

GESTÃO DE DADOS E A ABORDAGEM CAPITALIZAÇÃO DE CONHECIMENTO

Vanderlei Vilhanova Ortencio

PUC PR – Pontifícia Universidade Católica do Paraná.

Celepar – Companhia de Informática do Paraná.

Rua Mateus Leme, 1142, Centro Cívico. CEP: 81 830 010

email: vilanova@pr.gov.br, fones: 0xx-41-350-5449; 0xx-41-9983-0014.

Resumo

Na década de 1970, já era grande o movimento em torno da importância de Administração de Dados Corporativos. O dado persistente passa a ser visto como recurso, e portanto, patrimônio da organização.

Com a evolução tecnológica, este tema ganhou proporção ainda maior, suscitando a idéia de Gerência de Conhecimento, e o conceito de Capitalização do Conhecimento. Este trabalho apresenta uma introdução a este tema buscando uma integração de conceitos e a formulação de uma cultura organizada sobre estes assuntos.

Abstract

In the 1970 decade, it was a large movement around of the importance of Corporative Data Administration. The Persistent Data became to be considered as resource, and so, an Organization asset.

With the technologic evolution, this issue increased in proportion, rising the idea of Knowledge Management and the Knowledge Capitalization. This work item introduce an introduction to this matter seeking a concept integration and the formulization of a organized culture on these issues.

Palavras chave: Gestão de Dados, Capitalização de Conhecimento, Gestão Conhecimento, Bases de Conhecimento Corporativo.

1. Introdução

Uma característica altamente relevante no uso dos Computadores é que basicamente eles operam sobre elementos de Dados. Seja em aplicações científicas, militares, ou em áreas de negócio específicas, um determinado conjunto de dados está sempre sofrendo operações de modo a produzir novos dados e oferecer aos seus usuários serviços e informação [ORTE98].

Dados são representações de fatos e conceitos que se supõe existentes no mundo Real ou no mundo Ideal. Informação pode ser entendido como uma inferência possível a partir de um conjunto de dados inter-relacionados e que permitem uma decisão ou ação conseqüente tanto por humanos como por máquinas.

Um Serviço, neste contexto, pode ser a elaboração de dados para apresentação, como Gráficos estatísticos ou o acionamento e/ou monitoramento de outras máquinas, como lançamento de mísseis, acionamento de braço mecânico, etc.

Devido a essa característica, durante algum tempo, os centros de utilização de computadores foram designados como CPD (Centro de Processamento de Dados), o que denota uma conotação com processos automatizados manipulando Dados. De um modo geral ainda, a aplicação de Computadores está situada no contexto chamado de (Tecnologia da Informação). No Brasil, o termo Informática ganhou grande popularidade por ocasião da disseminação de uso de computadores motivada em grande parte pela evolução da tecnologia de redes, a explosão do consumo de microcomputadores e do salto tecnológico experimentado pela indústria de software oportunizada principalmente por este nicho de mercado [ORTE98].

Informática é definido no dicionário brasileiro como sendo a "ciência do tratamento racional e automático da informação, considerada esta como suporte do conhecimento e da comunicação" [FERR75]. Na língua inglesa este termo, "informatics", é considerado uma tradução da palavra russa "informátika" que o dicionário designa como "information science" [COST95], cuja tradução seria "Ciência da Informação", porém tendo a mesma definição que informática em português. Entretanto, "informatics" não tem a mesma consagração nos países de língua inglesa, sendo largamente utilizado o termo IT (Information Technology).

2. A Saga da Administração de Dados

Sendo os dados elementos fundamentais, por se constituírem em unidades atômicas que ao serem combinadas e contextualizadas, produzem informação, tem sido crucial o estudo e desenvolvimento de mecanismos para tratamento adequado sob vários pontos de vista, tais como: Persistência, Estrutura, Ciclo de Vida, Qualidade, Modelagem, etc. Surgiu com muita ênfase a função de Administrador de Dados, como uma atividade abrangente de gestão dos dados corporativos como sendo recursos fundamentais de uma organização. A abrangência da AD (Administração de Dados) era definida como instrumento de gestão de todos os dados da organização (não apenas dados em formato "machine readable" mantidos em mídia conveniente aos processos automatizados). O que tornava seu escopo, em essência, muito próximo ao que hoje se refere uma grande parte do contexto de que trata a Gestão de Conhecimento Corporativo.

Com o passar dos tempos, as organizações que investiram em AD colheram muitos benefícios, porém esta função até os dias de hoje encontra muitas dificuldades de sucesso [ORTE94]. As atividades funcionais que persistem mais facilmente nas organizações são aquelas associadas com objetivos de Administração de Banco de Dados, porque são mais operacionais, dependente de conhecimento técnico específico de SGBD (Sistemas Gerenciadores de Banco de Dados) e necessárias ao suporte para funcionamento do SGBD, e a manutenção da disponibilidade dos bancos de dados. As outras funções, associadas com a inteligência na gestão dos dados, para torná-los ativos corporativos e matéria-prima para decisões estratégicas tem sido negligenciadas. Como tudo que é abandonado à própria sorte, um ente acaba isolado ou diluído, num sistema onde a interação entre entes exige identidade para que o ente possa existir, então o definhamento é inevitável.

Além dos aspectos culturais, que implicam num entendimento diversificado sobre temas comuns, os problemas políticos tem sido a principal causa dos problemas de AD nas organizações [ORTE94]. O principal deles, a falta de vontade política dos escalões superiores de uma instituição para a AD, advém tanto da pobreza de visão como da pressão interna advinda de conflitos por poder nas hierarquia da Organização que obviamente está sujeito a atitudes tipicamente humanas, como os aforismos: "Se não for importante, porque implantar? e se for importante, é obvio que deveria ser minha atribuição". Além disso, os benefícios da AD devem se fazer sentir por toda a organização, de modo que todas as áreas passam a capitalizar os efeitos desses benefícios. Isto faz com que ao desempenhar sua missão, a AD não objetiva visibilidade para si, mas para os efeitos de sua ação. Assim, passa a ser visto apenas como um custo, ou quando muito, como um investimento. Com o tempo, a missão de AD, neste cenário, passa a depender mais do heroísmo de seus agentes e vai definhando até que apenas manutenções inadiáveis são as únicas ações possíveis.

Na maioria das vezes sem perceber, a organização vai ficando a mercê dos fornecedores de tecnologia e dos modismos ditados pela indústria e pela mídia. As decisões sobre investimentos em tecnologia para tratamento de dados são tomadas sem a participação de um agente de AD, algumas vezes porque este não existe formalmente, outras porque

não está em nível adequado na estrutura da organização para participar destas discussões. Porém após feito qualquer investimento, é reafirmado por qualquer fornecedor a importância da AD ou mais especificamente de suas funções para viabilização dos benefícios pretendidos.

Um exemplo disto, é o discurso de Datawarehouse. Muito pouco acrescenta aos objetivos historicamente perseguidos pelo AD, de definição de uma arquitetura de dados para aumentar a liquidez do dado através do compartilhamento destes por toda organização, maximizando o potencial de uso dos dados para tomada de decisão, dentre outros. Como a grande maioria das organizações não conseguiu cumprir estes objetivos, oportunizou um nicho de discurso e oportunidade para a indústria colocar no mercado uma porção de produtos de EIS (Executive Information System) encalhados. Isto acontecia porque seus usuários (executivos) percebiam que era por demais trabalhoso alimentar as bases de dados 'a mão', e aparentemente impossível fazê-lo a partir dos BD corporativos. O mercado de Datawarehouse traz então para o executivo uma nova linha de produtos. Para beneficiar-se deles, é necessário prover uma base de dados organizada, mapeada, dominada! Que represente os fatos da organização. Para isso, consultores oferecem serviços para apoiar no processo de Datawarehousing. Primeiro, faz-se uma cópia da base de dados corporativa, que continua existindo como "dados operacionais", criando uma nova base de dados redundante, que tende a se tornar muito maior do que a base de dados original. Torna então todos estes dados acessíveis para aplicações de SAD (Sistemas de Apoio Decisão), e pode-se então utilizar aquelas ferramentas de EIS, e também, é claro, outras mais atualizadas, que implementam conceitos de OLAP (On Line Analytical Processing) para análise de dados selecionados do BD, ou ainda ferramentas para descoberta de conhecimento nos BD (KDD – Knowledge Discovery in Databases). Depois de algum tempo, muitos dólares investidos, pessoas alocadas, etc, uma constatação que vem no próprio discurso do fornecedor. É fundamental a Administração de Dados para evitar o caos.

Devido a novos fatos que são trazidos novamente a luz, este discurso deve ser novamente retomado, ainda que com outra roupagem. Por serem verdadeiros, os objetivos de AD, às vezes estendidos, outras vezes na sua própria essência, sempre são alvo de pesquisas e reflexões. Convém um trabalho de conciliação entre a realidade dos propósitos da AD com os idealismos dos discursos acadêmicos tanto no fórum filosófico (para fundamentação e visão) como no fórum científico (para experimentação, confrontação e viabilização).

Um tema que vem alcançando grande notoriedade, inicialmente desenvolvido por pesquisadores de IA (Inteligência Artificial), e mais recentemente desenvolvido na área de ciências humanas (mais notadamente Economia, Administração, Direito) gira em torno de gerência de conhecimento, capitalização de conhecimento, capital intelectual, etc. Entendemos este interesse como uma evolução dos trabalhos remanescentes que almejavam prover uma teoria, conceitos e instrumental para Administração de Dados, com um escopo convenientemente mais alargado. Obviamente, estes novos estudos se baseiam nas possibilidades e restrições tecnológicas atuais, estando muito mais ajustados à nossa realidade. Portanto, devem substituir os preceitos antigos e concentrar esforços para a definição de um sólido arcabouço teórico que subsidie novos empreendimentos para uma efetiva administração de dados, informação, conhecimento em uma organização, como instrumentos para que essa instituição se torne mais inteligente. Investimentos significativos tem sido feitos na década de 1990, quando iniciou surgimento de fóruns de discussões que congregam especialistas do mundo todo. Por exemplo, desde 1993 existe um evento anual para tratar destes assuntos, o ISMICK (International Symposium for the Management of Industrial and Corporate Knowledge), e muitos outros eventos tem sido promovidos cada vez com maior especificidade e lucidez.

3. Alguns Fundamentos

Neste artigo pretendemos iniciar um corpo de conhecimento para futuramente discutirmos uma abordagem para o tratamento de dados, para produzir informações de qualidade, agregar sucessivamente conhecimento ao acervo corporativo, e capitalizá-los como patrimônio da Instituição, para que esta se torne mais “inteligente” no seu ambiente.

Para trabalharmos sobre estes assuntos, convém iniciarmos com a definição de alguns conceitos chave que serão utilizados a seguir [FUKU95]:

a) Dados: Percepção da realidade expressada em símbolos (palavras, etc) ou imagens (figuras, etc) que sejam tornados persistentes, geralmente através de registros em arquivos e bancos de dados.

b) Informação: [Dado + Seletor]; Seletor = [interesse do sujeito].
Informação é uma forma de dados, adquirido como resultado de uma seleção de dados que sejam úteis para a atividade que o sujeito esteja interessado em desempenhar. Está relacionada com a quebra da incerteza. Uma informação deve ajudar na eliminação de incertezas.

c) Conhecimento: [Informação + Refinaria]; Refinaria = [sistematização para reuso].
Conhecimento é o resultado de um refinamento de informação. Neste processo, informações com potencial para serem utilizadas novamente são sistematizadas.
Nota: Esta definição está voltada para o conhecimento dito Explícito. Existe ainda o conhecimento implícito ou tácito, que não podem ser simbolizados (traduzidos em palavras ou outro sistema de representação física).

d) Teoria: É a explicitação do conhecimento generalizado, organizado.

e) Inteligência : É a capacidade de aplicação dos meios disponíveis para resolução de problemas de maneira satisfatória (satisfatória, do ponto de vista do sujeito). Também se aplica ao aprendizado (que pode ser visto como um caso particular de problema)

Abaixo (figura 1) está uma representação dos níveis de abstração, cujos conceitos foram definidos acima e mostra que uma camada é construída a partir de outra inferior.

CONCEITO	AÇÃO SOBRE NÍVEL INFERIOR
Inteligência	Aplica teoria, conhecimento,... para resolver problemas.
Teoria	Generaliza, organiza Conhecimento
Conhecimento	Sistematiza Informação
Informação	Dá um significado aos Dados
Dados	Coleta e simboliza Realidade
Coisas e eventos da realidade	Existem no mundo real ou ideal

Figura 1 - Níveis de construção dos conceitos.

Na literatura corrente sobre KS (Knowledge System) "conhecimento, é a soma de percepções de um indivíduo sobre aspectos de algum denominado universo de discurso num tempo específico" ou seja, tudo que um indivíduo (ex. pessoa) sabe sobre um específico universo de discurso num dado tempo. Então neste contexto, um conhecimento deve estar sempre associado com estes três elementos: (um indivíduo, um domínio, um

tempo). Isto implica que por exemplo duas pessoas tem diferentes conhecimentos no mesmo domínio. Também que distintos domínios geram distintos conhecimentos, e que um conhecimento (de um indivíduo num domínio) muda com o tempo. Quando falamos de indivíduos, e isto inclui principalmente humanos, aspectos filosóficos tem importância, tais como crenças individuais e coletivas (tribais), que influenciam o conhecimento, e consciência. Do ponto de vista de KS, estas questões são deixadas de fora do seu escopo [MATT91]. Entretanto, num ambiente do mundo real, não poderão ser ignoradas.

Conhecimento, para nós pessoas, é importante porque praticamente tudo que fazemos está de alguma maneira baseado no conhecimento que temos armazenado na memória. As pessoas aplicam conhecimento para atingir seus objetivos, portanto, o processo de resolver problemas é uma rotina de aplicação de conhecimento para encontrar soluções. A aplicação do conhecimento é uma das quatro atividades no ciclo cognitivo: uma série de atividades por que passa o conhecimento na resolução de problemas (figura 2).

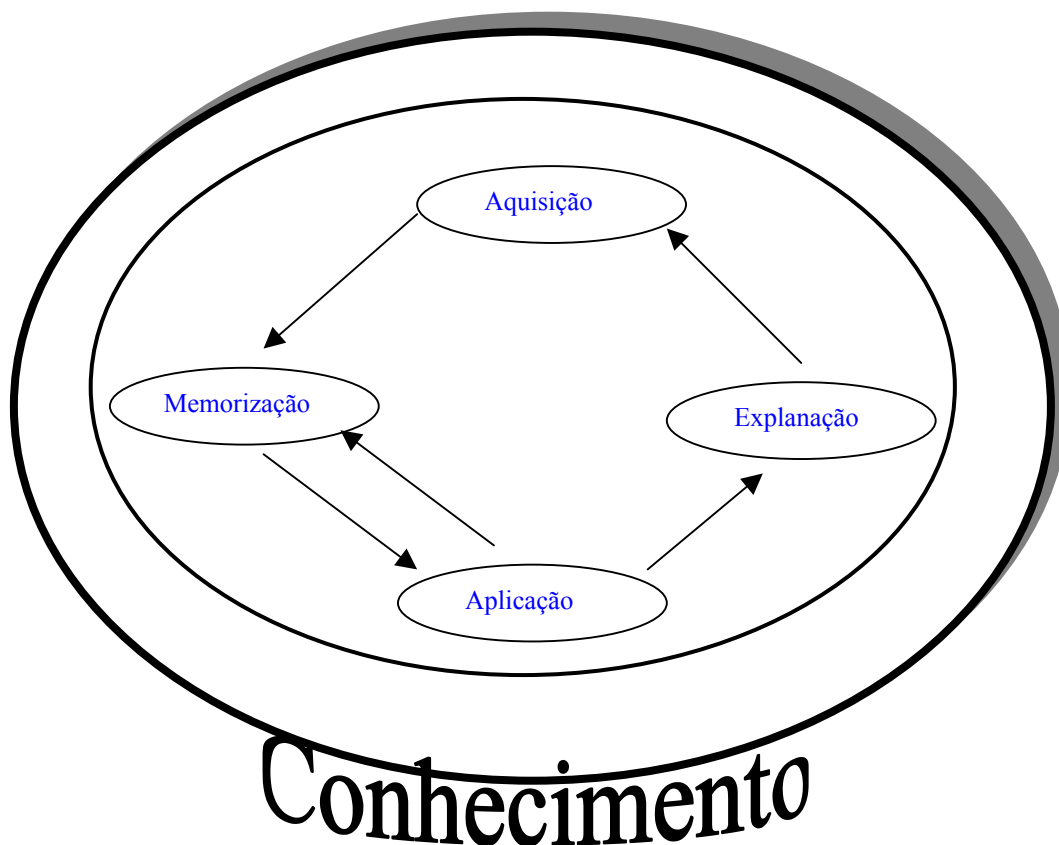


Figura 2 - Ciclo Cognitivo

A primeira fase importante neste contexto é o processo de aquisição de conhecimento, ou seja, uma metodologia de enriquecimento de uma "expertise" (conjunto de conhecimento usado para resolver problemas). Envolve coletar informações de muitas fontes e assimilá-las. Entra então a segunda fase que é a representação e organização deste conhecimento numa memória para que fique disponível para ser aplicado quando necessário: é o processo de memorização. Uma vez memorizado, está pronto para ser aplicado na resolução de problemas, pelo pensamento que é a transformação ativa de conhecimento memorizado, para criar um novo, que é usado para conseguir o objetivo

(resolver o problema). Automaticamente, este conhecimento produzido é assimilado e também memorizado. Entretanto, é importante a transmissão deste conhecimento para outros indivíduos, o que implica na explanação de como este conhecimento foi adquirido. À medida em que este conhecimento está sendo explanado, uma vez mais estará sendo adquirido, de modo que todo o ciclo se completa (figura 2) [MATT91]. Isto significa que, ao refletirmos sobre um domínio de conhecimento, estamos oportunizando a aquisição de novos conhecimentos.

Muitas teorias e métodos tem sido empregados para representação de conhecimento. Desde o desenvolvimento de vários paradigmas de lógica matemática, metodologias de inferências, até linguagens de programação, métodos e técnicas para representação de domínios específicos de conhecimento em máquinas. Sistemas Especialistas são exemplos de aplicação, onde se busca representar conhecimento de um especialista num domínio de problema. Mas esta área é muito mais ampla, e soluções tecnológicas tem sido buscadas para viabilizar grandes bases de conhecimento baseadas em computador. Classicamente temos duas maneiras de representação de conhecimento em computadores: sistemas convencionais, que se compõe de Algoritmos e Dados e sistemas ditos especialistas ou baseados em regras, que se compõe de Dados, Regras e Controle.

Os sistemas ditos convencionais tem ainda um uso maior, e são amplamente aplicados para automatização de processos, recuperação de informações e outras aplicações específicas. Por isso, teve um desenvolvimento significativo no desenvolvimento de tecnologias para gerenciamento de dados. Esta classe de tecnologia continua merecendo interesse e investimentos, propiciando a oportunidade de evolução para disponibilização de meios para gerenciamento de dados de propósito geral, de modo a contemplar requerimentos de aplicações baseadas em regras. Com isto, poderão ser produzidos gerenciadores de Bases de Conhecimento, que estarão em camadas acima da camada de gerenciamento de dados, mas com múltiplas funcionalidades similares e expandidas para prover transparência de uso e controle.

Vale dizer que sempre haverá a presença humana, como ingrediente fundamental no ciclo cognitivo de um sistema maior, talvez chamado de corporação, organização, etc. O próprio conhecimento de como tornar persistente o conhecimento em sistemas, tais como metodologias, sistemas de representação simbólica, etc, são tipos específicos de conhecimento. Por exemplo o conjunto de conhecimentos necessários para aprender matemática é muito diferente do conjunto de conhecimentos necessários para aplicar conhecimentos de matemática, ou de como trocar o pneu de um carro, ou decidir sobre oportunidades de negócio. Até mesmo a geração de taxonomias de conhecimento envolve a aplicação de conhecimentos específicos. Logo, o objetivo de criar mecanismos para gerenciamento de conhecimentos é uma ambição muito grande, envolvendo muita complexidade. Entretanto, pela relevância de seus efeitos, é um objetivo a ser buscado, ainda que os resultados sejam alcançados paulatinamente, a partir da organização dos elementos que o compõe. Por exemplo, uma evolução na gestão de dados persistentes em arquiteturas corporativas com capacidade de tratamento de dados para aplicações heterogêneas já será um grande passo.

Outro passo importante será a categorização do conhecimento como um bem, que compõe a principal parte do patrimônio de uma organização. O entendimento deste fato pelas autoridades de uma corporação poderá viabilizar os trabalhos necessários para o bom gerenciamento dos conhecimentos da organização, gerando um novo tipo de conhecimento: O conhecimento sobre si mesma.

4. Administração do Conhecimento

4.1 Capital Intelectual

A teoria neoclássica de crescimento econômico resumia este fator (de crescimento) a duas variáveis: Força de Trabalho e Capital. Na Segunda metade da década de 1980, entretanto, esta versão passou a ser contestada, por uma teoria que enfatiza outros elementos tais como Informação, Conhecimento, Prática (skill) como uma visão mais abrangente de capital do que meramente o capital físico [STEW94, STEW95, STOK95]. Esta parte é chamada de Capital Intelectual. O capital intelectual é considerado um produto interno (endógeno), que é gerado por um processo no estilo industrial do tipo "refinaria" de informações. Para tanto, utiliza como recurso primário: experiência prévia, conhecimento estabelecido, e um conjunto de informações não estruturadas (possivelmente oriundas de um conhecimento tácito). Assim, a construção de um conjunto de produtos de conhecimentos variados e integrados, formam o capital intelectual de uma organização. Na figura 2 é apresentado uma comparação entre as duas abordagens[STOK95].

Teoria Neoclássica	Teoria de Crescimento Endógeno
Foco = Capital Físico Parte tangível	Foco = Capital Intelectual Parte Intangível
Capital = [(oferta / demanda) + Valor agregado]	Capital = [Desenvolvimento Tecnológico + Conhecimento Acumulado]
GDP = [A * f (Qc, QI)]	GDP = [f (Qk, Qc, QI)] Obs. Qk = 80%; (Qc+QI) = 20%
Glossário: GDP = Gross Domestic Product A = Constante Qc = Capital QI = Trabalho	Glossário: GDP = Gross Domestic Product Qk = Conhecimento Qc = Capital QI = Trabalho

Figura 2 - Comparação entre a abordagem Neoclássica e a nova abordagem de crescimento endógeno.

Levando em conta a nova abordagem (Teoria de Crescimento Endógeno), temos duas variáveis para capital: O Capital Físico (CF) e o Capital Intelectual (CI). O Patrimônio da Organização (PO) é então a soma de ambos: PO = (CF + CI). A necessidade de definir uma medição para contabilização dos elementos de CI tem sido objeto de estudos, e demonstra dificuldades devido ao seu aspecto eminentemente intangível. Considerando sua peculiaridade endógena, os critérios de medição buscados são definidos como sendo conhecimento, competitividade e criação de valor agregado, e que vão influenciar todos os fatores externos, tais como Capital Físico (baseado na lei da oferta e procura) e o negócio ou missão da organização.

Um problema advindo disso é que o capital intelectual não pode ser considerado um "commodity" (produto) como outro qualquer no domínio do capital físico. Por exemplo, quando uma pessoa (A) transfere um produto físico para outra pessoa (B), este produto passa a pertencer apenas à segunda pessoa (B), ficando a primeira pessoa (A) sem acesso a este bem. Entretanto, com conhecimento, é diferente. Quando uma pessoa (A) transfere um produto intelectual para outra pessoa (B), este produto passa a pertencer a ambos (A e B), uma vez que a primeira pessoa (A) não perde o acesso a ele. Este é um exemplo das implicações que o paradigma de crescimento endógeno traz para ser tratado no domínio da engenharia econômica, e requer reflexão em outros domínios pois requer uma nova álgebra.

Um outro problema está em como medir objetivamente tais objetos intangíveis. Uma das soluções apresentadas para medição de capital intelectual [STOK95] busca valorar a contribuição que este pode dar (valor agregado) à competitividade de uma organização. Para isso, tem que ser mapeados os fatores de competitividade, seus requerimentos, e então associar com a demanda de conhecimento envolvido, almejando uma definição da estrutura de conhecimento requerido. Esta é uma área emergente para o desenvolvimento de metodologias e novos conceitos tanto no domínio das ciências tecnológicas, como das ciências Econômicas, de Administração, Contábeis, Psicológicas e Antropológicas.

4.2 Bases de Conhecimento Corporativo

As organizações modernas estão se tornando "baseadas em conhecimento" (Knowledge based organization). Novas experiências tem demonstrado que a habilidade de cooperação (via compartilhamento de conhecimento) tanto entre organizações, como entre diversas áreas de uma organização, ou mesmo entre indivíduos e fronteiras profissionais pode ser o bem mais importante que a organização pode possuir [STEW94, STEW95, STOK95, HALP98]. Este atributo pode ser um indicador e um fator crítico de sucesso para uma organização. Entretanto, fatores culturais e sistemas ultrapassados de distribuição de créditos são os principais inibidores destas atitudes. Não havendo incentivo ao compartilhamento de conhecimentos, qualquer base de conhecimentos será muito fraca [HALP98].

Não é novidade que os grandes gargalos de uma organização advém de causas psicológicas e organizacionais e não de limitações tecnológicas [HALP98]. A própria dinâmica de distribuição de cargos e autoridade numa organização tem um empirismo caótico, raramente comprometida com o mérito, contrariando o próprio mito que o sistema propaga: "do crescimento na hierarquia, por merecimento" [ADAM96]. Ao lado de políticas de avaliação de desempenho individual com critérios de retribuição do mérito, quando existentes, geralmente suspeitos na aplicação, estes fatos suscitam uma crise de confiança, não declarada, que permeia toda a organização. Isto é suficiente para inibir o compartilhamento de informação no dia a dia, quanto mais a motivação para construir bases corporativas de conhecimento. Entretanto, este problema transcende a nossa possibilidade de superação. Está fora de nossa área de competência resolvê-lo, pois advém da complexidade da natureza humana, logo, tem embutido em si leis naturais, físicas e antropológicas. É, pois, uma constante, ou uma variável no problema, que não deve ser ignorada. Deve ser, portanto, um requerimento de qualquer enunciado para empreendimentos corporativos, que condições de contorno sejam desenvolvidas para minimizar estes efeitos.

O conhecimento corporativo, é formado por todo seu acervo (independente de mídia) de documentos, bases de dados e sistemas de informação da Organização, e ainda o conhecimento de humanos que nela trabalham como gerentes ou agentes numa atividade qualquer. Gerentes são agentes em atividades de controle sobre outros agentes. Os

agentes possuem conhecimentos explícitos, que podem ser formalizados e transmitidos, e também conhecimentos implícitos ou tácitos, sendo que muitas vezes nem o próprio sujeito tem consciência acerca deste conhecimento. Em algumas atividades e momentos, estes dois tipos de conhecimentos se fundem podendo agregar valor significativo a uma classe de solução, perceber riscos ou oportunidades ocultos ou criar coisas inovadoras e surpreendentes. Estas são habilidades humanas difíceis de serem imitadas por outros meios, mas são a essência de muitas áreas de pesquisa em IA (Inteligência Artificial).

Um dos objetivos da gestão e capitalização de conhecimento é transferir o conhecimento de humanos para uma base de conhecimentos formal. Seja diretamente a partir da explicitação dos conhecimentos pelo próprio sujeito, ou pela garimpagem e compilação de diferentes tipos de informações, oriundas de documentos. Outra forma, é a construção de conhecimentos a partir de informações obtidas de bases de dados, ou ainda de regras de negócios que estão embutidos nos programas de aplicações, ou associados aos dados em gerenciadores de bases de dados. Entretanto, o maior objetivo não deve ser a transcrição das diversas representações de conhecimento, de uma forma para outra, mas a construção de um referenciamento do conhecimento, ou meta-conhecimento. Em outras palavras, conhecer o conhecimento que a organização possui.

Neste ponto, surgem os problemas de tangibilidade do conhecimento. As abordagens que tem sido tentadas, formalmente neste contexto, são baseadas em soluções simples, como estruturação de relatórios contendo histórico de resolução de problemas num domínio de conhecimento [SIMO95] e uso de hiperdocumento (editores de textos com ligações automáticas de palavras com outros textos) [MARC95]. Uma observação destas abordagens é que elas se apresentam como um passo inicial para uma apropriação futura destes conhecimentos em outras formas de representação, e conseqüente aplicação de métodos de gerenciamento de conhecimento e raciocínio automatizados.

Uma outra abordagem bastante propagada é a apropriação de dados persistentes fartamente já disponíveis em todas organizações. Consiste em organizar estes dados de modo a permitir maior número de inferências, formando assim um grande potencial de informações a serem sistematizadas, principalmente por humanos, que poderão então registrar este conhecimento apropriadamente. Alguns destes conceitos foram cunhados por Inmon [INMO92], que propôs um modelo de referência para uma arquitetura de dados em uma organização. Esta abordagem implica em extrair dados de bases corporativas e formar uma nova base de dados chamada de Datawarehouse, orientado para Sistemas de tomada de Decisão em uma Organização. Um Datawarehouse (DW) é uma base de dados ditos primitivos ou atômicos, de toda a corporação. Os dados de um DW serão derivados para outras estruturas de dados contextualizados e orientados à extração de informações de interesse Corporativo, Departamental, ou Individual. Esta abordagem vem alcançando grande aceitação pelo fato de não afetar as bases de dados existentes, e influenciado pelo barateamento de dispositivos de armazenamento, e um forte marketing de produtos.

Geralmente associada ao conceito de Datawarehouse, está o conceito de Data Mining (garimpagem de dados) e KDD (Knowledge Discovering in Databases - Descoberta de Conhecimento em Bases de Dados) [FAYY96]. Entretanto, estes tem uma personalidade própria e um espectro ainda maior. Trata de um conjunto de conceitos e conhecimentos aplicados para extração automatizada de dados e conhecimentos em bases de dados existentes. Uma premissa é de que o valor dos dados armazenados depende de nossa habilidade de acessá-los e utilizá-los, e que devido ao crescente volume, é humanamente impossível fazê-lo manualmente. Na verdade esta área está mais voltada para a construção de ferramentas, utilizando inteligência para descoberta de conhecimento em vários meios onde estes possam estar persistentes. Tem interseção com várias outras

áreas de pesquisa, tais como: Bancos de Dados, Aprendizado por Máquina, Reconhecimento de Padrões, Estatística, Inteligência Artificial e raciocínio com incertezas, Aquisição de Conhecimentos para sistemas especialistas, Visualização de dados, etc. [FAYY96]

A abordagem tradicional para tratamento de dados tem sempre sido voltado para dados transacionais, ou OLTP (OnLine Transact Processing). Esta abordagem vem desde os tempos em que surgiram as tecnologias de Banco de Dados e Comunicação de Dados, que permitiram o processamento online. É, portanto, o paradigma atual, uma evolução sobre o paradigma anterior: o processamento BATCH (hoje quase esquecido). Uma nova abordagem para tratamento de dados, que vem sendo bastante empregada é baseada no conceito de OLAP (OnLine Analytical Processing). Se baseia numa fundamentação teórica para manipulação de conjuntos de dados, proposto por Edward F. Codd, em uma evolução dos conceitos da Teoria Relacional, proposta pelo mesmo autor na década de 1970 (e que revolucionou a tecnologia empregada para Sistemas Gerenciadores de Bancos de Dados). Esta nova proposta de Codd, demonstra que os dados podem ser modelados segundo uma lógica baseada na metáfora de dimensões através das quais os fatos poderão ser analisados a qualquer momento. Os dados são estruturados em dimensões, como tempo, espaço, taxonomias, etc. Por esta razão os Bancos de Dados criados com este propósito são chamados de BD Multidimensionais. As ferramentas (softwares) destinados a Sistemas de Apoio a Decisão, ou Inteligência Corporativa (Business Intelligence) tem aplicado tecnologias dentro destes conceitos e muitas são vulgarmente conhecidas como “ferramentas OLAP”. É um conjunto de ferramentas que permitem análise de dados multidimensionais, geralmente com interfaces muito agradáveis, empregando recursos visuais, como Tabelas de várias dimensões, gráficos tridimensionais, etc.

Muitos estudos, formulando e reformulando conceitos, estão evoluindo e contribuindo para a formulação de uma teoria geral e até mesmo provendo ferramental para viabilização de uma base de conhecimentos em uma organização. Como já citado, os maiores volumes de dados persistentes com potencial para informação e formulação de conhecimento, estão em bases de dados ou datawarehouse. Estes dados guardam relação entre si, relações estas que se constituem em dados e regras e são representados segundo linguagens representacionais chamadas de modelos de dados. Modelos de dados são geralmente construídos utilizando alguma ferramenta CASE (Computer Aided System Engineering) e são armazenadas em Repositórios. Estes repositórios contém então conhecimento acerca de determinado ponto de vista da organização, que estão representados nos modelos. Estes modelos são representações de uma dada realidade, e contém uma alta densidade de informações sobre o contexto almejado. Como complemento do modelo existe uma coleção de elementos de dados que precisam ser conceituados. A descrição destes conceitos formam uma base de metadados que é a explicitação um domínio de conhecimento através de uma declaração sobre a interpretação sobre estes termos na organização. Uma evolução deste conceito nos leva à Ontologia, conforme veremos sucintamente a seguir.

4.3 Ontologia

Um corpo de conhecimento representado formalmente é baseado numa conceitualização: Os objetos, conceitos, e outras entidades que são assumidas para existir em alguma área de interesse e os relacionamentos que guardam entre si. Uma conceitualização é uma visão abstrata simplificada do mundo que desejamos representar para algum propósito. Toda base de conhecimento, sistema baseado em conhecimento ou agente (a nível de conhecimento) está de acordo com alguma conceitualização, explícita ou implicitamente [GRUB93].

Uma Ontologia é uma especificação explícita de uma conceitualização. Este termo vem da filosofia, onde Ontologia é uma contagem sistemática de existência. Para sistemas de IA (Inteligência Artificial), o que existe é o que pode ser representado. Quando o conhecimento de um domínio é representado em um formalismo declarativo, o conjunto de objetos que podem ser representados é chamado de universo de discurso. Este conjunto de objetos e os descritíveis relacionamentos entre eles são refletidos no vocabulário representacional com o qual um programa baseado em conhecimento representa um conhecimento. Assim, no universo de IA a ontologia de um programa pode ser descrita pela definição de um conjunto de termos representacionais. Assim, em tais Ontologias, definições associam nomes das entidades no universo de discurso com textos legíveis para humanos, descrevendo o que os nomes significam, e axiomas formais que assegurem a interpretação e o bem formado uso destes termos. Formalmente, uma Ontologia é a definição de uma teoria lógica. [GRUB93]

Usamos ontologias comuns para descrever acordos ontológicos para um conjunto de agentes de modo que eles possam se comunicar sobre um domínio de discurso sem necessariamente operar numa teoria globalmente compartilhada. Dizemos que um agente concorda com uma ontologia se suas ações observáveis são consistentes com as definições na ontologia. A idéia de acordos ontológicos é baseada na perspectiva do nível de conhecimento. O nível de conhecimento é um nível de descrição do conhecimento de um agente que é independente da representação do nível simbólico usado internamente pelo agente. Assim, conhecimento é atribuído aos agentes por meio da observação de suas ações. Um agente sabe algo se ele age como se tivesse a informação e agisse racionalmente para conseguir suas metas. Pragmaticamente, uma ontologia comum define o vocabulário com o qual perguntas e asserções são intercambiadas entre agentes. Acordos ontológicos são concordâncias para usar o vocabulário compartilhado de maneira coerente e consistente. Agentes compartilhando um vocabulário não precisam nem mesmo compartilhar a mesma base de conhecimento, porque se eles estão comprometidos com uma ontologia podem se comunicar e cada um responde o que sabe ao outro. [GRUB93]

Estas considerações sobre ontologia nos levam a perceber que são conceitos bastante genéricos e que podem ser praticados tanto no domínio restrito de sistemas cujas aplicações são escritas no paradigma de Multiagentes, como em outro paradigma qualquer. Mas podemos aplicar no sentido mais amplo para uniformizar a linguagem específica de áreas de negócio a exemplo do domínio de discurso, para intercâmbio de conhecimento de humanos e até para a especificação de requerimentos para o desenvolvimento de sistemas e construção de bases de dados convencionais.

5. Conclusão

Um passo importante será a categorização do conhecimento como um bem, que compõe a principal parte do patrimônio de uma organização. O entendimento deste fato pelas autoridades de uma corporação poderá viabilizar os trabalhos necessários para o bom gerenciamento dos conhecimentos da organização, gerando um novo tipo de conhecimento: O conhecimento da organização sobre si mesma.

Vale dizer que sempre haverá a presença humana, como ingrediente fundamental no ciclo cognitivo de um outro sistema chamado de corporação, organização, etc. O próprio conhecimento de como tornar persistente o conhecimento em sistemas, tais como metodologias, sistemas de representação simbólica, entre outros, são tipos específicos de conhecimento. O objetivo de criar mecanismos para gerenciamento de conhecimentos é uma ambição grande, envolvendo muita complexidade. Entretanto, pela relevância de seus efeitos, é um objetivo a ser buscado, ainda que os resultados sejam alcançados paulatinamente, começando por metas modestas, como a organização dos dados,

melhoria dos sistemas de informações, e formalização do conhecimento, explicitando o conhecimento humano, extraindo conhecimento através de heurísticas aplicadas nas bases de dados existentes, e integrando esses conhecimentos em bases de conhecimento com arquitetura bem projetada. [Vale lembrar os conceitos no início deste artigo, onde inferimos que os dados são os elementos atômicos constitutivos das estruturas cada vez mais complexas, até a formulação do conhecimento. Portanto, a gestão de dados é condição primária necessária para se alcançar os benefícios em estruturas mais complexas como a gestão de conhecimento.](#) Esta pode ser uma alternativa de atitude para a preparação dos ambientes informatizados de uma nova era. Do contrário seremos atropelados, como tem sempre acontecido a exemplo da assimilação da tecnologia para gerência de dados e do dispendioso saneamento contra o “bug do milênio”.

6. Referências Bibliográficas

- ADAM96 - ADAMS, Scott. O PRINCÍPIO DILBERT. 5ª Edição. Tradução de Texto Talita M. Rodrigues, Ediouro Publicações, Rio, 1996.
- COST95 - Costello, Robert B. and Posner Deborah M. Webster's College Dictionary. Random House Inc, New York, USA, 1995.
- FAYY96 - FAYYAD, Usama; SHAPIRO, Gregory Piatetsky; SMITH, Padhraic. The KDD Process for Extracting Useful Knowledge from volumes of Data. Communications of the ACM, Novembro/1996.
- FERR75 - Ferreira, Aurélio Buarque de Holanda. Novo Dicionário da Língua Portuguesa. 1. Edição, 15. Impressão. Editora Nova Fronteira, Brasil, 1975.
- FUKU95 - FUKUDA, Yutaka. Variation of Knowledge in Information Society: in SMICK'95. Institut International pour L'intelligence Artificielle. Compiègne, France, 1995.
- GRUB93 - GRUBER, Thomas R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Technical Report KSL93-04, Knowledge Systems Laboratory, Stanford University, Revisão 23/agosto/1993.
- HALP98 - HALPER, Mark. Integrando o Pool de Conhecimento: in Computerworld Global Innovators, 12/janeiro/1998.
- INMO92 - INMON, William H. DATA ARCHITECTURE: The Information Paradigm. 2ª edição, QED Technical Publishing Group, USA, 1992.
- MARC95 - MARCHAND, Yannick; GUÉRIN, Jean-Luc. The use of Hipertext to organize Knowledge in the Domain of Telecommunications: in SMICK'95. Institut International pour L'intelligence Artificielle. Compiègne, France, 1995.
- MATT91 - MATTOS, Nelson Mendonça. An Approach to Knowledge Base Management. Lectures Notes in Artificial Intelligence. Kaiserslautern, Alemanha, 1991.
- ORTE94 - ORTÊNCIO, Vanderlei Vilhanova. Fatores Críticos de Sucesso para a Administração de Dados: Um Estudo Preliminar. Curitiba, 1994.
- ORTE98 - ORTÊNCIO, Vanderlei Vilhanova. Evolução de Tecnologia aplicada a Gerenciamento de Dados Persistentes. Curitiba, 1998.
- SIM95 - SIMON, Gaële; GRANDBASTIEN, Monique. Corporate Knowledge: a case study in detection of metallurgical flaws: in SMICK'95. Institut International pour L'intelligence Artificielle. Compiègne, France, 1995.
- STEW94 - STEWART, Thomas A. Your Company's Most Valuable Asset: INTELECTUAL CAPITAL; in FORTUNE October 3, 1994.
- STEW95 - STEWART, Thomas A. Trying to Grasp the Intangible; in FORTUNE October 2, 1995.
- STOK95 - STOKKE, Per R. Endogenous Growth Theory and the Management of Industrial and Corporate Knowledge: in SMICK'95. Institut International pour L'intelligence Artificielle. Compiègne, France, 1995.