

Capítulo de livro publicado

TERRA, J. C. ; BAX, M. P. Portais corporativos: instrumento de gestão de informação e de conhecimento. In: Isis Paim. (Org.). A Gestão da Informação e do Conhecimento. 1 ed. Belo Horizonte, 2003, v. , p. 33-53.

Portais corporativos: instrumento de gestão de informação e de conhecimento

José Claudio Cyrineu Terra

jcterra@yahoo.com

Professor MBA em Gestão do Conhecimento – FEA/USP e PUC/PR

Marcello Peixoto Bax

bax@eci.ufmg.br

www.bax.com.br

Escola de Ciência da Informação

1 Introdução

Os portais de informação empresarial ou portais corporativos assumem importância cada vez maior para os negócios, traduzindo todo o manancial de conteúdo informacional proveniente das variadas fontes da empresa em informação útil para a tomada de decisões nos três níveis: estratégico, tático e operacional.

O portal corporativo tornou-se importante ferramenta para as empresas, principalmente as grandes, que competem em mercados globais onde o acesso eficiente a informações é requisito básico.

Embora alguma confusão ainda persista sobre o significado exato do conceito “portal corporativo”, muitas empresas estão implementando ou planejam implementar tais soluções. Este artigo, além de colaborar para esclarecer o conceito, destaca os benefícios, vantagens e potencialidades, decorrentes da implantação de um portal corporativo, considerando-o como um novo instrumento de gestão de informação.

Focalizam-se principalmente os aspectos relativos à melhoria do acesso à informação corporativa e a minimização do problema de excesso de informações, ou seja, como os portais corporativos lidam com as dificuldades de acesso à informação e, ao mesmo tempo, com a sobrecarga de informações.

Além disso, analisam-se, em particular, a importância do processo de personalização e o papel dos mecanismos de busca nos portais corporativos. Descrevem-se ainda problemas envolvidos no processo de categorização e taxonomia, fundamentais para a organização da informação espalhada por toda a empresa.

2 Definindo portais corporativos

Os portais corporativos, também chamados de *EIP's (Enterprise Information Portals)*, são aplicações visualmente similares aos portais encontrados na Internet. Embora, em geral, sejam aplicações mais complexas que encontram justificativa no apoio à missão, às estratégias e aos objetivos da organização e colaboram para a criação e o gerenciamento de um modelo sustentável de negócios.

Seu objetivo primário é promover eficiência e vantagens competitivas para a organização que o implementa. A idéia por trás desses portais é a de desbloquear a informação armazenada na empresa, disponibilizando-a aos utilizadores através de um único ponto de acesso. Esse ponto de acesso único, que lhe confere o signo de “portal”, disponibiliza aplicações e informação personalizadas, essenciais para a tomada de decisões nos níveis estratégico (de negócio), tático e operacional.

Os portais corporativos são instrumentos essenciais ao esforço, cada vez mais importante, em se compartilhar informação e conhecimento no seio das organizações.

Representam mudança necessária no sentido de se estabelecer uma plataforma única para o *e-Business* (B2C, B2B, etc), podendo também prover a integração, em tempo-real, de diversos sistemas de informação; o que representa mudança substancial no modo de como será apresentada a informação e em como os muitos sistemas de *back-end* precisarão comunicar entre si.

A capacidade dos portais corporativos em capturar, organizar e compartilhar informação e conhecimento explícito é interessante especialmente para empresas intensivas em conhecimento. O Portal Corporativo deve permitir a superação de alguns desafios relacionados ao gerenciamento estratégico de informação presentes no ambiente corporativo pré-Web e pré-Portal. Esses desafios incluem, para citar apenas alguns:

- Presença de sistemas não integrados e formatos de arquivos proprietários e incompatíveis;
- Dificuldade de acesso ágil à informação atualizada; e, ao mesmo tempo, sobrecarga de informação;
- Redundância e duplicação de informações através das redes;
- Informações e documentos publicados de modo desorganizado, sem controle de fluxo de aprovação;
- Diversidade de caminhos, métodos e técnicas diferentes para buscar e acessar a informação;
- Dificuldade para as pessoas publicarem informações acessíveis à empresa como um todo;
- Dificuldade de definição ou ausência de políticas de segurança;
- Usuários “não técnicos” excessivamente dependentes do departamento de TI para gerar, divulgar e obter informação;
- Arquiteturas proprietárias e caras dificultam a integração de diferentes tipos de informação.

Além dos pontos acima, muitos outros poderiam ser citados para ilustrar o nível de complexidade da gerência de informação em um ambiente corporativo. Sem os instrumentos de apoio, sejam tecnológicos ou gerencias, o problema é dificilmente suportável no médio prazo; a organização pode não sobreviver à concorrência.

Outro risco que vale a pena ser mencionado é que tal complexidade dificulta a definição de políticas de segurança, podendo levar a organização a sofrer ataques tanto externos quanto internos. Isso leva algumas organizações a simplesmente bloquearem diversos canais de comunicação, procurando fugir do problema, porém regredindo às vezes até ao estado anterior à implantação das redes locais.

As soluções de EIP orientam-se pelo fato de estarem as organizações atuais desafiadas a se tornarem mais ágeis, hábeis e inteligentes para obterem sucesso. Com efeito, as empresas estão sob imensa pressão dos desafios e oportunidades da economia atual,

que evolui em ritmo acelerado. Diante de vasta quantidade de fontes de informação para gerenciar, expectativas crescentes dos clientes, equipes cada vez mais multidisciplinares, os colaboradores precisam ter acesso à informação relevante e personalizada – em seu contexto apropriado – para a tomada de decisão bem embasada.

3 Minimizando a dificuldade de acesso às informações

3.1 Integrando dados estruturados e não estruturados

Um portal corporativo provê um ponto central de acesso aos recursos de informação de uma empresa (bases de dados e sistemas de informação, documentos, e-mails, etc) e utiliza metadados XML para integrar dados estruturados e não estruturados espalhados pela empresa.

Dados estruturados estão presentes em banco de dados e arquivos usados por sistemas transacionais recentes e legados e constituem, em geral, apenas uma pequena parte (algo em torno de 10%, em média) dos dados, informações e conhecimentos relativos aos negócios de uma organização; o restante é composto por dados não estruturados presentes em documentos, relatórios, gráficos e imagens, ou em formatos de áudio e vídeo. Tais fontes de dados não estruturados não são facilmente acessíveis aos sistemas de informação que utilizam banco de dados relacionais. Os portais corporativos utilizam metadados XML para integrar ambas as categorias de dados (estruturados e não estruturados) para possibilitar o acesso integrado a todas as fontes de informação.

3.2 Ponto único de acesso generalizado e personalizado

Outra vantagem percebida com a instalação do portal corporativo reside no fato de proporcionar, como já mencionado, um recurso único centralizado onde os usuários podem encontrar, extrair e analisar toda a informação que prolifera no ambiente corporativo, muitas vezes geograficamente distribuído. Além disso, ele expande o âmbito atual da informação corporativa para um cenário em que os usuários finais, que consomem a informação, possam ser tanto empregados, como clientes ou fornecedores.

Nesse contexto, os usuários da informação corporativa se beneficiarão de dados e informações, transferindo-os, minerando-os e reutilizando-os em diferentes cenários e aplicações.

As empresas, por sua vez, poderão configurar cada usuário do portal de acordo com níveis de segurança e autorizações de acesso individuais. Os usuários têm a possibilidade de personalizar seus portais para encontrar, aceder e pesquisar mais facilmente os recursos disponíveis na empresa. Muitos usuários já estão familiarizados com a personalização oferecida atualmente pela maior parte dos portais Internet.

Ao incluir softwares de análise de informação estruturada e permitir o acesso mais rápido à informação não/semi estruturada relevante e precisa, a partir de um ponto de acesso único, o portal apóia o processo de tomada de decisão na empresa.

3.3 Várias aplicações acessíveis de um único sistema

Segundo artigo da companhia Merrill Lynch¹, de novembro de 1998, os portais são constituídos de amálgamas de ferramentas de software que analisam, consolidam, gerenciam, e distribuem informação no seio das empresas e para o mundo exterior. A mesma consultoria, num estudo exaustivo sobre esse assunto elaborado em 1998, dividia os EIPs em três segmentos de produtos: sistemas de “Gestão de Conteúdo”, “Armazéns de dados (*Data Warehouses*)” e “Inteligência de Negócios (*Business Intelligence*)”. Nessa época, o artigo descritivo do estudo previa que esses segmentos de software convergiam para uma única aplicação de informação empresarial.

Aos sistemas de gestão de conteúdo cabe a tarefa de capturar, arquivar, indexar e distribuir informação externa e interna para a criação de um repositório de informação corporativo. Os sistemas de *Data Warehouse* ou *Data Marts* criam um ambiente de armazenamento onde os dados são orientados a diversas análises de desempenho corporativo. São usados para suportar soluções de BI corporativas complexas. Por outro lado, as aplicações de Business Intelligence (BI) potencializam as fontes de bases de dados empresariais de forma a que toda a empresa tenha disponível informação precisa e direcionada. Esses sistemas incluem funcionalidades de processamento analítico (*On-Line Analytical Processing*, ou OLAP), mineração de dados (*Data Mining*) etc.

3.4 Além dos dados e informações: processos

Enquanto os armazéns de dados focalizam sobre dados e informação, o portal corporativo deve também apoiar os processos e fluxos de trabalho (*workflows*). Dessa maneira ele provê um ponto central de acesso via navegador Internet a todos os sistemas que cada colaborador da empresa precisa operar e consultar no seu dia-a-dia.

Um desafio para a implementação bem sucedida de um portal é a relativa pobre organização das informações corporativas, o que atrapalha o trabalho eficiente em conjunto. Capacitados para apoiar a identificação e captura, armazenamento e recuperação de informação de múltiplas fontes, os portais corporativos, ao invés de aplicações isoladas que operacionalizam assuntos isolados, são ambientes que proporcionam uma visão de todo, fundamental para as definições em nível estratégico de qualquer negócio.

¹ Organização de pesquisa e consultoria econômica.

4 Lidando com o excesso de informações

4.1 Definindo o problema

Segundo Bawden (2001), sobrecarga informacional (“*information overload*”) é um estado no qual a informação disponível e potencialmente útil torna-se um obstáculo (ou atraso) ao invés de uma ajuda. O excesso de informação está associado à perda de controle sobre a informação e à incapacidade em usar efetivamente a informação. Como resultado tem-se trabalho ineficiente e eventualmente até risco para a saúde.

A questão não é nova, existem registros na literatura que atestam referência ao problema em 1852 (Rada 1991). Segundo alguns autores, por volta dos anos 50 do século XX a própria ciência corria risco. Porém, apenas ao final dos anos 50 e início dos anos 60 do século XX é que o estado de sobrecarga informacional é compreendido como um problema real. Nos anos 90 torna-se um problema grave, particularmente no mundo dos negócios, com a influência cada vez maior de novas tecnologias de informação e comunicação: celulares, palms etc; e particularmente a Internet provendo correio eletrônico, *chats* e Web. Como indica Shenk (1997), trata-se de um problema importante para indivíduos e organizações e as técnicas usadas no passado não são mais efetivas hoje. Atualmente o problema afeta diretamente a eficiência do trabalho e a produtividade das organizações.

Pesquisa da Reuters (Bird 1997) junto a gerentes revelou que a sobrecarga informacional, para dois terços deles, tinha causado perda de satisfação no trabalho e comprometido suas relações pessoais. Quase a metade acreditava que decisões importantes tinham sido atrasadas e afetadas de forma adversa pelo excesso de informação.

O relatório da Reuters, citado por Bawden (2001), reconhece o problema e argumenta que muitas informações, se mal organizadas, atrapalham o conhecimento. Com efeito, por um lado parece que o grande desafio para as empresas hoje é como gerenciar as informações de que dispõem, e não propriamente como obter novas informações.

Algumas empresas estariam tratando o conhecimento como “*commodity*” e não como um processo de melhoria da inteligência empresarial como um todo, que se traduziria mais apropriadamente em habilidades superiores de interpretação dos fatos e de tomada de decisão. Isto não quer dizer que as empresas não devam ter uma gestão pró-ativa do conhecimento. Pelo contrário, essa gestão deve envolver mudanças nos indicadores utilizados para medir o desempenho dos funcionários, refletindo alterações mais profundas na cultura organizacional e nas práticas gerenciais.

Neste mesmo sentido, Stewart (1998) cita o trabalho dos sociólogos Sara Keisler e Lee Sproull publicado em seu livro *Connections* (Sproull, 1993), mostrando que, paradoxalmente, o desejo de parecer racional tem levado a um comportamento não racional, que se traduz em uma preocupação excessiva com a informação por parte das pessoas. Outros autores, por sua vez, também distinguem conhecimento de informação.

Conhecimento é informação interpretada, o que faz com que a simples transferência de informação não aumente o conhecimento ou a competência (Nevis et al, 1995; Nonaka, Takeuchi, 1995; Sveiby, 1997; Davenport, Prusak, 1998; entre outros).

Nas "organizações que aprendem", as informações e os sistemas de informação têm que ser precisos, imediatamente disponíveis no tempo e espaço, e têm que estar disponíveis em um formato que facilite o uso. Na prática, duas questões ajudam as empresas a desenharem seus sistemas de informação: Como determinada informação agrega valor a determinado processo de decisão? Como ela pode chegar às pessoas que precisam dela?

Embora o problema da sobrecarga informacional não seja novo, podendo ser encontrado de forma recorrente ao longo da história, suas maiores conseqüências são recentes. Segundo Bawden (2001), a maneira de se lidar com o problema no passado estaria ligada à forma de estruturação da organização, particularmente a organização hierárquica, de padronização e racionalização. Arquivos e gavetas para papel, formulários padronizados etc. Esses métodos são, aos poucos, superados em tempos de globalização intensa, com as organizações menos fortemente estruturadas utilizando novas tecnologias de armazenagem e comunicação. A *Forrester Research* realizou estudos atestando que a capacidade média de armazenamento está crescendo ao ritmo de 50% ao ano.

Assim, parece não existir uma técnica ou ferramenta que poderá resolver o problema como um todo. As soluções propostas aparentemente caem em duas categorias: gerencial e tecnológica.

Do ponto de vista gerencial trata-se de retomar o controle. Como visto, a perda de controle sobre a informação é o principal sintoma da sobrecarga. Retomar o controle, tanto no nível individual quanto no organizacional, pode ser parte importante do remédio. Controle não no sentido coercitivo, mas de alinhamento a um padrão que implique o progresso em direção a um objetivo planejado e pré-determinado.

4.2 Tentativas de quantificar o problema

Segundo estudo realizado pelo *Institute of the Future*, funcionários já recebem em média 192 mensagens/dia nas seguintes formas: email, cartas, voice mail, livros, revistas, etc. Entrevista do *Gartner Group* mostrou que os executivos de empresas norte americanas estão gastando em média cerca de 108 minutos/dia apenas lidando com mensagens de email. Outras pesquisas apontam, por exemplo, para o crescimento de 6 vezes no uso do email em corporações americanas nos últimos 5 anos. Na grande maioria dos casos as Intranets cresceram tão rapidamente a ponto de se tornarem difíceis de gerenciar e ineficientes. A HP, por exemplo, antes de implementar um EIP, tinha 4.700 URL's em sua Intranet. Existiam mais de um milhão de URL's na Intranet da Intel nos primeiros meses do ano de 2001.

Os EIP's vêm representando uma tentativa de se reduzir o excesso de informações que foi exacerbado pela generalização do uso de aplicações *desktop* nos últimos tempos, principalmente de programas de e-mail e web.

4.3 Sobrecarga de informação e os portais corporativos

Na tentativa de lidar com o problema do excesso de informações os EIPs focalizam suas funcionalidades em três aspectos particularmente relevantes: personalização, mecanismos de busca e categorização.

4.3.1 Personalização

Os EIPs procuram responder ao desafio crescente da sobrecarga informacional e da escassez de tempo, provendo níveis sofisticados de personalização. Personalização significa aqui melhor adaptação das ferramentas às necessidades dos usuários em acessar rapidamente a informação mais relevante para a execução das atividades profissionais do seu dia-a-dia.

Por exemplo, a partir de uma pesquisa com 22 mil funcionários, a *Johnson Controls* implantou um EIP e concluiu que a solução ajudou a economizar pelo menos uma hora por dia de cada funcionário. Ao invés de ter que abrir várias janelas diferentes e lançar as diversas aplicações cada vez que acessam suas estações de trabalho, os EIP possibilitam aos funcionários predefinir as aplicações que serão abertas simultaneamente a cada momento. Dessa forma, as várias aplicações mais freqüentemente usadas ficam acessíveis com um clique.

Os EIP são particularmente interessantes para pessoas que trabalham em múltiplos projetos simultaneamente e lidam com situações complexas que requerem variadas fontes de informação e opiniões. Os portais facilitam o compartilhamento e a visualização das mesmas aplicações simultaneamente por vários usuários, mesmo estando geograficamente muito distantes uns dos outros.

A personalização deve acontecer de forma transparente, adaptando e adequando o conteúdo e/ou a experiência individual com base em informações implícitas e explícitas, e incorporando-as a caminhos de navegação regulares que minimizam o esforço do usuário.

Os sites mais avançados combinam abordagens de personalização de dois tipos: *pull* (puxar) e *push* (empurrar). A abordagem *pull* é dependente de escolhas feitas pelos usuários, eles próprios comandam a personalização. A abordagem *push* pode variar de muito simples a muito sofisticada. As abordagens *push* mais simples consistem em personalizar o site de acordo com os perfis dos funcionários (seu papel na organização, sua localização física, seus direitos de acesso, etc). As mais sofisticadas abordagens de personalização "push" são dinâmicas, realizadas em tempo real. Baseiam-se no número de fontes de informações consultadas e acumuladas ao longo do tempo, incluindo os

caminhos de navegação, dispositivos usados para o acesso, posição geográfica atual do usuário e a largura de banda disponível, etc. Considerando tais informações, os mecanismos de personalização predizem de forma independente as preferências dos usuários, mostrando conteúdos com base em seus papéis, sua localização, e seu comportamento durante visitas ao portal.

Como a personalização pode ter níveis de sofisticação diversos (e variados orçamentos), as empresas devem decidir previamente que conteúdos devem ser permanentes para um grupo específico de usuários e que conteúdos podem ou devem ser personalizados. Idealmente, cada funcionário deveria ter um “papéis” associado ao uso que faz do EIP. Isso permite personalizações pré-definidas, configuradas com facilidade quando novos funcionários integram-se à empresa. Com o passar do tempo os funcionários desenvolvem uma experiência verdadeiramente personalizada com o portal. Além disso, essa prática permite que novos funcionários iniciem suas atividades acessando informações e *links* que são relevantes à realização de suas tarefas iniciais e responsabilidades. A associação de papéis também permite às empresas fazerem melhor avaliação de que níveis e áreas de personalização, colaboração, uso de informações e fontes de conhecimentos terão maior impacto no desempenho dos funcionários.

Portais sofisticados apresentam possibilidades interessantes de personalização. Dentre as mais avançadas, podem-se citar:

- Personalização da apresentação, na qual vários elementos das páginas adaptados em moldes (ou *templates*) podem ser alterados: *layout*, cor, fontes, etc;
- Permissão aos usuários para arrastar e modificar o tamanho das janelas;
- Criação de novos *templates* de páginas com o auxílio de assistentes e sem necessidade de programação;
- Personalização considerando uma combinação de atributos: papéis, regras de negócios, padrões de navegação, filtragem colaborativa, o dispositivo de acesso utilizado, a largura de banda disponível etc;
- Compartilhar perfis de funcionários por várias aplicações no portal.

4.3.2 Mecanismos de busca

Os mecanismos de buscas são ferramentas familiares a todos os que usam a Internet. Estatísticas mostram que centenas de milhões de pessoas usam mecanismos de busca diariamente. Em meados de 2001, os mecanismos Google e Altavista, líderes de mercado, relataram em média 100 milhões de buscas realizadas por dia. Segundo o *Nielsen/NetRatings* o Google obteve 341 milhões de visitas/dia em 2002. Obviamente, isso não implica que as pessoas estejam encontrando as informações de que necessitam, com a rapidez que seria desejável.

No contexto corporativo onde realizam suas atividades profissionais, os funcionários não têm tempo para “navegar” na web. Além disso, precisam procurar não apenas páginas

web, mas também um número variado de fontes de informação, incluindo informações estruturadas (em geral presentes nos bancos de dados) e não ou semi-estruturadas (aplicações de *desktop* como e-mail, documentos *word*, e Intranet, etc).

Um dos objetivos centrais da gestão do conhecimento e, por conseguinte, dos EIPs é auxiliar os funcionários a encontrarem facilmente as fontes de informações corporativas mais relevantes no momento exato em que mais necessitam delas. Nesse contexto fica claro que os mecanismos de busca são ferramentas extremamente úteis. Eles devem prover, no menor tempo, os resultados mais relevantes para uma dada consulta.

Entretanto, na maioria dos casos os usuários ainda precisam de assistência para a definição mais precisa do que desejam procurar, quais descritores usar? Como usar a pesquisa avançada? Além disso, devido às ambigüidades inerentes às línguas naturais, com freqüência documentos referem-se a assuntos ou conceitos similares usando diferentes termos e níveis de profundidade variados, o que pode gerar resultados com baixa precisão². É também importante notar que pessoas diferentes procurarão por informações de maneiras diferentes, conforme vários fatores, tais como:

- Nível de experiência com mecanismos de busca;
- Nível de conhecimento prévio em um domínio ou assunto específico;
- Tipo de personalidade: cérebro-esquerdo ou cérebro-direito³, visual ou textual; e
- Preferência por navegação ou busca direta.

Em resposta a essas variadas circunstâncias individuais e também ao crescimento exponencial da quantidade de informações, tipos de dados e fontes de conhecimento, ocorreram vários avanços interessantes em tais mecanismos nos últimos anos. Muitas opções de busca, funcionalidades e níveis de sofisticação estão hoje disponíveis. Essas funcionalidades são representativas da evolução dos mecanismos de buscas e podem ser organizadas, segundo (Broader, 2002), em três gerações:

- **Mecanismos de busca de 1ª geração:**
 - o Busca por palavra chave ou frase exata, funcionalidades muito comuns nos mecanismos de buscas que listam os resultados com base na ocorrência da palavra ou frase. Este tipo de mecanismo foi bastante usado em bibliotecas e permitem aos usuários focar suas pesquisas por áreas específicas dos documentos tais como, título, palavras-chave e autor;
 - o Busca booleana retorna resultados com base na aplicação de operadores lógicos, tais como AND, OR e NOT;

² Ou excesso de resultados, poucos resultados ou irrelevantes.

³ Tem sido discutido que a Web e outros sistemas hipermídia possuem modelo semelhante ao modo como o cérebro processa informação (Small e Ferreira, 1994). Pesquisadores descobriram a existência de relação entre estilos de aprendizagem (cérebro direito ou esquerdo) e abordagens no uso do hipermídia.

- Buscas com filtros colaborativos, baseados na idéia de que indivíduos que compartilham interesses comuns considerarão documentos similares relevantes. O módulo de sugestão de livros da livraria *Amazon* é um exemplo desse tipo de mecanismo;
- **Mecanismos de busca de 2ª geração:**
 - Buscas baseadas em popularidade nas quais se altera o *ranking* dos *links* que serão mostrados, aplicando-se vários algoritmos que avaliam a perspectiva dos usuários em relação aos documentos ou aos sites retornados. Pode-se, por exemplo, considerar o número de vezes que usuários prévios visitaram o link ou quantos estabeleceram *links* para o site específico (o *Google* e muitos outros usam tal tecnologia). Podem-se também apresentar os documentos baseando-se em uma avaliação pró-ativa do usuário com relação ao documento ou site;
 - Busca por conceitos retornam resultados que podem não conter as palavras da busca. Isso ocorre porque o mecanismo procura por palavras correlacionadas com base em um thesaurus. Essa funcionalidade representa uma evolução dos mecanismos de busca considerados de 3ª geração.
- **Mecanismos de busca de 3ª geração:**
 - Buscas contextualizadas: certos mecanismos de busca consideram o contexto em que se apresenta a necessidade informacional do usuário. Isso é muito útil quando as palavras possuem significados bastante diferentes dependendo do contexto em que são usadas (e.g. a palavra “leão” pode representar uma bebida ou um animal). A esse fenômeno lingüístico específico dá-se o nome de polissemia; existem diversos outros tipos de ambigüidades inerentes às línguas naturais. Marisa Bräscher (Bräscher, 2002) apresenta um estudo aprofundado sobre esses fenômenos lingüísticos no português e como eles podem ser tratados pelos mecanismos de busca para que estes aumentem sua eficácia.
 - Buscas em linguagem natural: permitem aos usuários fazerem consultas em linguagem natural. Um exemplo é o mecanismo *AskJeeves*. Claro que este último não pode ser considerado de 3ª geração por apresentar apenas essa funcionalidade.
 - Buscas em bases de conhecimento: Muito semelhante às buscas em linguagem natural, este tipo de mecanismo cria automaticamente uma base estruturada de casos que permitem ao usuário obter respostas prévias a perguntas semelhantes. O mecanismo aprende a cada vez que uma sessão de interação tipo Pergunta/Resposta ocorre. Se o usuário não obtém a resposta apropriada, então ele é direcionado automaticamente pelo programa para falar com uma pessoa que tenta responder-lhe (este tipo de busca vem sendo muito usado em *Call Centers*);

- Busca por afinidade são similares aos filtros colaborativos. Entretanto, esse tipo de busca leva à idéia de associar pessoas e documentos às últimas conseqüências. Muitos níveis de similaridade baseados nas pessoas, autores, leitores e mantenedores dos documentos. Pode ser particularmente útil para grandes organizações para relacionar pessoas com interesses e competências similares;
- Buscas com mapeamento visual permitem aos usuários visualizar os resultados de suas buscas de forma gráfica. Documentos são ajuntados e relacionados formando clusters de informações relacionadas, de acordo com algoritmos de categorização que permitem navegação facilitada.

O processo de seleção do mecanismo de busca específico depende das necessidades de cada organização e da complexidade e tamanho das fontes de informação que devem ser pesquisadas. Isso pode fazer variar muito o custo do projeto. Soluções padronizadas de indexação e busca podem ser bastante baratas. Por outro lado, soluções completas e sofisticadas para grandes empresas podem significar milhares de reais em licenças de software, hardware, além dos custos de consultoria e desenvolvimento.

A integração de mecanismos de busca com a solução de EIP proporciona funcionalidades interessantes relacionadas à gestão de conhecimentos:

Notificação

- Usuários são notificados quando novos documentos são adicionados à base de conhecimentos criados por outra pessoa na organização;
- Usuários e/ou administradores podem programar rotinas de spiders e crawlers; o mecanismo de busca pode ser programado para vasculhar fontes de informação e gerar resultados para grupos específicos;
- Notificação baseada em eventos de exceção (vendas caíram 20%); o usuário pode definir onde ser notificado (e-mail, pager, etc);

Processo de busca e Interface

- O mecanismo permite ao usuário escolher em que fontes realizar pesquisas (diretórios, web, base de dados, etc);
- Salvar os resultados de busca e compartilha-los com outras pessoas na organização;
- Resultados podem ser personalizados de acordo, papel níveis de acesso dos usuários;
- O mecanismo de busca corrige erros ortográficos automaticamente;
- Mecanismo de busca considera sinônimos;
- Agrupa resultados de pesquisa;

- Permite busca multi-líguas.

Desempenho do mecanismo e *backend*

- Integração do mecanismo de busca com aplicações de business intelligence;
- Procura por tags XML;
- Mecanismo de categorização automática;

5 Desenvolvendo Categorizações e Taxonomias

Buscando simplificar a procura e a navegação, organizações precisam desenvolver estruturas de informação e categorias que sejam facilmente compreendidas pelos usuários da comunidade que utilizará o sistema. Independentemente do poder do mecanismo de busca empregado, na concepção de seu portal toda organização deve lidar com a questão da categorização da informação. A categorização adiciona informação fundamental para a classificação aos documentos, para que estes sejam organizados de acordo com uma taxonomia e facilmente encontráveis mais tarde.

Taxonomias são regras de alto nível para organizar e classificar. A disciplina não é nova e embora tenha sido evidenciada pelo desenvolvimento da Web, sempre foi o núcleo dos estudos de tratamento de informação em biblioteconomia. Atualmente novas ferramentas de software para categorização automática emergem e se aperfeiçoam rapidamente.

A formação de taxonomias se beneficia da marcação (*tagging*) dos documentos. A marcação é um passo importante, não só para garantir que mecanismos de busca encontrarão os documentos mais facilmente, e os organizarão em classes, como também para a disseminação dos documentos com base em regras de personalização. O conceito de metadados, ou dados que descrevem dados, base das linguagens de marcação como HTML e XML, é central para a evolução da Internet, tanto em sua parte pública quanto na corporativa. Os metadados provêm os contextos de significados necessários à informação e torna os documentos mais facilmente localizáveis por mecanismos de busca, auxiliando-os a resolver diversas ambigüidades das línguas naturais que trazem ruídos ao processo de busca.

Metadados também desempenham o importante papel de permitir que diferentes documentos sejam agrupados mais facilmente. A captura de metadados deveria ocorrer de preferência logo após a criação do conteúdo. Isso vem sendo facilitado graças ao uso de ferramentas semi-automáticas de apoio, que realizam uma combinação de processos de extração automáticos (tipicamente, autor, nome, data, etc) e manuais (palavras-chave, categorias, etc).

A criação dos diversos padrões de representação XML (*XLink*, *XPath*, *XSL*, *XQuery*, etc) vem possibilitando a criação de hiper-espacos de informação muito mais sofisticados. Espacos que permitem o tratamento, por software, de documentos marcados com significados mais livres, com semânticas mais específicas, e, ao mesmo tempo, universais

(Berners-Lee, 2003). Essa idéia de marcação semântica com o uso de identificadores universais (*URI*) está na base da concepção do que vem sendo denominado “Web Semântica” e é atribuída a Tim Bernes-Lee, considerado o inventor da Web.

Taxonomias e categorizações assumem geralmente formas hierárquicas. Entretanto, novas formas “mais visuais” como grafos ou árvores hiperbólicas se tornam comuns. O leitor poderá consultar *websites* tais como: Antartic.ca, TheBrain ou InXight.

Arquitetos de informação, com o auxílio dos geradores ou extratores semi-automáticos de conteúdos, são geralmente responsáveis pela criação das taxonomias a serem utilizadas na organização. Ferramentas de classificação automática simplificam e agilizam o processo.

O verdadeiro teste para qualquer taxonomia é o grau de eficiência que esta proporciona ao grupo de usuários para quem ela foi desenhada: os usuários estão sendo capazes de encontrar informação relevantes de forma eficiente, ou seja, sem desperdício de tempo?

Alguns dos desafios encontrados na criação e manutenção de uma taxonomia advêm dos seguintes fatores:

- É preciso estar ciente de que a ambigüidade é parte do processo e estará sempre presente, podendo no máximo ser minimizada;
- O certo ou errado universal não existem: assim pessoas diferentes irão desenvolver taxonomias diferentes. Conseqüentemente, ajuda muito criar múltiplos caminhos para se encontrar a mesma informação;
- Categorizações automatizadas baseadas na freqüência de palavras podem não gerar bons resultados. Preferencialmente, a construção de uma taxonomia deveria iniciar-se com algumas entradas desenvolvidas conjuntamente pelos geradores de conteúdo e os arquitetos de informação. Como vimos, algumas novas ferramentas de software permitem hoje recategorizações semi-automatizadas baseadas em *feedbacks* dos usuários;
- A marcação é uma tarefa intensiva e requer cooperação entre os criadores de conteúdo. Aqui também novas ferramentas de extração automática ajudam a minimizar as entradas manuais, automatizando a geração de marcas padronizadas que descrevem os documentos, tais com autores, filiação, data de criação, palavras-chave, etc;
- Quem ou qual grupo irá manter a integridade da taxonomia à medida que ela evolui no tempo?

Embora, na maioria dos casos, integradores de sistemas externos ou firmas de desenvolvimento web sejam responsáveis pelo desenvolvimento de uma taxonomia inicial, claramente a manutenção de um esforço perene relativo à arquitetura informacional exigirá novas funções além do uso de ferramentas sofisticadas para automatizar grande parte do esforço de classificação.

6 Considerações finais

Os portais corporativos já causam e continuarão a causar mudanças cada vez mais fundamentais nos processos gerenciais e no modo como as empresas funcionam. A implementação de portais corporativos deve ser considerada para servir, principalmente, às necessidades de funcionários no que tange ao compartilhamento de informações e conhecimentos. As redes corporativas e as comunidades de práticas que extrapolam as paredes físicas das organizações também são beneficiadas.

Dessa forma, as plataformas de portais corporativos mais avançadas vão além do acesso à informação estruturada e não (ou semi) estruturada. Elas provêm soluções que se adequam às necessidades diárias de informação e de colaboração de cada funcionário e/ou parceiros de negócio: personalizam o acesso à informação, automatizam e melhoram os ciclos de decisão complexos de trabalhadores do conhecimento e podem incentivar níveis mais profundos de colaboração entre empregados.

Com a vantagem de prover um único ponto de contato para todas as fontes de informação, o portal corporativo assume o papel, sem precedente, de integrador universal dentro das organizações. Por esse e muitos outros motivos comentados no decorrer deste texto, os portais corporativos constituem novos instrumentos de gestão de informação e conhecimento nas organizações.

7 Referências

- Bawden, D. "Information Overload", in *Library & Information Briefings*. Issue 92. Jan 2001. Disponível em: <<http://litc.sbu.ac.uk/publications/libs/libs92.pdf>>. Acesso em: 5 jun. 2003.
- Rada, R. "Hypertext: from text to expertext", McGraw Hill, 1991.
- Shenk, D. "Data smog: surviving the information glut", HarperEdge, New York, 1997.
- Bird, J. "The Reuters guide to good information strategy", Reuters, London, 1997.
- Stewart, Thomas A. *Capital intelectual: a nova vantagem competitiva das empresas*. 2a ed. Rio de Janeiro: Campus, 1998.
- Sproull, Lee, and Sara Keisler. *Connections: New Ways of Working in the Networked Organization*. Cambridge, Mass.: MIT Press, 1993.
- Nevis, E.C., DiBella, A.J., and Gould, J.M., 1995. *Understanding Organizations as Leaning Systems*, *Sloan Management Review*, 36, Winter, 73-85.
- Nonaka, I. and Takeuchi, H. (1997). *Criação de conhecimento na empresa: como as empresas japonesas geram a dinâmica da inovação*. Rio de Janeiro: Campus.
- Sveiby, Karl E. *A nova riqueza das organizações - gerenciando e avaliando patrimônios de conhecimento*. Rio de Janeiro: Campus, 1998.

Davenport, Th. and Prusak, L. Conhecimento empresarial - como as organizações gerenciam o seu capital intelectual. Rio de Janeiro: Campus, 1998.

Broder, Andrei. "A taxonomy of web search", ACM SIGIR Forum, v36, n2, ACM Press, 2002.
Disponível em: <<http://doi.acm.org/10.1145/792550.792552>>. Acesso em: 10 jun. 2003.

Small, R. and Ferreira, S. "Multimedia vs. Print Information Resources: Information Location and Use, Motivation, and Learning Patterns for Children and Adults". In Proceedings of the 16th National Convention of the Association for Educational Com. and Tech. Washington. 1994.

Bräscher, M. "A Ambiqüidade na Recuperação da Informação". DataGramZero - Revista de Ciência da Informação - v.3 n.1, Fev 2002. Disponível em:
<http://www.dgz.org.br/fev02/Art_05.htm>. Acesso em: 5 jun. 2003.

Berners-Lee, Tim. "WebArchitecture from 50,000 feet". Disponível em:
<<http://www.w3.org/DesignIssues/Architecture.html>>. Acesso em: 5 jun. 2003.