parciais a serem atingidas.

Num segundo momento, após o planejamento integrado, inicia-se o desenvolvimento do *data warehouse*, que vai ser desenvolvido de forma evolutiva.

Posteriormente, a partir de atingido os objetivos da primeira etapa, inicia-se a povoação dos *data marts*, juntamente com o início do desenvolvimento das ferramentas OLAP, EIS e *data mining*. Note que o *data mining* vai sendo realizado diretamente no *data warehouse*.

A partir daí, todos os sistemas evoluem em conjunto, cumprindo metas parciais, produzindo resultados imediatos a curto prazo; resultados esses em forma de aumento produção, aumento de vendas, melhorias de *marketing*, etc., dependendo do foco inicial dado pelo planejamento integrado e da efetividade com que essas informações serviram às pessoas que tomam decisão na empresa.

Da forma como a teoria propõe, os sistemas são desenvolvidos numa seqüência vertical (Figura 1). No modelo aqui proposto, o desenvolvimento seria de modo também sequencial, mas horizontal, se comparado com a mesma figura. Isso faz com que a complexidade do desenvolvimento da forma como é proposto na teoria seja distribuída no decorrer das etapas planejadas, com a diferença de que já a partir da segunda etapa de desenvolvimento, as ferramentas de apoio à tomada de decisão podem estar atuando.

Não é uma tarefa simples seguir este modelo. É preciso muito raciocínio e conhecimento dos dados e das informações da empresa e um bom planejamento. Deve haver pessoas na equipe que se dediquem a desenvolver suas criatividade e conhecimentos, que não são contemplados no ambiente de trabalho atual, onde torna-se cada vez mais necessário um modelo de gestão do conhecimento associado às pessoas que irão implementar este novo modelo de desenvolvimento de sistemas.

A complexidade do modelo teórico existente se dá pelo fator tempo que é requerido em grande quantidade. Já para esse novo modelo a complexidade se dá no conhecimento, e não no tempo, necessário para seu desenvolvimento.

Ou seja, a empresa que souber gerenciar o conhecimento de forma efetiva desenvolverá sistemas mais rapidamente e com resultados parciais que podem trazer juntos retornos sobre o investimentos elevados a curto prazo.

### **6.1. A Questão do Capital Humano**

A empresa e as pessoas precisam assimilar a cultura de que essas tecnologias são melhor aproveitadas quando as pessoas que as utilizam se dedicam a evoluir o conhecimento acerca dos diversos atores institucionais e a realizar atividades com prazer e realização pessoal.

A gestão do conhecimento é necessária para que se tenha resultados com efetividade (eficiência e eficácia). Portanto, esta é um ponto crítico para se ter diferencial competitivo, e a inteligência é o fator principal para conquista desse diferencial, que passa a ser



comparada com o trabalho e o capital quando esses eram os diferenciais competitivos anteriores.

Dentro desse contexto, outro ponto importante é a aprendizagem contínua do capital humano, ou seja, as pessoas que produzem conhecimento dentro da organização. Com a aprendizagem contínua as pessoas aumentam a capacidade de criar o futuro, sentir as mudanças no cenário competitivo, criar inovações, extrair aprendizagem da experiência diária (dos sucessos e fracassos), e criar um clima que alimente a criatividade das pessoas.

Ainda dentro da gestão do conhecimento, o *benchmarking* deve ser visto como forma sistemática de aprender mais rápido, e a organização deve aprender a aprender.

Surge, então, um novo patamar além do dado, da informação e do conhecimento: o saber. De acordo com Santos (1998):

- **Dado** é um registro a respeito de um determinado evento (um sinal) para o sistema;
- **Informação** é um conjunto de dados com determinado significado para o sistema;
- **Conhecimento** é a informação que, devidamente tratada, muda o comportamento do sistema; e
- **Saber** é um conjunto de conhecimentos a respeito de um determinado tema utilizado para a resolução de problemas no sistema.

O conhecimento passa a ser um recurso primário para esse novo ambiente. A gestão desse recurso deve fazer com que as pessoas o trabalhem para gerar o saber, num processo contínuo, visto que a tendência vislumbrada é de uma mudança de cenário a cada curto intervalo de tempo, o chamado ciclo da obsolescência, onde as pessoas tem sempre que estar se reciclando para não ficarem com os conhecimentos ultrapassados.

Por fim, um ponto que merece destaque é a questão da importância do capital humano nesse novo cenário. É fundamental as empresas manterem essas pessoas em seus quadros, visto que os dados, a informação e o conhecimento são algo que podem ser compartilhados, porém o saber e a competência estão inerentes a cada indivíduo, e quando o mesmo sai da empresa todas essas qualidades inerentes vão juntos, e a empresa pode sofrer o impacto dessa perda.

## 6. Conclusões

Mudanças tecnológicas, principalmente quando é vislumbrada uma nova filosofia de organização do trabalho, e redesenho da informação e da empresa, requerem mudanças comportamentais como fator crítico de sucesso da tecnologia.

Mudanças de comportamento como um fenômeno organizacional tomam tempo e não são espontâneas. Requerem a definição de uma estratégia clara, que leve em consideração vários fatores, além de conscientização do usuário para a necessidade pelo engajamento nas novas filosofias de trabalho baseado no conhecimento e na tecnologia, como suporte, para a sobrevivência do negócio da empresa e descoberta de novos mercados.

Uma organização para atingir seu objetivo deve dar “sinais” claros da importância do “capital humano” que ela detém; deve buscar a satisfação dos seus colaboradores com o ambiente no qual trabalham.

Para lidar com o fator tempo (mudanças de comportamento levam tempo para se efetivarem como um fenômeno organizacional homogêneo), é necessário que a organização estabeleça metas parciais no desenvolvimento da própria infraestrutura de informática de suporte aos negócios (*data warehouse, data marts, OLAP, EIS, data mining, etc...*).

Metas parciais indicam progresso, mostram resultados concretos e estabelecem uma relação de custo/benefício positiva, conquistam o usuário e o suporte da alta gerência, disseminando uma cultura voltada para a integração e padronização de dados.

Portanto, considera-se o modelo proposto adequado a atender às necessidades de otimização do tempo de implementação dessas tecnologias de suporte à decisão, visto que o fator tempo é substituído pelo fator conhecimento. A gestão do conhecimento, e não do tempo, torna-se o fator crítico, e sendo bem conduzida pode trazer retornos imediatos.

## 7. Referências

- BRACKETT, Michael H. **The Data Warehouse Challenge: Taming Data Chaos**. USA: Wiley, 1996. (ISBN 0-471-12744-2).
- GROTH, Robert. **Data Mining: A hands-on approach for business professionals**. New Jersey, USA: Prentice Hall, 1998.
- HARRISON, Thomas H. **Intranet Data Warehouse**. São Paulo: Berkeley, 1998. (ISBN 85-7251-460-0).
- INMON, W. H. **Como Construir o Data Warehouse**. Rio de Janeiro: Campus, 1997. (ISBN 0471-14161-5).
- KIMBALL, R. INMON, W. H. **The Data Warehouse Toolkit: practical techniques for building dimensional data warehouses**. USA: John Wiley & Sons, 1996. (ISBN 0-471-15337-0)
- LAUDON, K. C., LAUDON, J. P. **Management Information Systems: Organization and Technology**. New Jersey: Prentice Hall, 1996, 4ª edição. (ISBN 0-13-213778-X)
- MEYER, Don, CANNON, Casey. **Building a Better Data Warehouse**. NJ/USA: Prentice Hall, 1998. (ISBN 0-13-890757-9).
- SANTOS, Neri dos. **Notas de Aula**. Disciplina: Introdução à Inteligência Competitiva. IN: <http://www.eps.ufsc.br/~neri>. EPS/UFSC: Florianópolis-SC, 1998.
- WATSON, Hugh J., RAINER Jr, R. Kelly, HOUDESHIEL, George. **Building Executive Information Systems and other Decision Support Applications**. USA: John Wiley & Sons, 1997. (ISBN 0-471-06930-2)