

Correspondência para/  
Correspondência para/  
Correspondence to  
Dorzeli S. Trzeciak  
Av. Lédio J. Martins,  
999/21A - Bairro Kobrasol,  
São José, SC.  
CEP. 88122-000  
e-mail: dorzeli@eps.ufsc.br

Artigo recebido em: 18/03/03  
Aprovado em: 16/04/03

# MODELAGEM DA INFORMAÇÃO: UMA ABORDAGEM VOLTADA À INTEGRAÇÃO DE DADOS

## MODELADO DE LA INFORMACIÓN: UN ABORDAJE DIRIGIDO A LA INTEGRACIÓN DE DATOS

### INFORMATION MODELING: A DATA INTEGRATION APPROACH

*Dorzeli S. Trzeciak, M.Sc.*

Pesquisadora do IGT1/EPS/UFSC  
e-mail: dorzeli@eps.ufsc.br

*Manoel Agrasso Neto, M.Sc.*

Doutorando em Engenharia  
de Produção - UFSC  
e-mail: manoel@eps.ufsc.br

*Mauro Pacheco Ferreira, M.Sc.*

Professor da Universidade do Sul  
de Santa Catarina - UNISUL  
e-mail: mauro@floripa.com.br

*Nilce Miranda Ayres, M.Sc.*

Professora da Universidade do  
Sul de Santa Catarina - UNISUL  
e-mail: nilce@eps.ufsc.br

*Pedro Felipe de Abreu, Ph.D.*

Consultor e Professor -  
Artfactu Tecnologias  
e-mail: aabreu@floripa.com.br

*Vinícius M. Kern, Dr.*

Professor da Universidade do  
Vale do Itajaí - UNIVALI  
e-mail: kern@eps.ufsc.br

**Palavras-chave:**  
banco de dados,  
modelagem de dados,  
sistemas de recuperação da  
informação - administração.

**Palabras clave:**  
banco de datos,  
modelado de datos,  
sistemas de recuperación  
de la información -  
administración.

**Key words:**  
database,  
data modeling,  
information retrieval  
systems - administration.

**RESUMO:** Este artigo trata da integração de informações nas organizações com respaldo teórico em ontologias e no tratamento de metadados, adotando uma orientação menos tecnológica e mais voltada à visão do negócio, e na aplicação prática desses conceitos. Uma das causas frequentes da falta de integração de dados nas organizações advém da diferenciação conceitual sobre dimensões-chaves do negócio entre áreas funcionais das organizações. Como consequência tem-se: duplicação de esforços, falta de visão compartilhada, informações conflitantes, dificuldade em reunir e consolidar informações para suportar a tomada de decisão. A solução técnica relacionada ao uso de gerenciadores de banco de dados (SGBD) ou a aplicação de técnicas de modelagem de dados durante o desenvolvimento de sistemas de informação não é suficiente para sanar ou evitar problemas de integração de dados. Sugere-se que, antecedendo a modelagem de dados dos sistemas de informação de uma organização, haja uma definição e disponibilização dos conceitos-chave da organização.

**RESUMEN:** Este artículo trata de la integración de informaciones en las organizaciones, con respaldo teórico en ontologías y en el tratamiento de metadatos, adoptando una orientación menos tecnológica y más dirigida a la visión del negocio y a la aplicación práctica de esos conceptos. Una de las causas frecuentes de la falta de integración de datos en las organizaciones adviene de la diferenciación conceptual sobre dimensiones clave del negocio entre áreas funcionales de las organizaciones. Como consecuencia existe: duplicación de esfuerzos, falta de visión compartida, informaciones encontradas, dificultades en reunir y consolidar informaciones para sustentar la toma de decisión. La solución técnica relacionada al uso de administradores de banco de datos (SGBD) o la aplicación de técnicas de modelado de datos durante el desarrollo de sistemas de información no es suficiente para sanar o evitar problemas de integración de datos. Se sugiere que, antecedendo al modelado de datos de los sistemas de información de una organización, haya una definición y disponibilidad de conceptos clave de la organización.

**ABSTRACT:** This article deals with the integration of information in organizations. It has a theoretical basis in ontologies and the treatment of metadata, adopting a position that is less technological and more focused on the business vision and on the practical application of these concepts. One of the frequent causes of the lack of data integration in organizations arises from the conceptual differentiation of key business dimensions among the functional areas of organizations. The consequences are: duplication of efforts, lack of a shared vision, conflicting information and a difficulty in combining and consolidating information to support decision-making. The technical solution relating to the use of database managers (SGBD) or the application of data modeling techniques during the development of information systems is not sufficient to resolve or avoid problems of data integration. It is suggested that prior to the data modeling of an organizations' information systems, the key-concepts of the organization are defined and made available.

## 1. INTRODUÇÃO

As organizações habitualmente projetam e desenvolvem seus sistemas de informação de modo a se enquadrarem nas práticas existentes, disso resultando que muitas das aplicações são funcionalmente orientadas. Conforme expressa Davenport (1994), sistemas de marketing resolvem problemas de marketing, sistemas de vendas resolvem problemas de vendas, sistemas de fabricação resolvem problemas de fabricação. Tais sistemas "fracionados" acabam por não proporcionar uma visão de conjunto da organização, aprisionando os dados dentro das funções.

Uma orientação voltada ao desenvolvimento de sistemas de informações independentes, sem considerar as necessidades mais abrangentes do negócio, tende a criar uma "teia" de informações conflitantes e de difícil integração no futuro. Esta orientação normalmente se traduz em variações de conteúdo e significado dos dados organizacionais, dificultando o compartilhamento de informações entre as áreas/processos organizacionais, e a reunião e consolidação de informações para suportar a tomada de decisão nas empresas (BRACKETT, 1996; INMON, 1997).

Observa-se que há uma centralização de enfoque na tecnologia para suprir tais deficiências. Mas o problema da integração de dados organizacionais é mais que uma solução técnica relacionada ao uso de um sistema gerenciador de banco de dados (SGBD) ou a aplicação de técnicas de modelagem de dados durante o desenvolvimento de sistemas de informação. A prática em desenvolvimento de sistemas e em administração de dados mostra que nem sempre o uso da modelagem de dados e a existência de uma tecnologia de bancos de dados impedem a falta de integração entre os dados de uma organização. As aplicações, não raro, acabam por serem desenvolvidas levando em conta a visão particular de uma área da empresa ou a necessidade de desempenho ou extração de informação de uma comunidade em particular dentro da organização.

Assim, antecedendo a modelagem de dados dos sistemas de informação de uma organização, propõe-se que haja uma definição e disponibilização de alguns conceitos-chave da organização. Optou-se por buscar um respaldo a essa questão nas ontologias comuns, onde há a formalização de conceitos-chave da organização, e numa atenção maior ao tratamento dos metadados, numa orientação menos tecnológica e mais voltada à visão do negócio. Seguindo esta visão, este artigo primeiramente analisa alguns aspectos conceituais sobre integração de dados, ontologias e metadados. Num segundo momento, apresenta a consolidação de experiências na busca pelo compartilhamento de uma visão comum de dados, com o intuito de facilitar a integração de dados tratados pelos sistemas de informação, vivenciadas em empresas. Ao final, são apresentadas conclusões referentes ao tema em questão.

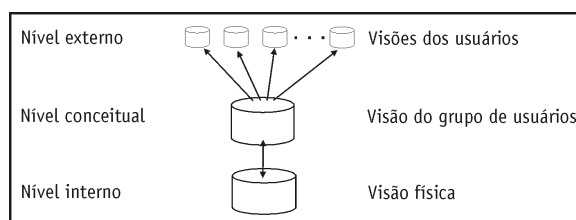
## 2. INTEGRAÇÃO DE DADOS NAS ORGANIZAÇÕES

Segundo Nolan (apud AYRES, 2000), a integração dos dados ocorre dentro das organizações num processo evolutivo. De acordo com o amadurecimento das empresas no processamento de dados, verifica-se que há uma evolução na forma de planejamento e controle dos dados e aplicações, buscando sua integração com os processos e estratégias organizacionais. De acordo com o autor supra-citado, esta evolução se dá em seis estágios, descritos a seguir.

- **Iniciação:** fase da introdução dos computadores na organização. Fase de crescimento lento e com ênfase na aprendizagem da tecnologia. As aplicações que são desenvolvidas são simples, geralmente de natureza contábil ou financeira.
- **Expansão:** fase da proliferação de aplicações na organização, com substancial perda de controle por parte da gestão.
- **Controle:** fase de controle por parte da gestão, com ênfase na documentação das aplicações existentes e no planejamento formal de sistemas de informação. As aplicações deixam de ter um foco estritamente operacional, passando a atender também aspectos gerenciais.
- **Integração:** fase caracterizada pela integração das aplicações existentes, originalmente concebidas de forma independente, tirando assim proveito das tecnologias de bancos de dados e telecomunicações.
- **Administração de dados:** fase em que a tecnologia de bancos de dados já está consolidada, surgindo a função de administração com o intuito de planejar e controlar o uso dos dados. A ênfase está em sistemas integrados, a partir do compartilhamento de dados e aplicações comuns.
- **Maturidade:** fase em que as aplicações são integradas e espelham o fluxo de informações da organização. A tecnologia da informação está alinhada com os processos gerenciais da organização e os recursos de dados são planejados a atender as necessidades estratégicas.

Do ponto de vista estritamente tecnológico, a integração de dados está intimamente associada à integração de sistemas de informação. A arquitetura ANSI/SPARC (TSICHRITZIS; KLUG, 1978), ilustrada na figura 1 e amplamente adotada na construção dos bancos de dados utilizados pelos sistemas de informação, permite integrar aplicações, mas não inibe o problema da integração de dados.

Figura 1 - Arquitetura ANSI/SPARC



Fonte: Adaptada de Tsichritzis e Klug (1978)

De acordo com Tsichritzis e Klug (1978), essa arquitetura preconiza que um banco de dados pode ser entendido a partir de três perspectivas:

- a visão dos usuários, ou nível externo do banco de dados, que reúne as várias concepções que os usuários podem fazer sobre o que é o banco de dados. Estas visões são **tendenciosas** por definição;
- a visão do grupo de usuários, ou nível conceitual, que é um mapeamento das visões de usuários para um único documento, compatível com todas as visões de usuário;
- a visão física, ou nível interno, que trata do banco de dados como é implementado no computador, com todas as características de gerência de banco de dados.

A visão do grupo de usuários é expressa em documentos como modelos ou diagramas de dados, ou em *scripts* de definição de dados. Estes documentos expressam a organização de um banco de dados, ou dados sobre os dados, ou **metadados**. Os metadados devem contemplar todas as necessidades de informação dos usuários e compatibilizar a terminologia e significados, o que pode ser feito através de ontologias. Tanto os metadados quanto as ontologias são descritos mais adiante neste artigo.

## 2.1 EVOLUÇÃO DA TECNOLOGIA DA INFORMAÇÃO NAS ORGANIZAÇÕES

O surgimento do problema relacionado à integração de dados está, em boa parte, associado ao próprio processo histórico-evolutivo da tecnologia da informação, que praticamente a cada 10 ou 15 anos provoca uma revolução (INPRISE, 2001).

No início da era da informática, nos anos 50, até o final da década de 70, presenciou-se o aparecimento de sistemas de informação formados por programas e dados fortemente independentes. Cada uma das aplicações compunha um sistema de informação fechado, centralizado, e continham toda uma lógica de apresentação, de negócio e de acesso a dados particulares, desenvolvendo uma cópia particular dos seus dados. Assim, apesar de diferentes sistemas utilizarem as mesmas fontes de dados, as organizações acabavam por armazenar cópias redundantes dos dados em múltiplos sistemas.

Em torno de 1970 surgiu uma nova tecnologia de armazenamento e acesso a dados - o disco ou DASD (*Direct Access Storage Device*, ou dispositivo de armazenamento de acesso direto). E com ele, um novo tipo de *software* conhecido como SGBD (Sistema Gerenciador de Bancos de Dados), que trazia o conceito de "banco de dados" definido por Martin (1976) como um conjunto de dados inter-relacionados, de redundância controlada, servindo a uma ou mais aplicações. O uso dos SGBD se constituiu muito mais no aproveitamento de um método de armazenamento e recuperação de dados mais poderoso, do que efetivamente

num repositório único de dados compartilháveis.

No início da década de 1980, novas tecnologias, como os PCs (*Personal Computer* ou computadores pessoais), as redes de grupos, as ferramentas de quarta geração começaram a despontar, colocando nas mãos dos usuários finais a possibilidade de controlar dados e aplicações (*end user computing*). Esta nova "onda" veio agravar o quadro de redundância de dados, conflito de resultados e conhecimento dos dados (INMON, 1997; LAUDON; LAUDON, 1996).

## 2.2 O PROBLEMA DA INTEGRAÇÃO DE DADOS

Este histórico tecnológico fortalece o aparecimento de uma lacuna na concepção e construção dos sistemas de informação, onde os sistemas são projetados sem levar em conta a definição das necessidades interfuncionais, e sim com visão específica de nível mais baixo, representando a unidade de análise ou de negócio em que será construído.

Existem vários métodos, na literatura, para ajudar a estruturar a definição dessas ligações entre os sistemas. Um exemplo é a metodologia BSP (*Business System Planning*), desenvolvida pela IBM, e que foi amplamente aplicada em diversas grandes empresas. A aplicação desses métodos, no entanto, não é simples, envolvendo muitas pessoas, tempo e dinheiro. Os resultados, comumente, são interrompidos pela falta de visão do objetivo a alcançar com tal dispêndio, por parte dos usuários. Outras vezes são abandonados já no processo de desenvolvimento das aplicações especificadas, pois as aplicações definidas já não mais traduzem a realidade vigente na organização, ou existe uma falta de entendimento do que foi definido, em função do processo estar sendo conduzido por pessoas diferentes daquelas que originalmente fizeram a especificação (MCGEE; PRUSAK, 1994; DAVENPORT, 1998).

O uso de técnicas e metodologias de modelagem da informação, a identificação das necessidades de informação no projeto de bancos de dados, não introduz automaticamente uma visão de dados integrados. A modelagem da informação se preocupa com "qual" dado é necessário e seus relacionamentos com outros, expressos num modelo de dados. Observa-se que esta técnica define com clareza as entidades de dados, os seus relacionamentos e algumas regras do negócio, mas não garante a integração dos dados, pois o foco normalmente é uma área ou processo específico.

## 3. ONTOLOGIAS PARA MODELAGEM DA INFORMAÇÃO

Ontologia é um termo derivado e emprestado da Filosofia e embora o interesse nele esteja aumentando, é de difícil definição e não é utilizado da mesma maneira por diferentes pessoas. Ontologia tem sido usado em Ciência da Computação com o significado de uma "especificação explícita de uma

conceitualização" (GRUBER 1994). Isto implica que, para construir uma ontologia em um certo domínio, é necessário definir de forma explícita, abrangente e não-ambígua os conceitos importantes no domínio.

A Filosofia define ontologia como sendo o estudo da realidade e dos objetos independente do conhecimento que se possui deles (CHERMONT, 1999; KERN, 1997). Nesse contexto, ontologia expressa algum modo de visão do mundo, ou seja, uma perspectiva, um modo de pensar sobre algum aspecto do mundo natural ou artificial. Pode ser encarada como um paradigma que utiliza uma série de conceitos e definições. Mas que visão de mundo é essa?

Problemas ocorrem quando termos similares são usados com significados diferentes ou semelhantes, podendo causar ambigüidade e imprecisão. Devido às ambigüidades, faz-se necessário definir e explicitar essa visão do mundo que o modelo de negócio está baseado.

Uma ontologia deve atuar como um meio que permita a comunicação entre diferentes pessoas através da organização; pessoas e sistemas computacionais implementados e diferentes sistemas computacionais. Nesse sentido, a abordagem de **ontologia** tem agregado o adjetivo **comum**. Ontologias comuns são fundamentais para a integração de sub-sistemas.

É importante também que a ontologia abranja a aquisição, representação e manipulação do conhecimento do negócio que está sendo modelado. Por esse motivo, não deve ser considerada estática, mas sempre ser revista, ampliada e corrigida.

Uma ontologia deve ser construída de forma a maximizar sua futura reutilização e aumento de escopo, que pode ser atingido pelo balanceamento correto entre ser específico o suficiente para a execução das tarefas, mas não tão específico que não seja de utilidade para outros. Esta visão proporcionada pela ontologia delinea a necessidade de uma visão conceitual, de espectro organizacional, no suporte à modelagem da informação. A identificação de conceitos-chave para a organização é uma das formas de se conseguir integração entre os dados organizacionais. A modelagem será mais genérica e passível de integração quanto mais genéricas forem as ontologias de que se utiliza (CHERMONT, 1999).

#### 4. METADADOS PARA MODELAGEM DA INFORMAÇÃO

Informações sem um significado comum na organização, normalmente, resultam em interpretações divergentes sobre seu significado. Existem vários métodos, na literatura, para ajudar a estruturar a definição das necessidades informacionais de uma empresa ou usuário, possibilitando a obtenção de um ambiente de dados integrados.

Os dicionários de dados, ferramenta centrada na visão técnica, são repositórios de dados sobre os dados. Estes "dados sobre os dados" são

denominados "metadados", e são de grande importância no tratamento e conhecimento dos dados organizacionais.

É vital para qualquer organização possuir termos e significados acordados. Esta visão tem como ponto central a articulação de um "dicionário" corporativo e a definição do vocabulário especial da organização, ou seja, os termos básicos que deverão ser comuns e relativamente invariáveis em todas as transações da empresa. Diferentemente de um dicionário comum, neste caso, não é aceitável haver ambigüidade ou uso de várias acepções para um termo.

Segundo Inmon (apud AYRES, 2000) os metadados devem atender a dois tipos básicos de necessidades: técnica e de negócio. Estratégias de metadados devem permitir aos usuários localizar e entender os dados, de acordo com as perspectivas do negócio. A definição de metadados compõe a base ideal para uma definição comum e compartilhamento dos dados, acrescenta o autor.

Para Souza (1997), metadados são descrições de dados armazenados em banco de dados, ou como é comumente definido: dados sobre dados a partir de um dicionário digital de dados. Esse dicionário de dados normalmente é utilizado para organizar os metadados, contendo uma seção que descreva, numa visão geral, como os dados são subdivididos em arquivos, que campos de registros se relacionam e ainda, possuir tópicos como, convenções adotadas em sua definição. Uma seção principal desse dicionário de dados deveria conter os metadados, assim como as descrições de cada campo. Para cada campo poderiam ser incluídos os itens: nome do campo, descrição do campo; tipo de dados; formato, etc. A finalidade principal dos metadados é documentar e organizar de forma estruturada os dados das organizações com o objetivo de minimizar a duplicação de esforços e facilitar a manutenção dos dados.

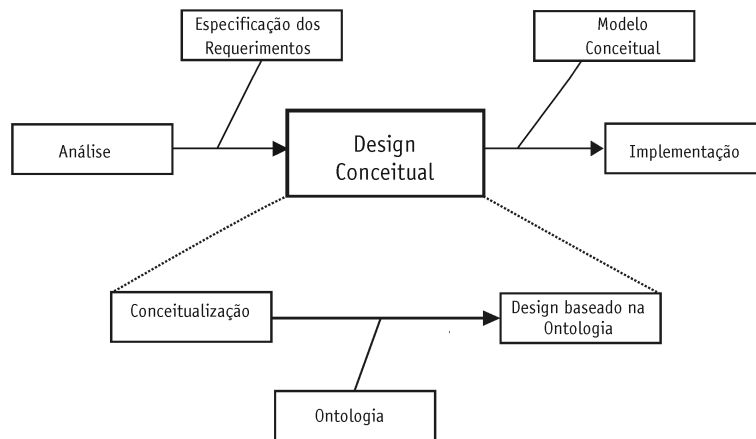
Uma das questões iniciais a serem estabelecidas na gestão dos dados está na definição de padrões para sua definição e nomenclatura, e num repositório que concentre os metadados. Também é necessário definir os critérios de segurança, proteção e integridade dos dados.

## **5. A PRÁTICA NA BUSCA DE UMA CONCEITUAÇÃO ORGANIZACIONAL**

A visão proporcionada pela ontologia, no que tange a definição de uma conceituação organizacional e o gerenciamento de metadados como fatores diferenciais na integração de dados, também se apóia na percepção advinda da prática da administração de dados em organizações. Alguns autores, como Meis e Ostermayer (1996), advogam a necessidade de se incluir uma etapa de modelagem ontológica no projeto de sistemas, conforme ilustra a figura 2.

Esta etapa permitiria resolver o problema da existência de termos com significados ambíguos conforme o setor funcional da organização, o que causa confusão.

Figura 2 - A necessidade da modelagem ontológica



Fonte: Meis e Ostermayer (1996)

A prática descrita a seguir é fruto da experiência vivenciada ao longo de alguns anos, por um dos autores deste artigo, na área de Administração de Dados de uma grande indústria de capital privado, atuante em todo o mercado nacional, e também no mercado externo, com várias unidades produtoras, de comercialização e distribuição de seus produtos. À área de administração de dados cabia a missão de prover um banco de dados que permitisse fornecer informações significativas e relevantes à operação e gestão dos negócios da empresa. Dentro desse enfoque, um dos objetivos era buscar a integração de dados e sistemas, pois como preconiza Garcia (1994) a administração de dados tem por fim coordenar e promover a harmonização das diferentes percepções do negócio e necessidades de informação, de modo que os conceitos e estruturas lógicas de representação sejam compatíveis e constituam um todo coerente, consistente, e sucessivamente enriquecido e partilhável de acordo com as necessidades e perspectivas de cada setor. Sua missão é assegurar a compatibilidade e qualidade de estruturas de representação e definições, para que essas sejam integráveis e integradas, revertendo assim em valor acrescentado para a Organização.

## 5.1 ASPECTOS OBSERVADOS NA FALTA DE INTEGRAÇÃO DE DADOS

Observou-se que havia falta de integração dos dados que compunham as bases de dados organizacionais, apesar da empresa possuir a infraestrutura



tecnológica considerada básica para um ambiente integrado, ou seja, redes de comunicação de dados e a existência de um SGBD, além do uso consistente de técnicas de modelagem na estruturação dos dados. Esta falta de integração resultava em dificuldades (às vezes, até a impossibilidade) de compartilhamento de dados entre os aplicativos, acarretando o desenvolvimento de funções de conversão de dados, transcrição de dados já registrados, atividades que em suma se traduziam em maiores custos e tempo no desenvolvimento de sistemas, além de uma menor confiabilidade e disponibilidade dos dados.

Ao longo do tempo, atuando na análise de diversos sistemas de informação da organização e analisando os diversos bancos de dados implantados, observou-se que o problema de redundância se concentrava em determinadas entidades de dados. Uma das principais causas dessa redundância, vinha de visões do negócio parcialmente diferentes entre os sistemas de informação que criavam/utilizavam estas entidades. A redundância podia ser reflexo de termos diferentes para expressar um mesmo objeto, pessoa, lugar ou conceito do mundo real, ou um mesmo conceito expresso por diferentes terminologias. Quando deparava-se com termos diferentes para uma mesma entidade tinha-se a dificuldade em identificar a existência daquela entidade de dados no banco de dados, gerando, em consequência redundância e o não compartilhamento. O caso contrário, de termos iguais para nomear entidades diferentes, tomava horas de discussão em relação a possibilidade de compartilhamento, culminando com o uso de uma entidade que não atendia as necessidades da área/processo ou a criação de mais uma "visão" da mesma entidade. Adicionalmente a estes problemas ainda existiam aqueles ligados às forças de poder da organização (pressões de prazo, áreas mais influentes que outras, etc.).

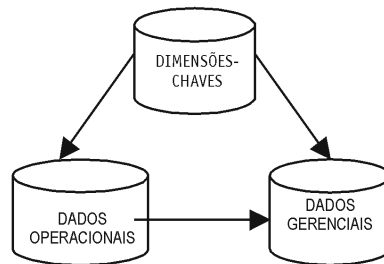
Observou-se, ainda, que as entidades que sofriam maior redundância eram aquelas que expressavam dimensões estratégicas do negócio, como "clientes", "produtos", "unidades organizacionais", etc. Como unidades importantes do negócio, eram necessárias em diversos aplicativos e a falta de um conceito comum dessas dimensões a toda a organização acabava gerando as diversas "visões" nos bancos de dados. A construção do consenso sobre o significado destas dimensões do negócio demonstrou ser tarefa não-trivial, eivada de conflitos de interesses, mas necessária para melhorar a qualidade das informações e o suporte à tomada de decisão.

## 5.2 UMA PROPOSTA DE SOLUÇÃO AOS PROBLEMAS OBSERVADOS

A solução para o problema supra delineado centrou-se em duas questões: (a) identificar um conjunto de dados críticos para o negócio, e para estes, buscar uma denominação e definição comum e, um consenso em toda a organização, quanto a sua definição e uso; (b) desenvolver bancos de dados

gerais e requerer sua utilização por todos os sistemas de informação, para aquelas entidades que representavam as principais dimensões do negócio. A figura 3 ilustra a arquitetura de dados estabelecida.

Figura 3 - Arquitetura de dados proposta



Esta proposta passou a facilitar a integração entre os sistemas de informação, a atender a operação das organizações, a facilitar a consolidação de dados e a atender as necessidades gerenciais no apoio à tomada de decisão. O compartilhamento das bases de dados tornou-se factível, pois os conceitos passaram a expressar uma visão da organização sobre o assunto e não mais uma expressão setorial. As necessidades específicas passaram a ser possíveis e incentivadas, complementando a visão mais geral estabelecida.

Estabelece-se, assim, uma "integração parcial" dos dados. Aquilo definido como "conceito organizacional" deve ser compartilhado, mas é possível acrescentar à visão específica de cada área/processo. O necessário é que haja uniformidade apenas em grau suficiente para se estabelecer a integração de dados. O objetivo deve ser harmonizar, sem homogeneizar. Esta é uma visão mais flexível e menos dispendiosa em relação àquela visão (utópica) de integração total dos dados, dominante nas organizações quando da introdução dos sistemas gerenciadores de bancos de dados e da administração de dados nas organizações, e tão raramente alcançada ou atingida a custos elevados.

Adicionalmente, observa-se que o resultado do trabalho de busca do significado e consenso sobre dimensões estratégicas do negócio possibilita um ganho muito além da integração de dados por sistemas de informação. Obtém-se uma visão mais clara e compartilhada do negócio da própria organização.

### 5.3 O PAPEL DOS METADADOS

A vivência também demonstrou que para apoiar todo o processo de integração de dados é necessário o registro dos metadados relativos as entidades-chave do negócio e seus atributos. A existência de metadados

é condição na identificação e avaliação da congruência de conceitos quando da análise dos sistemas de informação, com vistas ao compartilhamento dos dados e das próprias definições.

A existência de um dicionário de dados, como repositório central dos "dados sobre os dados" no suporte às definições estabelecidas, é verificada como uma ferramenta fundamental na organização, armazenamento, controle e divulgação do acervo de dados. O modelo adotado para este repositório, nas organizações, abrange a visão conceitual dos dados e aponta as estruturas físicas que implementavam tais dados. A documentação das estruturas físicas fica a cargo dos catálogos próprios dos bancos de dados.

Associado ao registro dos metadados deve haver a aplicação de uma convenção geral para nomear dados na organização, que minimize a geração de conflitos entre termos e significados. Neste sentido verificou-se na organização observada a aplicação de uma regra similar baseada em três requisitos na nomeação dos atributos das entidades: identificador do objeto (assunto) - ex: cliente; tipo de objeto - ex: nome; qualificadores (opcional) - ex: fantasia. Vale ressaltar a criação de um certo grau de automação na implementação dessas regras de nomeação dos dados. Observa-se, ainda, que esta automação deve ter sua efetivação no momento do registro de metadados no dicionário de dados.

#### 5.4 O PROCESSO DE INTEGRAÇÃO DE CONCEITOS-CHAVE

O processo de definição e uso de um "conceito organizacional" como base para a integração de dados adotado na prática observada pode ser resumido nas fases:

- seleção: identificação das entidades-chave do negócio;
- organização: determinação dos termos, conceitos e atributos-chave das entidades selecionadas e seu registro;
- limpeza: criação e/ou manutenção das bases de dados de dimensões-chaves do negócio;
- rotina: determinação das ações necessárias à manutenção e evolução dos "conceitos" organizacionais; e
- auto-disciplina: hábito para o cumprimento das definições e procedimentos estabelecidos.

### 6. CONSIDERAÇÕES FINAIS

Sistemas de informação são projetados e desenvolvidos de modo a refletirem nas práticas existentes em determinadas áreas funcionais, resultando em aplicações funcionalmente orientadas. Esta prática

apresenta duas deficiências: (a) aprisiona os dados dentro das áreas funcionais; e (b) proporciona uma falta de visão do conjunto da organização.

Alguns dados, criados independentemente dentro das áreas funcionais, sem considerar necessidades mais abrangentes do negócio, tendem a criar informações conflitantes e de difícil integração no futuro. Normalmente se traduzem em variações de conteúdo e significado dos dados organizacionais, dificultando o compartilhamento de informações entre as áreas/processos organizacionais, assim como seu agrupamento e consolidação para suportar a tomada de decisão nas empresas.

Buscando suprir tais deficiências, o enfoque tecnológico é utilizado. Tenta-se solucionar através da tecnologia um problema de natureza ontológica. Este problema requer mais que uma solução tecnológica ou a aplicação de técnicas de modelagem de dados, implica sim em uma solução da mesma natureza, ou seja, ontológica.

Por exemplo, **produto** pode ser definido pelo gerente industrial como aquilo que deixa a linha de produção. O gerente de vendas, talvez, veja produto como o que deixa o setor de despacho em direção ao cliente. O gerente de marketing, possivelmente, considere produto como o que é transformado em vendas menos eventuais perdas. A tentativa de contabilizar a produção, como se pode prever, está destinada à geração de conflitos devidos à **natural** divergência de pontos-de-vista sobre o significado de um termo usado em contextos distintos com sentidos diferentes. Uma abordagem ontológica deve atribuir um sentido não-ambíguo a produto, ao mesmo tempo em que define outros termos para os diversos significados anteriormente usados.

A proposta de uma definição e disponibilização de alguns conceitos-chave da organização, antecedendo a modelagem de dados dos sistemas de informação é fundamentada pela ontologia e por uma maior atenção ao tratamento dos metadados. Sua eficácia é transparente, pois a prática ratifica a teoria.

Desta forma, pode-se concluir que a integração de dados implica em administrá-los com visão organizacional ampla (não setorizada); com orientação voltada à visão de negócio (a tecnologia é apenas um meio); com um dicionário de dados (ferramenta fundamental para a organização); e com mecanismos adequados de armazenamento, controle e divulgação do acervo.

A busca de significado e consenso sobre dimensões estratégicas do negócio possibilita um ganho que extrapola a integração de dados por sistemas de informação, proporcionando uma visão mais clara e compartilhada do negócio da própria organização.

**1. INTRODUCCIÓN**V  
e  
r  
s  
i  
ó  
n  
e  
n  
E  
s  
p  
a  
ñ  
o  
I

Las organizaciones habitualmente proyectan e desarrollan sus sistemas de información de modo que se encuadren en las prácticas existentes resultando de eso que muchas de las aplicaciones son funcionalmente orientadas. Según expresa Davenport (1994): sistemas de marketing resuelven problemas de marketing, sistemas de ventas resuelven problemas de ventas, sistemas de fabricación resuelven problemas de fabricación. Tales sistemas "fraccionados" terminan no proporcionando una visión de conjunto de la organización, aprisionando los datos dentro de las funciones.

Una orientación dirigida al desarrollo de sistemas de informaciones independientes, sin considerar las necesidades más abarcadoras del negocio, tiende a crear una red de informaciones encontradas y de difícil integración en el futuro. Esta orientación normalmente se traduce en variaciones de contenido y significado de los datos organizacionales, dificultando la coparticipación de informaciones entre áreas/procesos organizacionales y en la reunión y consolidación de informaciones para sustentar la tomada de decisión en las empresas (BRACKETT, 1996; INMON, 1997).

Se observa que hay una centralización de enfoque en la tecnología para suplir tales deficiencias. Mas el problema de la integración de datos organizacionales es más que una solución técnica relacionada al uso de un sistema administrador del banco de datos (SGBD) o la aplicación de técnicas de modelado de datos durante el desarrollo de sistemas de información. La práctica en desarrollo de sistemas y en la administración muestra que ni siempre el uso del modelado de datos y la existencia de una tecnología de bancos de datos impiden la falta de integración entre los datos de una organización. Las aplicaciones, frecuentemente, terminan siendo desarrolladas teniendo en cuenta la visión particular de un área de la empresa o la necesidad de desempeño o extracción de información de una comunidad en particular dentro de la organización.

Así, antecediendo el modelado de datos de los sistemas de información de una organización, se propone que haya una definición y disponibilidad de algunos conceptos clave de la organización. Se optó por buscar un respaldo a esa cuestionen en las antologías comunes, donde existe la formalización de conceptos clave de la organización y en una atención mayor al tratamiento de los meta datos, en una orientación menos tecnológica y más dirigida a la visión del negocio. Siguiendo esta visión, este artículo primeramente analiza algunos aspectos conceptuales sobre integración de datos, ontologías y meta datos. En una segunda instancia, presenta la consolidación de experiencias en busca de la

coparticipación de una visión común de datos, con el objetivo de facilitar la integración de datos tratados por los sistemas de informaciones vivenciados en empresas. Finalmente, son presentadas conclusiones referentes al tema en cuestión.

## 2. INTEGRACIÓN DE DATOS EN LAS ORGANIZACIONES

Según Nolan (apud AYRES, 2000), la integración de datos ocurre dentro de las organizaciones en un proceso evolutivo. De acuerdo con la maduración de las empresas en el proceso de datos se verifica que hay una evolución en la forma de planeamiento y control de datos y aplicaciones, buscando su integración con los procesos y estrategias organizacionales. De acuerdo con el autor anteriormente citado, esta evolución sucede en seis etapas, descritas a seguir.

- **Iniciación:** fase de introducción de las computadoras en la organización. Fase de crecimiento lento y con énfasis en el aprendizaje de la tecnología. Las aplicaciones que son desarrolladas son simples, generalmente de naturaleza contable o financiera.

- **Expansión:** fase de la proliferación de aplicaciones en la organización con substancial pérdida de control por parte de la gestión.

- **Control:** fase de control por parte de la gestión, con énfasis en la documentación de las aplicaciones existentes y en el planeamiento formal de sistemas de información. Las aplicaciones dejan de tener un enfoque estrictamente operacional, pasando también a atender aspectos administrativos.

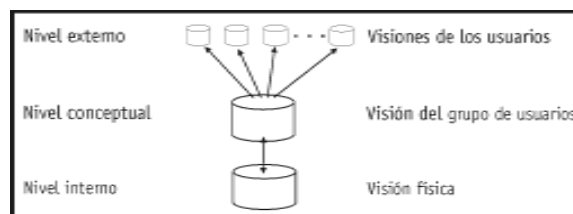
- **Integración:** fase caracterizada por la integración de las aplicaciones existentes, originalmente concebidas de forma independiente, sacando así provecho de las tecnologías de bancos de datos y telecomunicaciones.

- **Administración de datos:** fase en que la tecnología de bancos de datos ya está consolidada, surgiendo la función de administración con el objetivo de planear y controlar el uso de datos. El énfasis está en sistemas integrados a partir de la coparticipación de datos y aplicaciones comunes.

- **Madurez:** fase en que las aplicaciones son integradas y reflejan el flujo de informaciones de la organización. La tecnología de la información está alineada con los procesos administrativos de la organización y los recursos de datos son planeados para atender las necesidades estratégicas.

Desde el punto de vista estrictamente tecnológico, la integración de datos está íntimamente asociada a la integración de sistemas de información. La arquitectura ANSI/SPARC (TSICHRITZIS; KLUG, 1978) ilustrada en la figura 1 es ampliamente adoptada en la construcción de los bancos de datos utilizados por los sistemas de información, permite integrar aplicaciones, mas no inhibe el problema de integración de datos.

Figura 1 - Arquitectura ANSI/SPARC



Fuente: Adaptada de Tsichritzis y Klug (1978)

De acuerdo con Tsichritzis y Klug (1978), esa arquitectura preconiza que un banco de datos puede ser entendido a partir de tres perspectivas:

- La visión de los usuarios o nivel externo del banco de datos que reúne las varias concepciones que los usuarios pueden tener sobre lo que es el banco de datos. Estas visiones son **tendenciosas** por definición;
- La visión del grupo de usuarios o nivel conceptual que es un levantamiento de las visiones de usuarios para un único documento, compatible con todas las visiones de usuario;
- La visión física o nivel interno que trata del banco de datos como es implementado en la computadora, con todas las características de administración de banco de datos.

La visión del grupo de usuarios es expresa en documentos como modelos o diagramas de datos o en *scripts* de definición de datos. Estos documentos expresan la organización de un banco de datos sobre datos o **metadatos**. Los metadatos deben contemplar todas las necesidades de información de los usuarios y compatibilizar la terminología y significados, lo que puede ser hecho a través de ontologías. Tanto los metadatos cuanto las ontologías son descritos más adelante en este artículo.

## 2.1 EVOLUCIÓN DE LA TECNOLOGÍA DE LA INFORMACIÓN EN LAS ORGANIZACIONES

El surgimiento del problema relacionado a la integración de datos está en buena parte asociado al propio proceso histórico-evolutivo de la tecnología de la información, que prácticamente a cada 10 o 15 años provoca una revolución (INPRISE, 2001).

En el inicio de la era de la informática, en los años 50 hasta el final de la década del 70 se presenció el apareamiento de sistemas de información formados por programas y datos fuertemente independientes. Cada una de las aplicaciones componía un sistema de información cerrado, centralizado, y contenía toda una lógica de presentación, de negocio y de acceso a datos particulares, desarrollando una copia particular de sus datos. Así, a pesar

que diferentes sistemas utilizaban las mismas fuentes de datos, las organizaciones terminaban por almacenar copias redundantes de los datos en múltiples sistemas.

Alrededor de 1970 surgió una nueva tecnología de almacenamiento y acceso a datos - el disco o DASD (*Direct Access Storage Device*, o dispositivo de almacenamiento de acceso directo), y con él, un nuevo tipo de *software* conocido como SGBD (Sistema Administrador de Banco de Datos), que presentaba el concepto de "banco de datos" definido por Martín (1976) como un conjunto de datos Interrelacionados, de redundancia controlada, sirviendo a una o más aplicaciones. El uso de los SGBD se constituyó mucho más en el aprovechamiento de un método de almacenamiento y recuperación de datos más poderoso, de que efectivamente en un repositorio único de datos compartibles.

En el comienzo de la década de 1980 nuevas tecnologías como los PCs (*Personal Computer* o computadoras personales), las redes de grupos, las herramientas de cuarta generación comenzaron a despuntar, poniendo en las manos de los usuarios finales la posibilidad de controlar datos y aplicaciones (*end user computing*). Esta nueva "corriente" vino a agravar el cuadro de redundancia de datos, conflicto de resultados y conocimiento de los datos (INMON, 1997; LAUDON; LAUDON, 1996).

## 2.2 EL PROBLEMA DE LA INTEGRACIÓN DE DATOS

Este historial tecnológico fortalece el apareamiento de una laguna en la concepción y construcción de los sistemas de información, en la cual los sistemas son proyectados sin tomar en cuenta la definición de las necesidades interfuncionales y sí con una visión específica de nivel más bajo, representando la unidad de análisis o de negocio en que será construido.

Existen varios métodos en la literatura para ayudar a estructurar la definición de esas conexiones entre los sistemas. Un ejemplo es la metodología BSP (*Business System Planning*) desarrollada por la IBM y que fue ampliamente aplicada en diversas grandes empresas. La aplicación de estos métodos no es simple, involucrando muchas personas, tiempo y dinero. Los resultados son frecuentemente interrumpidos por la falta de visión de los objetivos a ser alcanzados con tal dispendio por parte de los usuarios. Otras veces son abandonados ya en el proceso de desarrollo de las aplicaciones especificadas, pues las aplicaciones definidas ya no traducen más la realidad vigente en la organización o existe una falta de entendimiento de lo que fue definido, en función de que el proceso esté siendo conducido por personas diferentes de aquellas que originalmente hicieron la especificación (MCGEE; PRUSAK. 1994; DAVENPORT, 1998).

El uso de técnicas y metodologías de modelado de la información, la



identificación de las necesidades de información en el proyecto del banco de datos no introduce automáticamente una visión de datos integrados. El modelado de la información se preocupa con "cuál" dato es necesario y sus relaciones con otros, expresados en un modelo de datos. Se observa que esta técnica define con claridad las entidades de datos o sus relaciones y algunas reglas de negocio, mas no garantiza la integración de datos, pues el enfoque normalmente es un área o proceso específico.

### 3. ONTOLOGÍAS PARA MODELADO DE LA INFORMACIÓN

Ontología es un término derivado y tomado prestado de la Filosofía y aunque el interés en él esté aumentado es de difícil definición y no es utilizado de la misma manera por diferentes personas. Ontología ha sido utilizado en Ciencia de la Computación con el significado de una "especificación explícita de una conceptualización" (GRUBER, 1994). Esto implica que para construir una ontología en un cierto dominio es necesario definir de forma explícita, abarcadora y no ambigua los conceptos importantes en el dominio.

La Filosofía define ontología como siendo el estudio de la realidad y de los objetos independiente del conocimiento que se posee de ellos (CHERMONT, 1999; KERN, 1997). En este contexto ontología expresa algún modo de visión del mundo, o sea una perspectiva, un modo de pensar sobre algún aspecto del mundo natural o artificial. Puede ser encarada como un paradigma que utiliza una serie de conceptos y definiciones. Pero ¿qué visión de mundo es esa?

Ocurren problemas cuando términos similares son usados con significados diferentes o semejantes, pudiendo causar ambigüedad e imprecisión. Debido a las ambigüedades se hace necesario definir y explicitar esa visión del mundo en la cual el modelo de negocio está basado.

Una ontología debe actuar como un medio que permita la comunicación entre diferentes personas a través de la organización; personas y sistemas computacionales implementados y diferentes sistemas computacionales. En este sentido, el abordaje de **ontología** ha agregado el adjetivo **común**. Ontologías comunes son fundamentales para la integración de subsistemas.

Es importante también que la ontología abarque la adquisición, representación y manipulación del conocimiento del negocio que está siendo modelado. Por este motivo no debe ser considerada estática, mas siempre ser revisada, ampliada y corregida.

Una ontología debe ser construida de modo que se maximice su futura reutilización y aumento de objetivo, que puede ser alcanzado por el balanceo correcto entre ser específico o suficiente para la ejecución de tareas, mas no

tan específico que no sea de utilidad para otros. Esta visión proporcionada por la ontología delinea la necesidad de una visión conceptual, de espectro organizacional en el apoyo al modelado de la información. La identificación de conceptos clave para la organización es uno de los modos de conseguir integración entre los datos organizacionales. El modelado será más genérico y pasible de integración cuanto más genéricas sean las ontologías que utiliza. (CHERMONT, 1999).

#### 4. METADATOS PARA MODELADO DE LA INFORMACIÓN

Informaciones sin un significado común en la organización, normalmente, resultan en interpretaciones divergentes sobre su significado. Existen varios métodos en la literatura para ayudar a estructurar la definición de las necesidades informativas de una empresa o usuario, posibilitando la obtención de un ambiente de datos integrados.

Los diccionarios de datos, herramienta centrada en la visión técnica, son repositorios de datos sobre los datos. Estos "datos sobre los datos" son denominados "metadatos" y son de gran importancia en el tratamiento y conocimiento de los datos organizacionales.

Es vital para cualquier organización poseer términos y significados convenidos. Esta visión tiene como punto central la articulación de un "diccionario" corporativo y la definición del vocabulario especial de la organización, o sea, los términos básicos que deberán ser comunes y relativamente invariables en todas las transacciones de la empresa. Diferentemente de un diccionario común, en este caso no es aceptable que haya ambigüedad o uso de varias acepciones para un término.

Según Inmon (apud AYRES, 2000) los metadatos deben atender a dos tipos básicos de necesidades: técnica y de negocio. Estrategias de metadatos deben permitir a los usuarios localizar y entender los datos, de acuerdo con las perspectivas del negocio. La definición de metadatos compone la base ideal para una definición común y coparticipación de los datos, amplía el autor.

Para Souza (1997), los metadatos son descripciones de datos almacenados en banco de datos, o como comúnmente es definido: datos sobre datos a partir de un diccionario digital de datos. Este diccionario de datos normalmente es utilizado para organizar los metadatos, conteniendo una sección que describa, en una visión general, cómo los datos son subdivididos en archivos, qué campos de registros se relacionan y aun poseer tópicos tales como: convenciones adoptadas en su definición. Una sección principal de ese diccionario de datos debería contener los metadatos, así como las descripciones de cada campo. Para cada campo podrían ser incluidos los ítem: nombre del campo, descripción del campo,

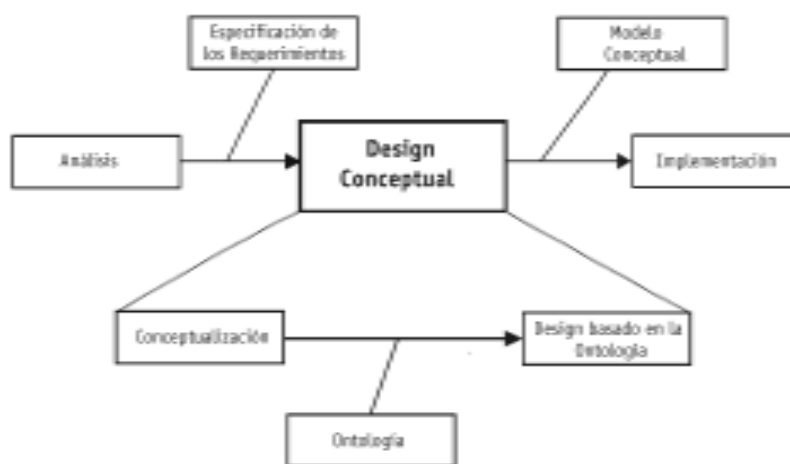
tipo de datos, formato, etc. La finalidad principal de los metadatos es documentar y organizar de modo estructurado los datos de las organizaciones con el objetivo de minimizar duplicación de esfuerzos y facilitar la manutención de datos.

Una de las cuestiones iniciales a ser establecidas en la gestión de los datos está en la definición de padrones para su definición y nomenclatura y en un repositorio que concentre los metadatos. También es necesario definir los criterios de seguridad, protección e integridad de los datos.

## 5. LA PRÁCTICA EN LA BÚSQUEDA DE UNA CONCEPTUACIÓN ORGANIZACIONAL

La visión proporcionada por la ontología en lo que atañe a la definición de una conceptuación organizacional y la administración de metadatos como factores diferenciales en la integración de datos también se apoya en la percepción advenida de la práctica de la administración de datos en organizaciones. Algunos autores, como Meis y Ostermayer (1986) defienden la necesidad de incluir una etapa de modelado ontológico en el proyecto de sistemas, según ilustra la figura 2. Esta etapa permitirá resolver el problema de la existencia de términos con significados ambiguos según el sector funcional de la organización, lo que causa confusión.

Figura 2 - La necesidad de modelado ontológico



Fuente: Meis y Ostermayer (1996)

La práctica descrita a seguir es fruto de la experiencia vivenciada a lo largo de algunos años por uno de los autores de este artículo en el área de Administración de Datos de una grande industria de capital privado actuante

en todo el mercado nacional y también en el mercado externo con varias unidades productoras de comercialización y distribución de sus productos. Al área de administración de datos poseía la misión de proporcionar un banco de datos que permitiese proveer informaciones significativas y relevantes a la operación y gestión de los negocios de la empresa. Dentro de este enfoque, uno de los objetivos era buscar la integración de datos y sistemas, pues como preconiza García (1994) la administración de datos tiene por finalidad coordinar y promover la armonización de las diferentes percepciones del negocio y necesidades de información, de modo que los conceptos y estructuras lógicas de representación sean compatibles y constituyan un todo coherente, consistente y sucesivamente enriquecido compatible de acuerdo con las necesidades y perspectivas de cada sector. Su misión es asegurar la compatibilidad y calidad de estructuras de representación y definiciones, para que las mismas sean integrables e integradas, revirtiendo así en valor aumentado para la Organización.

### 5.1 ASPECTOS OBSERVADOS EN LA FALTA DE INTEGRACIÓN DE DATOS

Se observó que había falta de integración de los datos que componían las bases de datos organizacionales, a pesar de la empresa poseer la infraestructura tecnológica considerada básica para un ambiente integrado, o sea, redes de comunicación de datos y la existencia de un SGBD, además del uso consistente de técnicas de modelado en la estructuración de datos. Esta falta de integración resultaba en dificultades (a veces hasta en la imposibilidad) de coparticipación de datos entre los aplicativos, acarreado el desarrollo de funciones de conversión de datos, transcripción de datos ya registrados, actividades que en suma se traducían en mayores costes y tiempo en el desarrollo de sistemas, además de una menor confiabilidad y disponibilidad de los datos.

A lo largo del tiempo, actuando en el análisis de diversos sistemas de información de la organización y analizando los diversos bancos de datos implantados, se observó que el problema de redundancia se concentraba en determinadas entidades de datos. Una de las principales causas de esas redundancias venía de visiones del negocio parcialmente diferentes entre los sistemas de información que creaban/utilizaban estas entidades. La redundancia podía ser reflejo de términos diferentes para expresar un mismo objeto, persona, lugar o concepto del mundo real, o un mismo concepto expresado por diferentes terminologías. Cuando se deparaba con términos diferentes para una misma entidad existía la dificultad en identificar la existencia de aquella entidad de datos en el banco de datos generando en consecuencia redundancia y la no coparticipación. El caso contrario de términos iguales para nombrar entidades diferentes llevaba horas de discusión

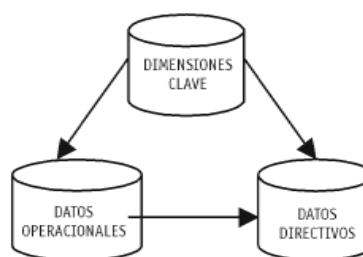
con relación a la posibilidad de coparticipación, culminando con el uso de una entidad que no atendía las necesidades del área/proceso o la creación de una "visión" más de la misma entidad. Adicionalmente a estos problemas aun existían aquellos conectados a las fuerzas de poder de la organización (presiones de plazo, áreas mas influyentes que otras, etc.).

Se observó también que las entidades que sufrían mayor redundancia eran aquellas que expresaban dimensiones estratégicas del negocio como "clientes", "productos", "unidades organizacionales", etc. Como unidades importantes del negocio eran necesarias en diversos aplicativos y la falta de un concepto común de esas dimensiones a toda la organización terminaba generando las diversas "visiones" en los bancos de datos. La construcción del consenso sobre el significado de estas dimensiones del negocio demostró ser una tarea no trivial, contaminada de conflictos de intereses, pero necesaria para mejorar la calidad de las informaciones y el apoyo a la tomada de decisión.

## 5.2 UNA PROPUESTA DE SOLUCIÓN A LOS PROBLEMAS OBSERVADOS

La solución para el problema anteriormente delineado se centró en dos cuestiones: (a) identificar un conjunto de datos críticos para el negocio y para que estos busquen una denominación y definición común y un consenso en toda la organización cuanto a su definición y uso; (b) desarrollar bancos de datos generales y requerir su utilización por todos los sistemas de información, para aquellas entidades que representaban las principales dimensiones del negocio. La figura 3 ilustra la arquitectura de datos establecida.

Figura 3 - Arquitectura de datos propuesta



Esta propuesta pasó a facilitar la integración entre los sistemas de información, a atender la operación de las organizaciones, a facilitar la consolidación de datos y a atender las necesidades administrativas en el apoyo a la tomada de decisión. La coparticipación de las bases de datos se hizo factible, pues los conceptos pasaron a expresar una visión de la organización sobre el asunto y no más una expresión sectorial. Las necesidades específicas pasaron a ser posibles e incentivadas complementando la visión más general establecida.

Se establece así una "integración parcial" de los datos. Aquello definido como "concepto organizacional" debe ser compartido, mas es posible agregar la visión específica de cada área/proceso. Lo necesario es que haya uniformidad apenas en grado suficiente para establecerse la integración de datos. El objetivo debe ser armonizar sin homogeneizar. Esta es una visión más flexible y menos dispendiosa con relación a aquella visión (utópica) de integración total de los datos, dominante en las organizaciones en la introducción de los sistemas administradores de bancos de datos y la administración de datos en las organizaciones, y tan raramente alcanzada a costes elevados.

Adicionalmente se observa que el resultado del trabajo de búsqueda de significado y consenso sobre dimensiones estratégicas del negocio posibilita una ganancia mucho más allá de la integración de datos por sistemas de información. Se obtiene una visión más clara y compartida del negocio de la propia organización.

### 5.3 EL PAPEL DE LOS METADATOS

La vivencia también demostró que para apoyar todo el proceso de integración de datos es necesario el registro de los metadatos relativos a las entidades clave del negocio y sus atributos. La existencia de metadatos es condición en la identificación y evaluación de la congruencia de conceptos en el momento del análisis de los sistemas de información con el propósito de compartir los datos y las propias definiciones.

La existencia de un diccionario de datos como repositorio central de los "datos sobre los datos" en el apoyo a las definiciones establecidas es verificada como una herramienta fundamental en la organización, almacenamiento, control y divulgación del acervo de datos. El modelo adoptado para este repositorio en las organizaciones abarca la visión conceptual de los datos y señala las estructuras físicas que implementaban tales datos. La documentación de las estructuras físicas queda a cargo de los catálogos propios de los bancos de datos.

Asociado al registro de los metadatos debe haber la aplicación de una convención general para nombrar datos en la organización, que minimice la creación de conflictos entre términos y significados. En este sentido se verificó en la organización observada la aplicación de una regla similar basada en tres requisitos en el nombramiento de atributos de las entidades: identificador del objeto (asunto) - ejemplo: cliente; tipo de objeto - ejemplo: nombre; calificadores (opcional) - ejemplo: fantasía. Vale resaltar la creación de un cierto grado de automatización en la implementación de estas reglas de nombramiento de datos. Se observa también que esta automatización debe tener su efectación en el momento del registro de meta datos en el diccionario de datos.

## 5.4 EL PROCESO DE INTEGRACIÓN DE CONCEPTOS CLAVE

El proceso de definición y uso de un "concepto organizacional" como base para la integración de datos adoptado en la práctica observada puede ser resumido en las fases:

- Selección: identificación de las entidades clave del negocio;
- Organización: determinación de los términos, conceptos y atributos clave de las entidades seleccionadas y su registro;
- Limpieza: creación y/o manutención de las bases de datos de dimensiones clave del negocio;
- Rutina: determinación de las acciones necesarias para la manutención y evolución de los "conceptos" organizacionales; y
- Autodisciplina: hábito para el cumplimiento de las definiciones y procedimientos establecidos.

## 6. CONSIDERACIONES FINALES

Los sistemas de información son proyectados y desarrollados de modo que reflejen las prácticas existentes en determinadas áreas funcionales, resultando en aplicaciones funcionalmente orientadas. Esta práctica presenta dos deficiencias: (a) aprisiona los datos dentro de las áreas funcionales; y (b) proporciona una falta de visión del conjunto de la organización.

Algunos datos, creados independientemente dentro de las áreas funcionales, sin considerar necesidades mas abarcadoras del negocio, tienden a crear informaciones encontradas y de difícil integración en el futuro. Normalmente se traducen en variaciones de contenido y significado de los datos organizacionales, dificultando la coparticipación de informaciones entre las áreas/procesos organizacionales, así como su agrupamiento y consolidación para apoyar la tomada de decisión en las empresas.

Buscando suplir tales deficiencias, el enfoque tecnológico es utilizado. Se trata de solucionar a través de la tecnología un problema de naturaleza ontológica. Este problema requiere más que una solución tecnológica o la aplicación de técnicas de modelado de datos, implica sí una solución de la misma naturaleza, o sea, ontológica.

Por ejemplo, **producto** puede ser definido por el gerente industrial como aquello que deja la línea de producción. El gerente de ventas, talvez, vea el producto como lo que deja el sector de despacho en dirección al cliente. El gerente de marketing, posiblemente, considere producto como el que es transformado en ventas menos eventuales pérdidas. La tentativa de contabilizar la producción, como se puede prever, está destinada a la generación de conflictos debidos a la **natural** divergencia de puntos de vista sobre el significado de un término usado en contextos distintos con sentidos diferentes.

Un abordaje ontológico debe atribuir un sentido no ambiguo al producto, al mismo tiempo en que define otros términos para los diversos significados anteriormente usados.

La propuesta de una definición y disponibilidad de algunos conceptos clave de la organización, antecediendo al modelado de datos de los sistemas de información es fundamentada por la ontología y por una mayor atención al tratamiento de los meta datos. Su eficacia es transparente, pues la práctica ratifica la teoría.

De esta forma se puede concluir que la integración de datos implica administrarlos con visión organizacional amplia (no sectorizada); con orientación hacia la visión de negocio (la tecnología es apenas un medio); con un diccionario de datos (herramienta fundamental para la organización); y con mecanismos adecuados de almacenamiento, control y divulgación del acervo.

La búsqueda de significado y consenso sobre dimensiones estratégicas del negocio posibilita una ganancia que sobrepasa la integración de datos por sistemas de información, proporcionando una visión más clara y compartida del negocio de la propia organización.

---

## REFERÊNCIAS BIBLIOGRAFICAS

AYRES, N. M. **Fatores condicionantes na estruturação de uma gestão estratégica da informação**: uma contribuição na evolução da administração da informação e da tecnologia nas organizações. Florianópolis, 2000. 170 f. Dissertação (Mestrado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2000.

BRACKETT, M. **The data warehouse challenge**: taming data chaos. New York: John Wiley & Sons, 1996.

CHERMONT, G. Reutilizando dados com a modelagem de negócios. **Developers' Magazine**, v. 4, n. 38, p. 30-34, out.1999.

DAVENPORT, T. **Reengenharia de processos**: como inovar na empresa através da tecnologia da informação. Rio de Janeiro: Campus, 1994.

DAVENPORT, T. **Ecologia da informação**: por que só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura, 1998.

GARCIA, A. Administração de dados na prática: o repositório de meta-informação. **Sistemas de Informação**, n. 2, p. 27-33, 1994.

GRUBER, T. Towards principles for the design of ontologies used for knowledge sharing. **International Journal of Human and Computer Studies**, v. 43, n. 5/6, p. 907-928, 1994.

INMON, W. **Como construir o data warehouse**. Rio de Janeiro: Campus, 1997.

INPRISE do Brasil Ltda. **O futuro do desenvolvimento de software no Brasil**. Disponível em: <[http://www.borland.com.br/artigos/artigo\\_phampshire\\_2.htm](http://www.borland.com.br/artigos/artigo_phampshire_2.htm)>. Acesso em: 01 maio 2001.

KERN, V. M. **Manutenibilidade da semântica de modelos de dados de produtos compartilhados em rede interoperável**. Florianópolis, 1997. 139f. Tese (Doutorado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 1997.

LAUDON, K; LAUDON J. **Management information systems: organization and technology**. Englewood Cliffs: Prentice-Hall, 1996.



- MARTIN, J. **Principles of data-base management**. Englewood Cliffs: Prentice-Hall, 1976.
- MCGEE, L.; PRUSAK, L. **Gerenciamento estratégico da informação**: aumente a competitividade e a eficiência de sua empresa utilizando a informação como uma ferramenta estratégica. Rio de Janeiro: Campus, 1994.
- MEIS, E.; OSTERMAYER, R. Recommendations to the STEP committee. **Kactus Consortium, Document KACTUS-01-RPK-D007**, v. 11, 28 Sept. 28 1996. Disponível em: <<http://swi.psy.uva.nl/projects/NewKACTUS/home.html>>. Acesso em 14 jul. 2002.
- SOUZA, T. B. et al. Metadados: catalogando dados na Internet. **TransInformação**, v. 9, n. 2, maio/ago. 1997.
- TSICHRITZIS, D.; KLUG, A. The ANSI/X3/SPARC DBMS framework report of the study group on database management systems. **Information Systems**, v. 3, p. 173-191, 1978.

