

**UNIVERSIDADE FEDERAL DE SANTA CATARINA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

Dorzeli Salete Trzeciak

**BASE DE CONHECIMENTO EM PORTAIS: PROPOSTA DE  
ORGANIZAÇÃO DE CONTEÚDO NAS ÁREAS DE INOVAÇÃO E  
TECNOLOGIA DA INFORMAÇÃO APLICADAS À GESTÃO DE  
NEGÓCIOS**

Dissertação de Mestrado

Florianópolis

2002

DORZELI SALETE TRZECIAK

**BASE DE CONHECIMENTO EM PORTAIS: PROPOSTA DE ORGANIZAÇÃO DE  
CONTEÚDO NAS ÁREAS DE INOVAÇÃO E TECNOLOGIA DA INFORMAÇÃO  
APLICADAS À GESTÃO DE NEGÓCIOS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina, como requisito parcial para obtenção do título de Mestre em Engenharia de Produção.

Orientadora: Aline França de Abreu, Ph.D.

Co-orientador: Pedro Felipe de Abreu, Ph.D.

Florianópolis

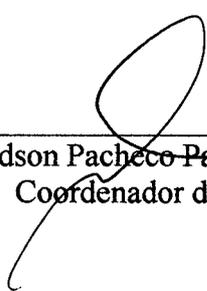
2002

DORZELI SALETE TRZECIAK

**BASE DE CONHECIMENTO EM PORTAIS: PROPOSTA DE ORGANIZAÇÃO DE  
CONTEÚDO NAS ÁREAS DE INOVAÇÃO E TECNOLOGIA DA INFORMAÇÃO  
APLICADAS À GESTÃO DE NEGÓCIOS**

Esta dissertação foi julgada e aprovada para a obtenção do título de  
**Mestre em Engenharia de Produção** no Programa de Pós-Graduação em Engenharia de  
Produção da Universidade Federal de Santa Catarina

Florianópolis, 13 de novembro de 2002.



---

Edson Pacheco Paladini, Dr.  
Coordenador do Curso

**BANCA EXAMINADORA**



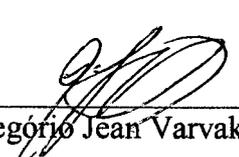
---

Aline França de Abreu, Ph.D.  
Orientadora



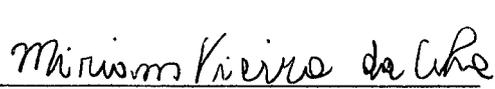
---

Pedro Felipe de Abreu, Ph.D.  
Co-orientador



---

Gregório Jean Varvakis Rados, Ph.D.



---

Miriam Vieira da Cunha, Dra.

Dedico esta dissertação à minha família,  
em especial à minha mãe, pelo exemplo  
de humildade, perseverança e honestidade.

## AGRADECIMENTOS

À Professora Aline, pela acolhida num momento difícil da minha vida, pela oportunidade de aprendizado e crescimento, pela confiança e dedicação, enfim por ter acreditado em mim e por ter sido muito mais que minha orientadora;

Ao Professor Pedro pelas constantes provocações, pelo apoio e confiança depositados;

Aos Professores Gregório e Miriam, por terem aceitado o convite a participar da banca examinadora, pelos comentários e sugestões;

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de mestrado concedida, que foi de primordial importância para a realização desta dissertação;

Ao IGTI, em especial ao Manoel, pela valorosa e constante contribuição;

Ao Alexandre, por não resistir a uma “*donzela em apuros*”;

À Carmen, Juliana e ao José Carlos que nunca mediram esforços para me ajudar;

À Celina, pelo *design* da base;

Ao Leonardo Lima, o Leozinho, que pacientemente me mostrou e ensinou o lado bom da tecnologia;

Às amigas Janice e Cíntia, sempre presentes nas horas boas e também nas horas ruins;

Ao Enio e sua família, que sempre me apoiaram nas minhas conquistas;

A Deus, pelas belezas que posso ver, pelas palavras que posso ouvir e por tudo que posso sentir. Por todas as vezes que caí e pelas vezes que levantei. Por tudo que acertei e pelos momentos que falhei. Pela liberdade de pensar, de desejar, de querer e de optar. De não ser nada, mas de poder ter me tornado quem realmente sou;

E a todos àqueles que, direta ou indiretamente, contribuíram para a realização desta dissertação, minha profunda gratidão e carinho.

“O conhecimento é de duas espécies. Podemos conhecer nós mesmos um assunto ou saber onde podemos encontrar informações a respeito.”

Samuel Johnson

TRZECIAK, Dorzeli S. **Base de conhecimento em portais: proposta de organização de conteúdo nas áreas de inovação e tecnologia da informação aplicadas à gestão de negócios.** 2002. 102 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

## RESUMO

O objetivo da presente dissertação é estruturar uma base de conhecimento nas áreas de inovação e tecnologia da informação, com ênfase na organização de conteúdo, visando armazenamento adequado, fácil recuperação e sua disponibilização na Internet. A revisão de literatura possibilitou levantar subsídios para a elaboração do modelo de uma base de conhecimento, concebida como um repositório de conhecimento. O modelo, proposto ao Portal IGTI, compreende as seções base bibliográfica, memória organizacional, *handbooks*, estado da arte, mapas de competências e páginas temáticas com análise de tendências, que disponibilizarão conteúdo (em forma de referência ou na íntegra) das áreas de abrangência do portal. Espera-se, como resultado, que o modelo proposto possibilite sistematizar o conhecimento nas áreas temáticas do Portal IGTI, dando ao usuário uma noção de quais informações estão disponíveis, como estão organizadas e como podem ser recuperadas, oferecendo então, informações relevantes, de forma eficiente e em tempo hábil, e que tenha uma repercussão positiva nas áreas acadêmica e de negócios, promovendo a geração e o compartilhamento de conhecimento de ponta e disponibilizando-o à sociedade via Internet.

**Palavras-chave:** Internet, Recuperação da informação, Base de conhecimento

TRZECIAK, Dorzeli S. **Base de conhecimento em portais: proposta de organização de conteúdo nas áreas de inovação e tecnologia da informação aplicadas à gestão de negócios.** 2002. 102 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

## **ABSTRACT**

The objective of the present dissertation is to structure a knowledge base in the innovation and information technology areas, with emphasis in content organization, aiming at suitable storage, easy retrieval and its availability on the Internet. The literature revision made it possible to raise means for the elaboration of a knowledge base model, conceived as a knowledge repository. The model, proposed to IGTI Portal, comprehends the sections: bibliographical base, organizational memory, handbooks, state of the art, competence maps and thematic pages with trends analysis, that will make available the content (in reference form or full text) of the interest areas of the portal. One expects, as result, that the considered model makes possible to systemize the knowledge in the thematic areas of the IGTI Portal, giving to the user a notion of which information are available, how it is organized and how it can be recovered, offering then, relevant information, in efficient way and in short time, and that it has a positive repercussion in the academic and business areas, promoting the creation and the sharing of advanced knowledge and making it available to the society via Internet.

**Keywords:** Internet, Information retrieval, Knowledge base

## LISTA DE QUADROS

Quadro 1 – Requisitos funcionais de um portal corporativo.....	p. 34
Quando 2 – Vantagens e desvantagens da linguagem natural.....	p. 50
Quando 3 – Vantagens e desvantagens da linguagem controlada.....	p. 51
Quadro 4 – Operadores booleanos.....	p. 61
Quadro 5 – Sinais utilizados para realizar buscas.....	p. 61

## LISTA DE FIGURAS

Figura 1 – Portal IGTI.....	p. 77
Figura 2 – Modelo proposto de base de conhecimento.....	p. 80
Figura 3 – <i>Interface</i> de descrição e acesso ao conteúdo das seções.....	p. 84
Figura 4 – <i>Interface</i> da seção estado da arte.....	p. 85
Figura 5 – Busca avançada na base de conhecimento.....	p. 86
Figura 6 – Tipos de busca.....	p. 87
Figura 7 – Resultado da busca na base de conhecimento.....	p. 88
Figura 8 – Detalhamento do registro.....	p. 89

## SUMÁRIO

<b>CAPÍTULO 1 - INTRODUÇÃO.....</b>	<b>p. 12</b>
1.1 TEMA E PROBLEMA DA PESQUISA.....	p. 12
1.2 OBJETIVOS .....	p. 14
1.2.1 Objetivo geral.....	p. 14
1.2.2 Objetivos específicos.....	p. 14
1.3 JUSTIFICATIVA.....	p. 15
1.4 PROCEDIMENTOS METODOLÓGICOS.....	p. 16
1.5 DELIMITAÇÕES DA DISSERTAÇÃO.....	p. 18
1.6 ESTRUTURA DA DISSERTAÇÃO.....	p. 19
<b>CAPÍTULO 2 - REVISÃO DE LITERATURA.....</b>	<b>p. 20</b>
2.1 GESTÃO DO CONHECIMENTO.....	p. 20
2.1.1 Gestão do conhecimento e a gestão de negócios.....	p. 22
2.1.2 O papel da tecnologia da informação na gestão do conhecimento.....	p. 25
2.2 PORTAIS.....	p. 28
2.2.1 Tipos de portais.....	p. 30
2.2.2 Componentes funcionais de um portal corporativo.....	p. 33
2.3 BASES DE DADOS.....	p. 36
2.3.1 Tipos de bases de dados.....	p. 39
2.3.2 Formatos de registro.....	p. 42
2.3.3 Indexação.....	p. 45
2.3.4 Linguagem natural e linguagem controlada.....	p. 49
2.3.5 Avaliação de bases de dados.....	p. 52
2.3.6 Bases de conhecimento.....	p. 54
2.4 RECUPERAÇÃO DE INFORMAÇÕES.....	p. 55

2.4.1 Mecanismos de busca.....	p. 57
2.4.2 Estratégias de busca.....	p. 63
2.4.3 <i>Interfaces</i> .....	p. 65
2.4.4 Dificuldades do usuário na recuperação de informações.....	p. 71
<b>CAPÍTULO 3 - MODELO PROPOSTO DE BASE DE CONHECIMENTO..</b>	<b>p. 75</b>
3.1 CONTEXTUALIZAÇÃO.....	p. 75
3.1.1 Núcleo de Estudos em Inovação, Gestão e Tecnologia da Informação – IGTI.....	p. 75
3.1.2 Projeto do Portal IGTI.....	p. 76
3.2 PREMISSAS BÁSICAS.....	p. 79
3.3 MODELO PROPOSTO DE BASE DE CONHECIMENTO.....	p. 80
3.3.1 Recuperação de informações na base de conhecimento.....	p. 85
3.3.1.1 <i>Tipos de busca</i> .....	p. 86
3.3.1.2 <i>Resultado da busca na base de conhecimento</i> .....	p. 88
3.4 RESULTADOS ESPERADOS.....	p. 89
<b>CAPÍTULO 4 - CONSIDERAÇÕES FINAIS.....</b>	<b>p. 91</b>
4.1 CONCLUSÕES.....	p. 91
4.2 CONTRIBUIÇÕES.....	p. 94
4.3 LIMITAÇÕES.....	p. 95
4.4 RECOMENDAÇÕES.....	p. 95
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>p. 96</b>

## **CAPÍTULO 1 - INTRODUÇÃO**

### **1.1 TEMA E PROBLEMA DA PESQUISA**

A informação pode ser considerada como o insumo básico que contribui para o desenvolvimento das nações, possibilitando o intercâmbio, o compartilhamento e a disseminação de conhecimento, oferecendo suporte aos tomadores de decisão.

E apenas mediante um processo de sistematização que administre a informação pertinente no momento oportuno, isto é, na tomada de decisão, é possível antecipar tanto as ameaças como as oportunidades de mudanças produzidas pela velocidade das transformações tecnológicas e pelo processo de globalização (TOMAÉL et al., 2001).

A Internet, convergência das tecnologias da computação e da comunicação, representa uma verdadeira revolução nos métodos de geração, armazenagem, processamento e disseminação da informação. Ela vem aumentando as alternativas de comunicação, modificando o modo como as pessoas encaram o mundo e criando novas formas de disponibilizar e recuperar informações (ARAÚJO, 1999/2000; TOMAÉL et al., 2001).

Mas, ao disponibilizar dados e informações a qualquer momento e de inúmeras fontes, a Internet gera uma sobrecarga de informações, de forma desorganizada e de difícil recuperação (ATAÍDE, 1997).

Essa sobrecarga de informações tem deixado muitos usuários frustrados. De um lado há muita informação disponível, e de outro, a falta de organização, que impede a recuperação de informações relevantes. Além disso, diante do caos informacional, os usuários acabam perdendo muito tempo para fazer as buscas e encontrar as informações que realmente precisam.

Desse modo, a organização, controle e recuperação de informações úteis na Internet são alguns dos maiores desafios enfrentados atualmente.

Bertholino (1999) e Cendón (2000) colocam que, diante da necessidade de facilitar a busca e o acesso às informações, as bases de dados e os mecanismos de busca vêm proliferando cada vez mais através da interação das redes de telecomunicações, constituindo-se como ferramentas fundamentais para recuperar, organizar, estruturar e disponibilizar informações, atendendo às necessidades da comunidade usuária.

Neste contexto, de localizar e facilitar o acesso a informações úteis e relevantes, surge a proposta do Núcleo de Estudos em Inovação, Gestão e Tecnologia da Informação (IGTI), de desenvolver um Portal de Conhecimento (Portal IGTI) para as áreas de inovação e de tecnologia da informação aplicadas à gestão competitiva de negócios, contemplando os ambientes conhecimento, inteligência competitiva, redes de relacionamento e apoio à pesquisa.

A presente dissertação está relacionada ao ambiente de conhecimento e busca sistematizar o conhecimento nas áreas temáticas do Portal IGTI, em uma base de conhecimento. Seu conteúdo, em forma de referência ou na íntegra, será disponibilizado nas seções base bibliográfica, memória organizacional, *handbooks*, estado da arte, mapas de competências e páginas temáticas com análise de tendências.

A motivação para o desenvolvimento do Portal IGTI prende-se ao fato do IGTI, desde a sua formação (1997), sentir necessidade de ferramentas eficientes que agilizassem e facilitassem o processo de busca de informações na Internet. Aliado a isso, havia a necessidade de disponibilizar material para os alunos (integrantes do IGTI) tanto presenciais quanto a distância. Somando-se a isto a criação do ensino à distância pelo Programa de Pós-Graduação em Engenharia de Produção.

Muitos trabalhos já foram e ainda estão sendo desenvolvidos nesta área, pelos integrantes do IGTI. Dentre eles pode-se destacar a tese de Helena Pereira da Silva, intitulada “Inteligência competitiva na Internet: proposta de um modelo”, onde é apresentado o desenvolvimento de um processo de inteligência competitiva utilizando a Internet como ferramenta. Este trabalho está servindo de referência para o desenvolvimento da dissertação relacionada ao ambiente inteligência competitiva.

O interesse da autora em desenvolver a presente dissertação, se deve ao fato de ser integrante do IGTI e de participar ativamente de grupos de estudo e das atividades desenvolvidas no Núcleo, mais especificamente no projeto do Portal IGTI.

Portanto, a partir dos fatores apresentados anteriormente e da proposta do IGTI, de desenvolver um Portal de Conhecimento em áreas temáticas e tendo como meta principal a recuperação, a disseminação e a disponibilização de informações relevantes de forma eficiente e em tempo hábil, surge a seguinte questão de pesquisa: como organizar o conteúdo de uma base de conhecimento?

## **1.2 OBJETIVOS**

### **1.2.1 Objetivo geral**

O objetivo geral da presente dissertação é:

Estruturar uma base de conhecimento nas áreas de inovação e tecnologia da informação, com ênfase na organização de conteúdo, visando seu armazenamento, sua recuperação e sua disponibilização na Internet.

### **1.2.2 Objetivos específicos**

Os objetivos específicos da presente dissertação são:

- identificar, na literatura, o potencial da tecnologia da informação e ambientes digitais na proposição de estruturas e serviços para armazenamento, recuperação e disponibilização de informações;
- identificar meios de armazenamento de conteúdo em bases de conhecimento;
- identificar meios de recuperação de conteúdo em bases de conhecimento;
- identificar meios de disponibilização de conteúdo em bases de conhecimento;

- elaborar modelo de base de conhecimento para organização de conteúdo;
- aplicar proposta ao Portal IGTI.

### 1.3 JUSTIFICATIVA

O crescente número de usuários, somado a ampla interconectividade, robustez, interatividade e facilidade com que os recursos informacionais podem ser criados e acessados, tornam a Internet um meio atraente de disponibilização de uma variedade enorme de informações (CENDÓN, 2000).

O acesso a estas informações se dá por diferentes pessoas, de diferentes idades, culturas, territórios e necessidades, sejam cientistas, estudantes, professores ou profissionais liberais, que possuem necessidades de informação diversificada, isto é, cultural, profissional, acadêmica, lazer ou entretenimento (BUENO; VIDOTTI, 2000).

Porém, diante do grande volume de informações disponíveis na Internet, é necessário um processo de organização destas informações, que contemple desde sua criação até seu uso final, para que os usuários possam encontrar o que realmente precisam, para satisfazer as suas necessidades, no menor espaço de tempo possível.

As bases de dados e os mecanismos de busca podem ser consideradas ferramentas úteis para recuperar, armazenar e disponibilizar informações relevantes e de forma eficiente aos usuários. E para que uma base de dados seja eficiente, é necessário considerar a qualidade das estruturas de armazenamento e recuperação e a qualidade do conteúdo da mesma. Sob o enfoque da qualidade das estruturas de armazenamento e recuperação, a participação da comunidade de usuários é fundamental. No tocante ao conteúdo, os mecanismos de coleta e registro são dimensões de suma importância, pois a funcionalidade destas *interfaces* será fundamental no processo de carga da base de dados (PEREIRA et al., 1999).

É importante destacar que informações úteis poderão ser utilizadas de diferentes formas, em diferentes momentos, para diferentes objetivos e por diferentes pessoas. Portanto, elas somente terão importância quando forem ao encontro dos interesses dos usuários (BORGES; CARVALHO, 1998).

No meio acadêmico a informação constitui o principal insumo para a geração de conhecimento e no meio empresarial é crescente o valor da informação para os negócios. Informações tanto sobre o ambiente interno quanto o ambiente externo, que possibilitam identificar oportunidades e ameaças, fundamentais para a tomada de decisões (ATAÍDE, 1997).

Portanto, uma base de dados bem estruturada pode agilizar o acesso a informações atualizadas, precisas, confiáveis, no menor espaço de tempo possível, oferecendo subsídios para a gestão de negócios e contribuindo para a geração de conhecimento novo.

Tão importante quanto o valor da informação e do conhecimento, é o seu gerenciamento. Essa percepção, ligada aos grandes avanços da tecnologia da informação, levou ao surgimento da gestão do conhecimento, como uma disciplina ou uma preocupação explícita que passa a fazer parte das estratégias organizacionais, preocupadas com o gerenciamento e a disponibilização de conhecimento, visando obter maior vantagem competitiva e conquistar novos mercados (TERRA; GORDON, 2002).

Neste contexto, para esta dissertação torna-se relevante consolidar, através de revisão de literatura, o conhecimento sobre os temas envolvidos, identificar requisitos para armazenamento, recuperação e disponibilização de conteúdo na Internet e aplicar os conceitos identificados na estruturação da base de conhecimento do Portal IGTI.

Desta forma, a presente dissertação terá repercussão positiva nas áreas acadêmica e de negócios, permitindo o aumento do escopo de acesso à informação na temática do Portal IGTI, a agilização do acesso a informações úteis, confiáveis e relevantes, a disponibilização de grande quantidade de conteúdo digital, em forma de referências ou texto integral, bem como a geração de conhecimento de ponta numa ação colaborativa e de compartilhamento, disponível à sociedade via Internet.

#### **1.4 PROCEDIMENTOS METODOLÓGICOS**

Cervo e Bervian (1996, p. 44) definem pesquisa como “uma atividade voltada para a solução de problemas através do emprego de processos científicos.” A pesquisa parte de “uma dúvida

ou problema e, com o uso do método científico, busca uma resposta ou solução.” Assim, uma pesquisa, além de buscar a atualização do conhecimento para uma nova tomada de posição, permite transformar em ação concreta os resultados de seu trabalho.

Os autores argumentam ainda, que qualquer que seja o tipo de pesquisa, exige-se uma pesquisa bibliográfica prévia, quer para o levantamento da situação em questão, quer para a fundamentação teórica ou ainda para justificar os limites e contribuições da própria pesquisa.

Esta dissertação enquadra-se como pesquisa aplicada, tendo como característica principal contribuir para fins práticos, ou seja, aplicar ou utilizar mais ou menos imediatamente os resultados na solução de problemas que ocorrem na realidade (CERVO; BERVIAN, 1996; MARCONI; LAKATOS, 1990).

A seguir, são apresentados os procedimentos adotados para a elaboração da presente dissertação.

No planejamento do trabalho houve a elaboração do projeto de pesquisa, onde foram definidos o tema, o problema, a questão de pesquisa, os objetivos, a justificativa e as delimitações da dissertação, bem como os principais conceitos e a estratégia para a fundamentação teórica da dissertação.

Na revisão de literatura, foram pesquisados os conceitos referentes a gestão de conhecimento, portais, bases de dados e recuperação de informações. Estes conceitos envolvem as áreas de Informática, Ciência da Informação e Administração. Para a realização da revisão, foram utilizados livros, dissertações, periódicos, anais e documentos disponíveis na Internet, procurando abranger os últimos dez anos.

Sendo esta dissertação parte do projeto do Portal IGTI, inicialmente foram realizadas reuniões periódicas para a definição das etapas do projeto. Estas reuniões eram compostas por alunos de mestrado, doutorado e graduação, tendo como coordenadores os professores Aline França de Abreu e Pedro Felipe de Abreu.

O Portal IGTI será composto por quatro ambientes (conhecimento, inteligência competitiva, redes de relacionamento e apoio à pesquisa), e esta dissertação está relacionada ao ambiente

conhecimento. Os ambientes conhecimento e inteligência competitiva gerarão como produto, duas dissertações e os ambientes redes de relacionamento e apoio à pesquisa, duas teses.

A estruturação do modelo da base de conhecimento, proposto ao Portal IGTI, se deu com base nos subsídios levantados na revisão de literatura, procurando atender a problemática apresentada no início da dissertação e os objetivos do próprio Portal.

É importante deixar claro, que o modelo que está sendo proposto não será implementado no momento. A sua implementação acontecerá com a conclusão do projeto do Portal IGTI. Assim, com a finalização desta dissertação, tem-se como resultado a parte teórica do ambiente conhecimento (desenho da base e suas funcionalidade). Os trabalhos referentes aos demais ambientes se encontram em andamento, sendo que as duas teses estão em fase de qualificação.

Com a implementação da base, é provável que sejam necessários alguns ajustes, visando o desempenho adequado da mesma. Contudo, as funcionalidades apresentadas no modelo serão mantidas.

A questão referente a tecnologia a ser utilizada para implementar a base de conhecimento, ainda está em discussão, já que envolve também os outros ambientes.

Atualmente, estão sendo realizadas reuniões para acompanhar o andamento dos trabalhos, visando a conclusão do projeto do Portal. Há também, um grupo de estudos, formado por alunos de graduação e pós-graduação, que vem trabalhando na modelagem de dados da base bibliográfica.

Após a apresentação do modelo, são expostas as considerações finais.

## **1.5 DELIMITAÇÕES DA DISSERTAÇÃO**

O escopo da presente dissertação compreende o desenho da base de conhecimento e suas funcionalidades. Não fazem parte do escopo da dissertação os seguintes aspectos:

- programação da base de conhecimento;

- disponibilização da base de conhecimento para usuários; e
- avaliação de satisfação do usuário.

## **1.6 ESTRUTURA DA DISSERTAÇÃO**

Esta dissertação está estruturada em quatro capítulos, descritos abaixo.

O primeiro capítulo é constituído de uma introdução, onde constam o tema e o problema, os objetivos (geral e específicos) e a justificativa, além dos procedimentos metodológicos, as delimitações e a estrutura da dissertação.

O segundo capítulo é composto pela revisão de literatura, onde são desenvolvidos os principais conceitos que permitiram o embasamento teórico da dissertação.

No terceiro capítulo são apresentados o local para o qual o modelo foi proposto, as premissas básicas que serviram para desenvolver o modelo, a estruturação, implementação, dinâmica de funcionamento e manutenção do modelo proposto de base de conhecimento e os resultados esperados.

As conclusões, as contribuições e as sugestões para futuros trabalhos são expostas no quarto capítulo.

Em seguida são disponibilizadas as referências bibliográficas.

## CAPÍTULO 2 - REVISÃO DE LITERATURA

A revisão de literatura prende-se aos conceitos necessários ao objetivo da dissertação. Assim, os conceitos aqui tratados formam a base de sustentação para o desenvolvimento do modelo proposto. São eles: gestão de conhecimento, portais, base de dados e recuperação de informações. A revisão de literatura sobre portais pode ser justificada pelo fato da base de conhecimento estar inserida no Portal IGTI. Quanto a base de dados e recuperação de informações, justifica-se por serem consideradas ferramentas que possibilitam a recuperação, a disseminação e a disponibilização de informações. Em relação a gestão de conhecimento, justifica-se pelo fato que portais e bases de dados podem ser considerados ferramentas que auxiliam o processo de gestão do conhecimento. Assim, além de organizarem e disponibilizarem o conteúdo, estarão contribuindo para a gestão do conhecimento do IGTI.

### 2.1 GESTÃO DO CONHECIMENTO

- O desenvolvimento da sociedade vem provocando mudanças nos contextos sociais e organizacionais, marcando a passagem deste milênio por uma nova revolução - a revolução digital, também denominada de revolução da informação ou revolução do conhecimento. Devido aos avanços da tecnologia, a informação, principal responsável pelo diferencial competitivo das organizações, pode ser acessada de uma maneira mais fácil e rápida, permitindo a criação, a utilização e a disponibilização de novos conhecimentos numa velocidade cada vez maior. Mas, em meio a tantas afirmações, é importante saber distinguir o que são dados, informações e conhecimento (ROMANI; BORSZCZ 2001).
- Para Davenport (1998, p. 18), dados são “simples observações sobre o estado do mundo, são facilmente estruturados, facilmente obtidos por máquinas, freqüentemente quantificados e de fácil transferência.” Informações são “dados dotados de relevância e propósito, exigem unidade de análise, consenso em relação ao significado e necessariamente a medição humana.” Por sua vez, conhecimento é “a informação da mente humana, inclui reflexão,

síntese e contexto, é de difícil estruturação, de difícil captura em máquinas e de difícil transferência.”

- Ao processo de localizar, organizar, transferir e usar informações dentro de uma organização, com a finalidade de transformá-las em conhecimento que contribuirá para que a organização obtenha vantagem competitiva, Davenport (1998) chama de gestão do conhecimento.
- Terra e Gordon (2002, p. 56) relatam que o conceito de gestão do conhecimento não é novo. Entretanto, em função das novas tecnologias, novas mídias, novos dispositivos e novas técnicas vem sendo esquematizado e disponibilizado de uma forma diferente. Para os autores, a gestão do conhecimento é “uma disciplina emergente que está se constituindo a partir da contribuição de vários campos teóricos e levará tempo para que evolua completamente e que suas oportunidades e seus efeitos sejam totalmente compreendidos.”

Segundo Teixeira Filho (2000), a gestão do conhecimento é

uma certa forma de olhar para a organização, em busca de pontos dos processos de negócio em que o conhecimento possa ser usado como vantagem competitiva. Conhecimento útil, oriundo da experiência, da análise, da pesquisa, do estudo, da inovação, da criatividade, além de conhecimento sobre mercado, concorrência, clientes, processos do negócio, tecnologia e tudo mais que possa trazer vantagem competitiva para a organização.

- De acordo com Davenport e Prusak (1998), a gestão do conhecimento é o conjunto de atividades relacionadas com a geração, codificação e transferência do conhecimento. Estes autores ainda colocam que a gestão do conhecimento procura melhorar os recursos existentes na organização, de forma orientada para o conhecimento. Este conhecimento é transferido nas organizações quer se gerencie ou não esse processo, pois transferências cotidianas fazem parte da vida organizacional, entretanto, localizadas e fragmentadas. Embora a transferência espontânea e não estruturada do conhecimento seja vital para o sucesso de uma organização, o termo gestão do conhecimento implica na transferência formalizada, sendo necessário, portanto, desenvolver estratégias específicas para incentivar essas trocas espontâneas.
- O’Dell e Grayson Júnior (2000, p. 23) conceituam gestão do conhecimento como “uma estratégia consciente de obter o conhecimento certo das pessoas certas no momento certo e de ajudar as pessoas a transformarem a informação em ação de modo que se empenhem em

melhorar o desempenho da empresa.” Os autores ainda afirmam que a gestão do conhecimento é a mentalidade gerencial, incluindo o uso de experiências passadas (bibliotecas, banco de dados, pessoas) como base e a criação de novos meios que permitam a troca de conhecimento (*sites* de Intranets, comunidades de práticas, redes).

Para estes autores, muitas organizações concentram esforços para criar, identificar, coletar e organizar as melhores práticas e conhecimentos internos para entender o que elas sabem e onde se encontra esse conhecimento. Entretanto, o fato de saber que o conhecimento e as melhores práticas existem não garante o seu uso e/ou a sua transmissão. Este processo deve concentrar-se em compreender e entender as práticas existentes, proporcionar meios para adaptação e aplicação das melhores práticas a situações novas, criar um novo conhecimento e colocá-lo em ação.

- ⇒ Para Thives Júnior e Angeloni (2000), a gestão do conhecimento é um campo multidisciplinar, que envolve a gestão da informação, a tecnologia da informação, a comunicação interpessoal, o aprendizado organizacional, a motivação, o treinamento e a análise de processos. Trata-se de um enfoque integrado para identificar, capturar, gerenciar e compartilhar o ativo informacional das organizações, incluindo documentos, bases de dados e outros repositórios, além da competência individual dos funcionários.
- ⇒ Com base nos autores citados, conclui-se que a gestão do conhecimento pode ser entendida como um processo que visa a criação, a identificação, a coleta, a utilização e o compartilhamento do conhecimento, tanto interno quanto externo, visando atingir plenamente os objetivos da organização.

### **2.1.1 Gestão do conhecimento e a gestão de negócios**

Para Barroso e Gomes (1999), no mercado que cada vez está mais competitivo, o sucesso dos negócios depende basicamente da qualidade do conhecimento que as organizações aplicam nos seus processos de tomada de decisão.

Neste contexto, utilizar o conhecimento existente na organização, visando a criação de vantagens competitivas torna-se um desafio decisivo.

Duarte e Szostak (1999) compartilham deste pensamento, ao relatarem que cada vez mais as organizações estão reconhecendo o conhecimento como seu principal ativo. Em sua maior parte, esse conhecimento é formado pelas informações técnicas necessárias para a realização das tarefas, pelas práticas de negócios das organizações, pelas legislações e normas aplicáveis às atividades e pelas experiências dos funcionários no exercício das suas atividades. Para estes autores, o maior problema reside no fato que a maior parte desse conhecimento está na mente das pessoas (conhecimento tácito) e não documentado de forma estruturada (conhecimento explícito).

→ Em relação a estes dois tipos de conhecimento, explícito e tácito, Nonaka e Takeuchi (1997) fazem a seguinte distinção: o explícito é o conhecimento formal contido nos manuais e nas normas de praxe da organização, podendo ser facilmente estruturado e verbalizado, e o tácito, é o conhecimento inerente às pessoas, ou seja, as habilidades que elas possuem. O conhecimento explícito é a parte do conhecimento que pode ser transportada, armazenada e compartilhada e o conhecimento tácito é a parcela não estruturada do conhecimento que não pode ser registrada e/ou facilmente transmitida. Os autores acreditam que o conhecimento acumulado externamente, se amplamente compartilhado dentro da organização, forma uma base de conhecimento organizacional que abastece a inovação e permite melhorias contínuas.

→ Ainda de acordo com Nonaka e Takeuchi (1997), a interação entre estes dois conhecimentos resulta em quatro modos de conversão de conhecimento: a socialização (processo de compartilhamento de experiências entre as pessoas de um grupo, se desenvolvendo freqüentemente através da observação, imitação e prática); a externalização (processo de organização do conhecimento tácito em conhecimento explícito através de metáforas, analogias, conceitos, hipóteses ou modelos); a combinação (processo de sistematização de conceitos em um sistema de conhecimento) e a internalização (processo de incorporação do conhecimento explícito ao conhecimento tácito).

→ Teixeira Filho (1998) também acredita que atualmente o recurso econômico básico para a estratégia de atuação das organizações é o conhecimento, e não mais o capital de longo prazo, os recursos naturais, ou o trabalho. Entretanto, é preciso que esse conhecimento seja compartilhado, beneficiando as estratégias de atuação das organizações, que devem passar da competição para a cooperação.

Nessa mesma linha de raciocínio, O'Dell e Grayson Júnior (2000), argumentam que atualmente o conhecimento está sendo amplamente aceito como ingrediente básico da vantagem competitiva. Para a gestão do conhecimento, isso representa um desafio, pois é preciso deduzir a existência do conhecimento, entender sua capacidade e observar seu impacto sobre outras forças (concorrência, inovação, avaliação do mercado acionário, produtividade dos funcionários e excelência do produto).

Assim, a gestão do conhecimento pode ser vista como uma ferramenta para intensificar uma atividade organizacional ou um processo de aprimoramento, levando as organizações a tomar melhores decisões de maneira mais rápida, a desenvolver melhores produtos e a conquistar novos mercados.

Terra e Gordon (2002) também argumentam que, qualquer que seja o tipo de organização, elas estão percebendo cada vez mais que o recurso de conhecimento se tornou a chave para estabelecer vantagens competitivas duradouras. Assim, o papel da gestão do conhecimento passa a ser considerado como o esforço para melhorar o desempenho humano e organizacional, garantindo que todos dentro da organização tenham acesso ao conhecimento, quando, onde e da forma que necessitam.

Percebe-se, que há um consenso entre os autores de que o conhecimento é um recurso essencial para as organizações obterem vantagem competitiva. Logo, este conhecimento deve estar associado a uma ação, e sua gestão só fará sentido se estiver voltada para a melhoria do desempenho da organização.

Portanto, incentivar o compartilhamento do conhecimento garante a execução das tarefas de uma maneira melhor, levando a organização a obter vantagem competitiva.

Diante disso, as organizações devem estimular os seus funcionários a identificar, registrar e compartilhar as melhores práticas e os conhecimentos nas fábricas e nas unidades de negócios, preenchendo as lacunas do desempenho, reduzindo os custos de produção e melhorando o desempenho global (O'DELL; GRAYSON JÚNIOR, 2000).

### 2.1.2 O papel da tecnologia da informação na gestão do conhecimento

Uma questão que surge quando se fala em gestão do conhecimento, é o papel que a tecnologia desempenha neste processo. Sem dúvida, a tecnologia pode ajudar as pessoas a se comunicarem e a se relacionarem, porém, não pode gerar conhecimento e nem criar as condições necessárias para que ocorra o compartilhamento do mesmo, se as pessoas não perceberem as vantagens ou a necessidade de fazê-lo. Assim a tecnologia é um recurso necessário para a solução, mas é somente parte da resposta.

Carvalho e Ferreira (2001) relatam que a finalidade principal da tecnologia da informação consiste em ampliar o alcance e acelerar a velocidade de transmissão do conhecimento.

Para O'Dell e Grayson Júnior (2000), o crescimento explosivo e a rápida adoção das tecnologias da informação foram um grande catalisador do processo de compartilhamento do conhecimento. Entretanto, os autores argumentam que a gestão do conhecimento não depende da tecnologia da informação para tornar os processos mais eficientes, ela depende do conhecimento existente na mente das pessoas, e usa a tecnologia para facilitar seu compartilhamento. A tecnologia é necessária, mas não suficiente para possibilitar a transmissão do conhecimento.

Davenport e Prusak (1998) colocam que as ferramentas de gestão do conhecimento devem auxiliar o processo de coleta e estruturação do conhecimento de grupos de indivíduos, disponibilizando-o em uma base compartilhada por toda a organização. Com os *softwares* de gestão do conhecimento é possível otimizar o fluxo de conhecimento através de redes de comunidades, transformando a tecnologia em um meio e o conhecimento em uma mensagem. Os autores salientam que a gestão do conhecimento é muito mais do que a tecnologia, mas, com certeza, a tecnologia faz parte da gestão do conhecimento.

De acordo com Carvalho e Ferreira (2001) e Teixeira Filho (2000), as principais tecnologias para a gestão do conhecimento são: videoconferência, *groupware*, painéis eletrônicos, grupos de discussão, bases de dados *on line*, *cd-rooms*, Internet, Intranet, sistemas especialistas, agentes de pesquisa inteligentes, *data warehouse*, *data mining*, gerenciamento eletrônico de documentos, portais, *workflow*, sistemas para construção de bases inteligentes de

conhecimento, *business intelligence*, sistemas de mapas de conhecimento, ferramentas de apoio à inovação.

Por outro lado, estes autores argumentam que a gestão do conhecimento é uma questão que envolve essencialmente pessoas e processos. Assim, as tecnologias úteis para a gestão do conhecimento são aquelas que proporcionam a integração das pessoas, facilitando a transposição das fronteiras entre unidades de negócios, ajudando a prevenir a fragmentação das informações e permitindo a criação de redes globais para o compartilhamento do conhecimento.

Teixeira Filho (2000) salienta que o desafio para a área de tecnologia da informação é identificar, encontrar e/ou desenvolver e implementar tecnologias e sistemas de informação que sejam úteis para apoiar a comunicação empresarial, bem como a troca de experiências, procurando facilitar e incentivar as pessoas a se unirem, a participarem, a fazerem parte de grupos e a se renovarem em redes informais. A tecnologia da informação deve ser utilizada para facilitar as atividades essenciais para a evolução da organização, fornecendo meios para que as pessoas possam compartilhar problemas, perspectivas, idéias e soluções. O autor conclui dizendo, que o maior desafio da tecnologia da informação passa a ser a migração de uma posição de suporte a processos para o suporte a competências.

O'Dell e Grayson Júnior (2000) colocam que comprar sistemas e implementar arquiteturas sofisticadas não garante, por si só, o compartilhamento do conhecimento e das melhores práticas. Muitas vezes, tecnologia demais ou o tipo errado de tecnologia pode prejudicar qualquer iniciativa de gestão do conhecimento. Por fim, a tecnologia da informação deve ser usada com inteligência, procurando combinar os tipos e as necessidades de conhecimento com as aplicações certas.

Pode-se afirmar então, com base em Barroso e Gomes (1999), que a prática da gestão do conhecimento inclui a identificação e o mapeamento dos ativos ligados à empresa, a geração de novos conhecimentos visando a competição no mercado, e a disponibilização de grandes quantidades de informações e conhecimento, através do compartilhamento das melhores práticas e da utilização de tecnologia que pode tornar tudo isso possível.

Atualmente, estão disponíveis muitas soluções de tecnologia para suporte ao fluxo de conhecimento. Portanto, na escolha de uma tecnologia de gestão de conhecimento é importante identificar quais são as necessidades específicas de geração, disseminação e utilização do conhecimento. Assim, a tecnologia usada de maneira adequada atuará como facilitadora de qualquer iniciativa nesta área.

Além da tecnologia, no processo de gestão de conhecimento, atenção especial deve ser dada às pessoas. Pessoas com capacidade de aprender, experimentar, observar, analisar e disseminar as lições aprendidas passam a ser fatores chave de sucesso para as organizações, levando-as a obter melhores resultados e maior competitividade.

Por fim, de acordo com Costa et al. (2000), para a gestão do conhecimento é preciso considerar as pessoas (geradoras do conhecimento), a informação (matéria-prima para gerar o conhecimento) e a tecnologia da informação (suporte à informação e ao conhecimento).

## 2.2 PORTAIS

As informações disponíveis na Internet diferem das informações disponíveis em outras fontes por sua acessibilidade via redes de computadores, sua estrutura, seu dinamismo e seus métodos de disponibilização. As peculiaridades das informações eletrônicas em rede podem gerar tanto vantagens como barreiras e desafios. De um lado o usuário tem acesso a um acervo sem tamanho, por outro lado, recuperar informações neste acervo exige, além dos aspectos técnicos de conexão, conhecimento sobre o funcionamento de vários aplicativos, de outros idiomas e uma dose muito grande de paciência para encontrar as informações que atendam as suas necessidades (CENDÓN, 2000).

Numa tentativa de facilitar o acesso às informações contidas em documentos espalhados nos *sites* da Internet surgiram os mecanismos de busca.

Inicialmente estes mecanismos localizavam informações a partir de buscas booleanas e navegação associativa entre *links*. Para reduzir mais o tempo de busca na Internet e ajudar os usuários menos experientes a localizar as informações, vários *sites* de busca passaram a filtrar

estes *sites* e documentos em grupos pré-configurados de acordo com o seu conteúdo (turismo, esporte, notícias, etc.). O passo seguinte foi a integração de outras funções, tais como listas de discussão, *chats* em tempo real, personalização dos *sites* de busca e acesso a conteúdos especializados e comerciais. Com a popularização dessa nova concepção de *sites* de busca na Internet, surgiu o termo portal (DIAS, 2001).

Moreira (2000) ressalta que o portal não é somente um mecanismo que possibilita localizar recursos na Internet por meio de suas bases de dados, mas também um local onde diversas informações factuais ou utilitárias, de diferentes naturezas podem ser encontradas. Ele é o responsável por conduzir o usuário aos meandros da Internet ao mesmo tempo em que pretende tornar-se seu 'porto seguro'.

Para Teixeira Filho (1999, p. 1) o interesse inicial nos portais vem do fato que há um grande volume de informações e *sites*, e qualquer recuperação manual se torna impossível. Assim, um portal se torna um ponto de partida natural de navegação, oferecendo um mecanismo de busca, que indexa por assunto e palavras-chave os *sites* existentes na Internet. O autor afirma que o conceito de portal se refere a "um ponto de entrada unificado no universo de informações sobre um determinado assunto, ou de uma determinada organização."

Kleinschmidt e Goodman (2002, p. 1) conceituam portal como "a primeira página que aparece no *browser* quando o internauta entra na Internet ou seu ponto de entrada para a Internet."

De acordo com Arce e Pérez (2001, p. 223), portal é "uma página na *web* que agrega conteúdos e funcionalidades, organizados de tal maneira que facilitam a navegação, proporcionam ao usuário um ponto de entrada na rede com um amplo leque de opções de serviços e produtos."

Freitas et al. (2000, p. 71) relatam que um portal é "um aplicativo ou dispositivo que possui uma *interface* personalizada e adaptada para as pessoas descobrirem, usarem e intercambiarem aplicativos e conteúdos relevantes."

Para Angulo e Albertin (2000, p. 2) um portal é "um *web site* tido como ponto inicial de 'alguma coisa', sendo que este 'ponto inicial' apresenta um grande número de usuários e também um conjunto de produtos e serviços que criam valor para os consumidores." Ao

mesmo tempo, os portais permitem às organizações anunciar ou vender seus produtos para um grande número de clientes, em um único *web site*. Por oferecerem diversos tipos de serviços e produtos, os portais para a Internet tornam a experiência na *web* mais simples para o usuário, uma vez que é preciso ir a um único destino para encontrar diferentes opções de serviços e produtos.

Diante dos conceitos apresentados, pode-se concluir que portal é uma porta de entrada para o conteúdo *on line*, onde as pessoas podem encontrar informações sobre todo e qualquer assunto, além de uma variada opção de produtos e serviços que atendam as suas necessidades, sejam profissionais, culturais ou de lazer.

Teixeira Filho (1999) argumenta que os portais genéricos como o Yahoo! o Excite, o Cadê e o Aonde apresentam problemas por não serem especializados por assunto, área da indústria, ou qualquer outro critério. Com isso, o nível de resposta a uma pesquisa específica na maioria das vezes, não é satisfatório. Recupera-se uma grande quantidade de *links*, porém, muitos não são relevantes para o assunto que está sendo pesquisado. Mesmo assim, estes portais acabam sendo usados por um grande número de usuários, sendo hoje talvez os *sites* de maior tráfego na Internet e, conseqüentemente, aqueles que atraem maior número de usuários.

De acordo com Dias (2001), a evolução dos portais chamou a atenção da comunidade corporativa, que vislumbrou a possibilidade de utilizar esta mesma tecnologia para organizar e facilitar o acesso às informações internas da empresa e para outras áreas informacionais e comerciais. Assim surgiu um novo conceito, o portal corporativo, que está se tornando uma ferramenta essencial para as empresas obterem vantagem competitiva.

Nesta mesma linha de raciocínio, Terra e Gordon (2002) enfatizam que os portais corporativos estão mudando a forma como a informação e as responsabilidades de colaboração são compartilhadas numa organização: de um foco estreito, funcional e não coordenado, para uma abordagem ampla e de colaboração.

Estes autores ainda colocam que os portais integram muitas características que estão relacionadas a processos de gestão do conhecimento, tais como personalização e busca, acesso a fontes de informações num ambiente propício à rede, informação interna e externa,

comunicação e colaboração, facilidade de publicação e acesso a uma grande quantidade de dados, informações e conhecimentos.

Considerando essa visão corporativa, portal pode ser compreendido como uma ferramenta para a solução dos problemas de distribuição de informações, tanto internas quanto externas, agilizando o processo de tomada de decisão da organização, aumentando a produtividade, reduzindo custos e, conseqüentemente, gerando maior competitividade.

### 2.2.1 Tipos de portais

Dias (2001) argumenta que por ser um conceito muito recente, a terminologia relacionada com os portais ainda não se estabilizou. Na literatura, termos como portal corporativo, portal de informações corporativas, portal de negócios e portal de informações empresariais são utilizados como sinônimos.

Segundo Firestone (2000 apud DIAS, 2001), o processo de definição de um portal é um processo político, ou seja, uma tentativa de convencer a comunidade de usuários e os investidores da área de tecnologia da informação a adotar uma definição que seja mais adequada que outra, favorecendo interesses de um ou outro consultor/fornecedor.

Dias (2001), através de extensa revisão bibliográfica sobre portais constatou que eles podem ser classificados de duas formas. Uma em relação ao contexto de sua atualização: público ou corporativo e a outra, em relação às suas funções: suporte à decisão e/ou processamento cooperativo.

O portal público, também denominado portal Internet, portal *web* ou portal de consumidores, oferece ao usuário uma única *interface* a toda rede de servidores que compõem a Internet, tendo como função principal atrair para o seu *site* o público que navega na Internet, estabelecendo um relacionamento unidirecional com os seus usuários, constituindo-se numa mídia adicional para o *marketing* de produtos (DIAS, 2001).

Na mesma linha de pensamento, Angulo e Albertin (2000) destacam que portal público (ou portal Internet) é a primeira página que aparece no *browser*, quando o usuário (internauta)

entra na Internet. Pelo fato de oferecerem serviços e produtos genéricos, estes portais não têm um público alvo determinado. Exemplos desta categoria são: Yahoo!, Netscape, America Online e Universo Online.

Por sua vez, o portal corporativo é considerado como uma evolução do uso das Intranets, com a incorporação de novas ferramentas que possibilitam a identificação, a captura, o armazenamento, a recuperação e a distribuição de grande quantidade de informação de diferentes fontes, internas e externas para as pessoas de uma instituição. Suporte à decisão e processamento cooperativo são as funções mais importantes deste tipo de portal (DIAS, 2001; REYNOLDS; KOULOPOULOS, 1999).

Os portais com ênfase em suporte à decisão auxiliam executivos, gerentes e analistas de negócios no acesso às informações corporativas para a tomada de decisões, dando pouca ou quase nenhuma ênfase ao processamento cooperativo. Entre eles destacam-se os seguintes:

- *portal de informação ou de conteúdo*: organiza grandes acervos de conteúdo, a partir dos temas ou assuntos nele contidos, conectando as pessoas às informações. Nesta categoria de portal são incluídos os mecanismos de busca e os portais públicos (DIAS, 2001; MURRAY, 1999);
- *portal de negócio*: fornece o ponto de partida central aos aplicativos do processo de tomada de decisão e de gerência de conteúdo, conectando os usuários a informações estruturadas e não estruturadas. A função deste tipo de portal é disponibilizar aos usuários corporativos informações necessárias para a tomada de decisões, tais como relatórios, pesquisas, correio eletrônico, etc. (DAVYDOV, 2000; DIAS, 2001);
- *portal de suporte à decisão*: permite aos usuários organizar e encontrar informações da corporação em um conjunto de sistemas que constituem a cadeia produtiva de informações de negócios. Este tipo de portal utiliza ferramentas inteligentes e aplicativos analíticos para capturar informações que se encontram armazenadas em bases de dados operacionais, no *data warehouse* da corporação ou em sistemas externos à organização. A partir destas informações gera relatórios, gráficos e indicadores de desempenho das operações de negócios e os distribui eletronicamente aos tomadores de decisão da empresa (WHITE, 1999 apud DIAS, 2001).

Os portais com ênfase em processamento cooperativo lidam com informações da cadeia produtiva tradicional, armazenadas e manipuladas por aplicativos corporativos e com informações geradas por grupos de indivíduos fora dessa cadeia. Esta categoria de portal foi a que menos se desenvolveu no mercado, até o momento (DIAS, 2001). Entre os portais com ênfase em processamento cooperativo destacam-se:

- *portal cooperativo ou para processamento cooperativo*: utiliza ferramentas cooperativas de trabalho em grupo (*groupware*) e de fluxo de tarefa/documento (*workflow*) para possibilitar o acesso às informações geradas por indivíduos ou grupos. Geralmente, as informações manipuladas por este tipo de portal não são estruturadas e personalizadas, encontrando-se sob a forma de textos, memorandos, mensagens de correio eletrônico, gráficos, boletins informativos, páginas *web* e arquivos multimídia (DIAS, 2001);
- *portal de especialistas*: propõe um meio de comunicação e troca de experiências entre pessoas especializadas em áreas específicas do conhecimento, por meio de comunicação em tempo real, educação a distância e manutenção de cadastro automático de especialistas (DIAS, 2001; MURRAY, 1999).

Os portais de suporte à decisão e processamento cooperativo são portais abrangentes. Conseguem aliar funções de suporte à decisão e processamento cooperativo e conectam os usuários às informações e às pessoas necessárias aos processos de negócios da empresa. Consolidam aplicativos de gerência de conteúdo, processamento de decisões, *groupware*, *workflow*, correio eletrônico, inteligência de negócios, sistemas especialistas, etc., num mesmo ambiente (DIAS, 2001). Entre estes portais destacam-se os seguintes:

- *portal de conhecimento*: é um ponto de convergência dos portais de informações, cooperativos e de especialistas, com capacidade de implementar tudo o que os outros tipos de portais implementam e de fornecer conteúdo personalizado conforme a atividade de cada usuário (DIAS, 2001; MURRAY 1999). Os portais do conhecimento podem ser a resposta para a imensa quantidade de informações das empresas, permitindo acesso a todo tipo de informação necessária e significativa ao usuário, de forma fácil e rápida (EVANS; YOKOYAMA, 1999). A informação disponível em um portal deve ser adaptada ao usuário, customizada. Portanto, portais de conhecimento devem ser mais do que simples *sites* de informação, devem ser eles mesmos

customizados, possuindo características que os diferenciem dos seus congêneres da rede mundial (STRAUHS; ABREU; RENAUX, 2000);

- *portal de informações empresariais*: integra dados não estruturados, mantidos em arquivos textuais, relatórios, mensagens de correio eletrônico, gráficos, imagens, etc. aos dados estruturados das bases de dados do *data warehouse*, fornecendo acesso às informações da instituição a partir de uma *interface* individualizada, disponível na rede hipertextual corporativa - a Intranet (DIAS, 2001). O potencial deste tipo de portal está no fato de proporcionar um recurso único e centralizado onde os usuários podem encontrar, extrair e analisar a informação que prolifera na empresa, acessando um conteúdo personalizado e essencial para a tomada de decisões nos negócios (EIP..., 2000). Muitas vantagens podem ser obtidas com a instalação de um portal de informações empresariais. Entre elas destacam-se: aumento da taxa de retorno de investimentos em projetos de *business intelligence*, aumento da eficiência da organização, aumento da produtividade, redução de custos, *e-commerce*, aumento dos poderes do usuário final, tornando-o mais independente e criação de novas perspectivas para a gestão da informação (ENTERPRISE..., 2001).

Diante dos tipos de portais apresentados, é possível concluir que são muitos os portais existentes, alguns até com características bastante semelhantes, sendo que, a característica principal, presente na maioria deles, está relacionada com a oferta de produtos e serviços ao público, atendendo desde as necessidades mais genéricas até as mais específicas. Desta forma, a definição por um tipo de portal estará sempre associada ao tipo de produto e/ou serviço que ele vai oferecer e ao público que ele pretende alcançar, procurando sempre satisfazer os interesses das organizações.

### 2.2.2 Componentes funcionais de um portal corporativo

Os componentes de um portal corporativo são os elementos que agregam benefícios funcionais distintos ao usuário. Para ajudar na escolha de um portal, muitos consultores têm publicado artigos e relatórios contendo requisitos mínimos que um portal corporativo deve ter.

O quadro 1 resume os principais requisitos esperados de um portal corporativo.

<b>Requisitos</b>	<b>Descrição</b>
Facilidade para usuários ocasionais	localização e acesso fácil a informações corretas independente do local onde estão armazenadas. Deve ser intuitivo, minimizando as necessidades de treinamento
Classificação e pesquisa intuitiva	capacidade de indexar e organizar as informações da organização. O mecanismo de busca deve refinar e filtrar as informações, suportar palavras-chave e operadores booleanos, além de apresentar o resultado da pesquisa de maneira compreensível
Compartilhamento cooperativo	permissão aos usuários para publicar, compartilhar e receber informações de outros usuários; possibilidade de interação entre pessoas e grupos na organização e permissão ao usuário para especificar, na publicação, quais os usuários e grupos que poderão acessar seus documentos
Conectividade universal aos recursos da informação	possibilidade de amplo acesso a todo e qualquer recurso informacional, suportando conexão com sistemas heterogêneos, tais como correio eletrônico, banco de dados, sistemas de gestão de documentos, servidores <i>web</i> , etc.; capacidade de gerenciar vários formatos de dados estruturados e não estruturados
Acesso dinâmico aos recursos informacionais	permissão de acesso dinâmico às informações nele armazenadas, por meio de sistemas inteligentes, possibilitando ao usuário receber sempre informações atualizadas
Roteamento inteligente	direcionamento automático de relatórios e documentos a usuários selecionados
Ferramenta de inteligência de negócios integrada	integração dos aspectos de pesquisa, relatório e análise dos sistemas de inteligência de negócios para atender a demanda de informações dos usuários
Arquitetura baseada em servidor	presença de uma arquitetura cliente-servidor, para suportar um grande número de usuários e grandes volumes de informações, serviços e sessões concorrentes
Serviços distribuídos	distribuição dos serviços por vários computadores ou servidores, visando melhor balanceamento de carga de processamento
Flexibilidade nas permissões de acesso	o administrador do portal deve ser capaz de definir permissões de acesso para usuários e/ou grupos de organizações, por meio dos perfis de usuário
<i>Interfaces externas</i>	capacidade de comunicação com aplicativos e sistemas externos
<i>Interfaces programáveis</i>	capacidade de ser "chamado" por outros aplicativos, tornando pública sua <i>interface</i> programável
Segurança	serviços de segurança, como criptografia, autenticação, <i>firewall</i> , etc., a fim de salvaguardar as informações corporativas e evitar acessos não autorizados
Fácil administração	gerenciamento de todas as informações corporativas e monitoramento do funcionamento do portal de forma centralizada e dinâmica. Facilidade na instalação, configuração e manutenção, aproveitando, sempre que possível, a base instalada de <i>hardware</i> e <i>software</i> existente na organização
Customização e personalização	o administrador do portal deve ser capaz de customizá-lo conforme políticas e expectativas da organização, assim como os próprios usuários devem ser capazes de personalizar sua <i>interface</i> para facilitar e agilizar o acesso às informações desejadas

Quadro 1 - Requisitos funcionais de um portal corporativo

Fonte: Dias (2001); Eckerson (1999)

A estes requisitos, Dias (2001) acrescenta as seguintes habilidades:

- gerenciamento do ciclo de vida das informações, estabelecendo níveis hierárquicos de armazenamento e descarte das informações ou documentos quando não são mais necessários;
- localização de especialistas na organização, de acordo com o grau de conhecimento exigido para o desempenho de determinadas tarefas;
- satisfação das necessidades de informação de todos os tipos de usuários da organização e possibilidade de troca de informações com clientes, fornecedores, revendedores, etc., fornecendo uma infra-estrutura informacional adequada também para o comércio eletrônico.

Mesmo sendo uma tecnologia recente, fornecedores e consultores têm apontado muitos benefícios associados aos portais corporativos. Entre eles pode-se destacar a facilidade de acesso às informações que se encontram nos diferentes sistemas, arquivos e bases de dados institucionais, que está intrinsecamente relacionada à sua facilidade de uso, aprendizado e satisfação do usuário, ou seja, à usabilidade da *interface web* (DIAS, 2001).

De acordo com a ISO 9241 (1998 apud DIAS, 2001) o conceito de usabilidade está relacionado com a capacidade de um produto ser usado por usuários específicos, procurando alcançar os objetivos definidos com eficácia, eficiência e satisfação num contexto específico de uso.

Reforçando o aspecto usabilidade, Shackel (1985), citado por Heemann (1997), relata que ela envolve os quatro componentes principais de uma situação de trabalho: usuário, tarefa, sistemas e ambiente. Assim, um bom projeto, que vise a usabilidade de um produto, depende da harmonia e da dinâmica de atuação destes quatro componentes.

De acordo com Terra e Gordon (2002), há no mercado dezenas de opções de *softwares* oferecendo soluções de portal, e escolher a melhor opção não é tarefa fácil. Portanto, como coloca White (1999 apud DIAS, 2001), na escolha da melhor tecnologia é necessário considerar a arquitetura básica de qualquer portal, que é composta por um navegador *web* (*software* que interpreta e apresenta na tela do computador as páginas *web*) e um servidor *web*

(computador integrante de uma rede cliente-servidor, responsável pelos serviços relacionados com a Internet/Intranet).

Por fim, em relação aos requisitos funcionais dos portais, nota-se que o enfoque maior está na relação entre o portal e o usuário. Assim, quando se projetam portais, estes requisitos devem ser levados em conta, visando sempre a otimização da interação entre usuário e portal, oferecendo informações de forma mais eficiente através de uma *interface* amigável.

### **2.3 BASES DE DADOS**

Ao disponibilizar dados e informações a qualquer momento e de inúmeras fontes, a Internet gera uma sobrecarga de informações, de forma desorganizada e de difícil recuperação.

Visando facilitar a busca e o acesso às informações e através da interação das redes de telecomunicações, as bases de dados vêm proliferando cada vez mais, incluindo materiais em suportes diferentes, disponibilizando a informação potencial dos arquivos e fichários manuais, podendo ser pesquisadas via terminais locais ou remotos (BERTHOLINO, 1999; HEEMANN, 1997; TEIXEIRA; SCHIEL, 1997).

O principal objetivo de uma base de dados, de acordo com Albrecht e Ohira (2000), é fornecer informação atualizada, precisa, confiável e de acordo com a demanda da comunidade usuária.

Assim, as bases de dados se constituem como ferramentas fundamentais para a recuperação, a organização, a estruturação e a disponibilização de informações (PEREIRA et al., 1999).

Antes de apresentar os conceitos sobre bases de dados, é importante definir o que são documentos.

Guinchat e Menou (1994, p. 53) definem documento como “um conjunto de suporte de informação e dos dados nele registrados, que podem servir para consulta, estudo ou prova.”

Para Davenport (1998, p. 56), documentos são “meios que estruturam e apresentam as informações.”

Para o propósito desta dissertação e com base nas definições destes autores, documento é todo e qualquer tipo de suporte que estrutura e apresenta as informações, podendo ser livros, dissertações/teses, jornais, revistas, *cd-rooms*, vídeo, relatórios, projetos, atas, *sites* e páginas da Internet e arquivos digitais. Quanto ao seu conteúdo, estes documentos contemplam texto, som e imagem.

De acordo com Cianconi (1987, p. 54), há uma grande confusão em relação aos conceitos e a terminologia na literatura, que trata indiscriminadamente banco de dados como base de dados. Para a autora, base de dados é “um conjunto de dados inter relacionados, organizados de forma a permitir a recuperação de informações” e banco de dados pode ser entendido como “um conjunto de bases de dados.”

Conforme Rowley (2002, p. 125), base de dados é “uma coleção geral e integrada de dados junto com a descrição deles, gerenciada de forma a atender a diferentes necessidades de seus usuários.”

Para Cunha (1994, p. 183) as bases de dados são “fontes de informação computadorizadas que podem ser pesquisadas em um modo interativo ou conversacional por intermédio de um terminal de computadores, telex ou mesmo um microcomputador.”

Camarão (1994 apud HEEMANN, 1997, p. 10) conceitua base de dados como

a consolidação de todos os dados, tidos como operáveis sob qualquer forma e armazenados em qualquer local; um conjunto de dados inter relacionados, armazenado num determinado arquivo, conforme um esquema predefinido e disponível para uma ou mais aplicações; um conjunto de dados fundamentais para um determinado sistema ou serviço; um arquivo onde são armazenadas informações permitindo atualização e recuperação dos dados.

Para Cunha (2001, p. 35), base de dados é “a expressão utilizada para indicar a coleção de dados que serve de suporte a um sistema de recuperação de informações.”

Ryan (1992) citado por Heemann (1997) argumenta que base de dados é uma coleção de itens de informação com um tema comum. O autor destaca também os seguintes elementos que fazem parte uma base de dados:

- *registros*: blocos numerados construídos nas bases de dados, também conhecidos por referências, documentos, artigos e itens;
- *campos*: partes dos registros, como títulos, descritores, etc., também chamados seções ou parágrafos;
- *termos*: palavras, frases, números e símbolos, ou combinações destes, que formam os campos, também chamados de conteúdo de campo.

Para Rowley (2002), os elementos que constituem uma base de dados são: registros, dados, elementos de dados e informações. O registro é uma informação que a base contém e diz respeito a um documento ou item, podendo ser composto de vários campos. Os tipos de campos utilizados, seu tamanho e a quantidade de campos de um registro são escolhidos de acordo com uma aplicação específica, e podem ser de tamanho fixo e variável. Os campos de tamanho fixos contêm o mesmo número de caracteres em cada registro, não sendo necessário sinalizar para o computador onde começa e termina cada campo. Por sua vez, um campo de tamanho variável armazena dados de tamanho variável, apresentando diferentes extensões em diferentes registros, impossibilitando o computador de reconhecer quando um campo termina e outro começa, sendo necessário então, sinalizar o início e o fim dos campos.

Diante dos conceitos sobre bases de dados apresentados pelos autores, pode-se concluir que base de dados é o local onde se armazenam dados, inter relacionados entre si e gerenciados de forma a permitir sua posterior recuperação, com acesso via terminais locais ou remotos. Assim, as bases de dados se constituem em ferramentas essenciais para a recuperação, armazenamento e disponibilização de informações, tendo como objetivo principal atender às necessidades informacionais dos usuários.

Graças ao desenvolvimento das tecnologias da informática e teleprocessamento, a partir da década de 70, o acesso à informação se tornou mais amplo, fazendo com que a mesma passasse a ser vista como um produto comercializável, além de um bem social. Surgem então, as bases de dados *on line*, comercializadas em larga escala nos Estados Unidos e na Europa, através de sistemas e serviços de armazenagem e recuperação de informações. Estes sistemas e serviços, desenvolvidos com recursos técnicos e materiais da Ciência da Computação, são

orientados e avaliados pela Ciência da Informação, visando eficiente armazenagem, recuperação e disseminação de informações à comunidade usuária (CIANCONI, 1987).

### 2.3.1 Tipos de bases de dados

Geralmente o tipo de base de dados é determinado pela natureza das informações que elas contêm.

De acordo com Rowley (2002), as bases de dados podem ser classificadas em bases de dados de referência e de fontes, e podem ser acessadas por meio de um serviço de buscas em linha, ou localmente, em *cd-room*. As mesmas são descritas a seguir.

*Bases de dados de referência:* encaminham ou remetem o usuário para uma outra fonte, como um documento, uma instituição, ou uma pessoa, a fim de obter informações adicionais ou acessar o texto integral de um documento. Entre estas bases destacam-se:

- *bases de dados bibliográficos:* incluem citações ou referências bibliográficas, podendo incluir também resumo informativo e palavras-chave do assunto pelas quais o documento foi indexado. Trazem a informação necessária para localizar um documento: o título, o autor, o nome do periódico, o volume, o número e a data. Estas bases informam sobre o que foi publicado e onde se publicou;
- *bases de dados catalográficos:* mostram os documentos que as bibliotecas e/ou centro de informação possuem no acervo, sem informações adicionais sobre o conteúdo dos documentos;
- *bases de dados referenciais:* referenciam informações ou dados, como cadastros e guias com nomes e endereços de instituições.

*Bases de dados de fontes:* contém os dados originais e constituem um tipo de documento eletrônico. Após consultar uma base de dados de fontes, o usuário terá em mãos as informações de que precisa, não precisando ir buscá-las numa fonte original. Os dados podem estar disponíveis tanto em formato legível por computador, quanto em formato impresso. Estas bases podem ser agrupadas segundo seu conteúdo, em:

- *bases de dados numéricos*: contém dados numéricos de vários tipos, inclusive dados estatísticos e de resultados de pesquisas;
- *bases de dados de texto integral*: contém notícias de jornal, especificações técnicas, programas de computador, etc. Disponibilizam o texto original completo;
- *bases de dados textuais e numéricos*: contém uma mistura de dados textuais e numéricos e dados de manuais;
- *bases de dados multimídia*: incluem informações armazenadas numa mescla de diferentes tipos de meios (som, vídeo, fotografias, textos, animação, etc.).

Rowley (2002) argumenta que, embora as bases de dados bibliográficos informatizadas existam há mais de vinte anos, seus elementos básicos ainda retêm a marca de sua origem impressa, que geralmente, era uma publicação de resumos ou um índice aos quais essas bases estavam vinculadas. Em relação ao conteúdo das bases de dados de fontes, a autora coloca que o mesmo pode ser tão variado quanto o conteúdo de um livro impresso. Estas bases se beneficiam do fato de não ter as mesmas limitações físicas dos materiais impressos, e sendo multimídia, podem incluir, além de textos e dados numéricos, tabelas, figuras, imagens, programas de computador, mapas e gráficos.

Guincht e Menou (1994) dividem as bases de dados a partir da natureza dos dados em:

- *bases de dados de referências bibliográficas*: artigos, teses, relatórios, patentes e normas;
- *bases de dados diretamente utilizáveis*: numéricas, factuais e textuais.

Amat i Noguera (1990 apud TARGINO; GARCIA 2000) classificam as bases de dados de acordo com o conteúdo que armazenam e disseminam em:

- *bibliográficas*: contém as referências bibliográficas dos documentos e algumas vezes, indicam sua localização e forma de acesso;
- *de referência*: listam as fontes e os centros que possuem a informação;
- *não bibliográficas*: reúnem documentos visuais, mapas, fotografias, patentes, dispositivos, etc., o que permite sua inclusão na categoria das bases de referência;
- *textuais*: contém o texto completo do documento ou parte dele;
- *factuais*: armazenam informações estatísticas, numéricas, séries cronológicas ou outro tipo de informação numérica ou alfanumérica.

Cianconi (1987), por sua vez, classifica as bases de dados utilizando outros critérios, além da natureza das informações:

- quanto a disponibilidade, em bases de acesso restrito ou privadas e de acesso público, ou públicas;
- quanto a origem, em bases de dados governamentais ou privadas;
- quanto a natureza das informações, em bases de dados bibliográficas ou referenciais e factuais ou fontes.

Para Cunha (2001) as principais bases de dados são:

- *bases de dados bibliográficas*: incluem referências bibliográficas e resumos;
- *bases de dados textuais*: incluem textos completos de artigos de periódicos, jornais ou outros tipos de documentos.

Cendón (2002) considera três tipos de bases de dados:

- *bases de dados bibliográficas ou referenciais*: contém registros bibliográficos que permitem ao usuário localizar determinada publicação, podendo fornecer também o resumo dos documentos;
- *bases de dados em texto completo*: contém o documento completo;
- *bases de dados factuais*: fornecem respostas a perguntas que não visam a obter uma resposta bibliográfica.

De acordo com Targino e Garcia (2000), com a evolução das novas tecnologias de informação e de comunicação, as bases de dados, referem-se a uma coleção de registros armazenados em suporte magnético, acessível via computador. Assim, as bases podem ser classificadas também de acordo com o suporte em:

- bases de dados *on line* ou de recuperação em linha;
- bases de dados privadas;
- bases de dados em videotexto;
- bases de dados em videodisco ou em *cd-room*.

Percebe-se, que a determinação do tipo de base de dados está ligada principalmente à natureza dos dados que ela contém, podendo ser bibliográficas, de referência, de fontes, de dados

numéricos, de texto integral. Além da natureza dos dados, é preciso ser levada em conta a finalidade da base e o público a que ela se destina.

### 2.3.2 Formatos de registro

O formato se refere a maneira segundo a qual os dados são formatados, para posterior processamento. Um registro catalográfico significa a informação tradicionalmente apresentada numa ficha catalográfica, incluindo a descrição de um item, as entradas principal e secundárias, o cabeçalho de assuntos e a classificação ou número de chamada (não necessariamente nessa ordem) (FURRIE, 2000).

De acordo com Rowley (1994) o formato de registro MARC foi desenvolvido pela *Library of Congress* e a *British Library* para a comunicação de descrições bibliográficas em formato legível por computador, de tal maneira que os registros pudessem ser reformatados para atender a qualquer objetivo pretendido. Os primeiros ensaios, realizados por volta de 1966, pela *Library of Congress*, utilizaram o formato MARC I, sendo substituído em 1967 pelo MARC II ou MARC, como é denominado geralmente.

O formato UNIMARC é um formato de intercâmbio internacional de registros MARC. As instituições nacionais que produzem registros MARC utilizam normas locais, na esfera dos respectivos países, e para intercâmbio com outros países reformatam seus registros em conformidade com o UNIMARC (ROWLEY, 1994).

Furrie (2000) destaca que utilizar um padrão de registro é importante para evitar duplicação de trabalho e permitir o melhor compartilhamento de informações. A opção pelo uso do MARC possibilita a aquisição de dados confiáveis e previsíveis, além de permitir a utilização de sistemas de automação disponíveis comercialmente. Há muitos sistemas de automação, para bibliotecas e/ou centros de informação de qualquer tamanho, que foram projetados com o formato MARC. Estes sistemas são mantidos e constantemente aprimorados pelos fornecedores, possibilitando maior benefício dos avanços da informática. O padrão MARC também permite que um sistema automatizado seja substituído por outro com a certeza que os dados continuarão compatíveis.

Robredo e Cunha (1994) informam que algumas variantes do formato MARC foram desenvolvidas em outros países, ultrapassando o âmbito geográfico dos Estados Unidos. No Reino Unido foi desenvolvido o formato UK MARC. O INTERMARC foi elaborado conjuntamente pela França, Bélgica e Reino Unido. O Canadá desenvolveu o Canadian MARC. No Brasil, foi desenvolvido o formato CALCO.

Conforme Rowley (1994) grande parte dos projetos centralizados e cooperativos de catalogação leva em consideração e provavelmente emprega o formato de registro MARC. Entretanto, este grau de padronização já não é a praxe fora desta área de aplicação específica. Há duas categorias diferentes de sistemas: bases de dados externas e sistemas locais suportados por pacotes de *software*.

Em relação às grandes bases de dados, esta autora ressalta que são poucas as pressões para que seja aceito um formato padrão, e geralmente cada produtor acaba adotando um formato de registro que seja adequado às bases que produz. Há casos em que uma mesma base de dados apresenta diferentes formatos de registro, conforme o hospedeiro onde se ache instalada. É comum também, a presença de bases de dados de texto integral, que exigem um formato de registro diferente dos registros bibliográficos para apresentar as informações de maneira adequada. Os formatos de registro que se encontram em sistemas locais suportados por pacotes de *software* são inúmeros e variados. Alguns desses pacotes oferecem sistemas que aceitam um formato de registro MARC ou que produzem registros compatíveis com o mesmo. Já outros não oferecem esta opção. Normalmente, os pacotes oferecem a seus clientes a oportunidade de desenvolver um formato de registro que seja adequado a uma aplicação específica.

Ainda de acordo com Rowley (2002), com o objetivo de facilitar a comunicação de dados bibliográficos entre os setores da comunidade de informação especializada, foi projetado o Formato Comum de Comunicação (*Common Communication Format - CCF*). Desenvolvido em 1978, pela UNESCO, o CCF, em comum com o MARC, constitui uma implementação específica da norma ISO 2709. Assim, o CCF:

- especifica um pequeno número de elementos de dados obrigatórios, aceitos como essenciais para a identificação de um item;

- proporciona elementos obrigatórios, suficientemente flexíveis para suportar maneiras diferentes adotadas na descrição bibliográfica;
- proporciona vários elementos optativos, úteis na descrição de um documento de uma forma mais coerente com as práticas da instituição que cria o registro;
- permite à instituição onde se originou o registro incluir elementos não-padronizados, porém, considerados úteis em seu sistema;
- proporciona um mecanismo de ligação ou conexão entre os registros, ou segmentos dos registros, sem impor à instituição que lhes dá origem nenhum método uniforme relativo ao tratamento de grupos de registros ou elementos de dados.

Por fim, Rowley (2002) destaca que há muitos formatos de registro, mas é muito difícil encontrar aqueles que sejam suficientemente parecidos para serem processados pelos mesmos programas de computador. Dependendo de sua origem, as descrições bibliográficas diferem muito entre si. Assim, tem sido difícil, ou praticamente impossível reunir num único arquivo registros bibliográficos oriundos de origens diferentes.

Na mesma linha de raciocínio, Robredo e Cunha (1994) ressaltam que, no Brasil, a falta de uma política de padronização foi uma das causas prováveis para o aparecimento de vários sistemas independentes, concebidos sem levar em conta os diversos padrões que pouco a pouco iam se desenvolvendo em nível internacional, até constituir um corpo coerente de normas e recomendações que possibilitassem o intercâmbio e a colaboração entre bibliotecas e/ou sistemas de informações.

Diante do exposto sobre formatos de registro, constata-se que utilizar um padrão de registro no desenvolvimento de uma base de dados é extremamente importante, pois isso irá garantir um melhor compartilhamento de informações e evitará duplicação de trabalho.

Cabe ressaltar, que não se pretendeu aqui, explorar a fundo os formatos de registros nacionais e internacionais existentes, e sim destacar a importância da padronização dos mesmos que possibilitará ou não, o intercâmbio e o compartilhamento de informações entre bibliotecas e/ou sistemas de informações.

### 2.3.3 Indexação

Segundo Guinchat e Menou (1994, p. 175), a indexação é “a operação pela qual escolhe-se os termos mais apropriados para descrever o conteúdo de um documento [...]. É a operação central do sistema para a armazenagem e pesquisa das informações.”

Para Rowley (2002, p. 162), indexação é “o processo de atribuir termos ou códigos de indexação a um registro ou documento, termos ou códigos esses, que serão úteis posteriormente na recuperação do documento ou registro.”

Conforme Guinchat e Menou (1994), o nível de indexação varia conforme as necessidades e as possibilidades da unidade de informação, podendo ser:

- *genérica*: assemelha-se muito a classificação, com a diferença que ela é geralmente múltipla (identifica vários assuntos), enquanto que a classificação é, em geral, única (identifica o assunto principal do documento);
- *profunda*: quando diz respeito a todos os assuntos descritos de forma bastante detalhada, com mais de dez descritores;
- *exaustiva*: relaciona-se a totalidade do documento, praticamente frase por frase;
- *seletiva*: quando retém apenas as informações que possam interessar aos usuários de uma unidade documental específica.

Yates (1996 apud LOH, 1997) destaca três tipos de indexação:

- *indexação tradicional*: onde uma pessoa determina os termos que descrevem ou caracterizam os documentos, os quais farão parte do índice de busca. Este tipo de indexação é encontrado na maioria das bibliotecas. O maior problema que apresenta está relacionado a intervenção humana que pode gerar falhas no índice, além de ser uma atividade difícil e que consome muito tempo;
- *indexação full-text (ou indexação do texto completo)*: onde todos os termos que compõe o documento fazem parte do índice. Este tipo de indexação procura indexar todos os termos, sem estruturas hierárquicas entre eles, gerando índices muito volumosos;

- *indexação por tags (por partes do texto)*: onde apenas algumas partes do texto (as consideradas mais importantes ou mais caracterizadoras) são escolhidas, automaticamente, para gerar as entradas no índice. Este tipo de indexação procura indexar somente as partes relevantes do documento, procurando por marcas (*tags*) que identificam estas partes mais importantes. O maior problema apresentado é que os tipos de documentos precisam ser analisados previamente, isto é, deve-se conhecer que tipo de conteúdo compõe os documentos e quais as marcas que identificam as partes relevantes dos mesmos.

Cendón (2000) coloca que a indexação por *tags*, utilizada para indexar páginas da Internet, é feita de maneira precária, pois as *tags* são adicionadas aos documentos pelo criador das páginas, e muitas vezes algumas nem chegam a ser indexadas.

Segundo Branski (2002), na Internet os *sites* são indexados por robôs, que percorrem a rede a partir de uma seleção de *sites* mais populares ou considerados melhores por seus administradores. Visitam os *sites*, lêem seu conteúdo, armazenando cada uma das palavras encontradas e seguem seus *links* para outros *sites*. Nestes *sites*, seguem os *links* para encontrar outros *links* e, sucessivamente adicionam os endereços aos bancos de dados dos mecanismos de busca. As informações são cadastradas sem qualquer classificação. Alguns mecanismos de busca indexam integralmente o conteúdo dos *sites*, outros somente o título e um resumo algoritmicamente construído, outros o título e as primeiras linhas do *site*.

Cendón (2000) relata que são poucos os recursos de técnicas de indexação de documentos na *web*, organizadas por profissionais com conhecimento. Apesar de alguns estudos estarem em andamento, a autora acredita que talvez os sistemas nunca se viabilizem de maneira global, devido a enorme quantidade de informações disponíveis na Internet.

De acordo com Guinchat e Menou (1994), a indexação pode ser realizada por uma única pessoa ou por várias. Estas pessoas são responsáveis pela extração de termos, pela formulação de termos em linguagem documental e pela verificação. Ela pode ainda ser realizada pelos responsáveis por cada um dos níveis (genérico, médio e profundo). Pode ser realizada no todo por uma única pessoa e depois revisada por outra mais experiente ou ainda por várias pessoas.

Esta autora ainda coloca que a indexação pode ser realizada automaticamente pelo computador, que manipula a linguagem documental, auxiliando na extração dos termos que representam o documento (a partir do título, resumo ou texto), ajudando a verificar se os descritores escolhidos existem e se estão escritos de maneira correta.

Em oposição ao processo de indexação humana, a indexação com o auxílio do computador ganha em rapidez, porém, exige cuidados maiores, como em casos de sinônimos e hierarquia de assuntos, onde, a partir do documento indexado, gera-se um índice geral que indica a quais documentos estão relacionados os termos extraídos da indexação (BUENO; VIDOTTI, 2000). Guinchat e Menou (1994), Robredo e Cunha (1994) apresentam as seguintes etapas da indexação:

- tomar conhecimento do documento;
- identificar o assunto principal do documento;
- identificar os elementos do conteúdo que devem ser descritos e extrair os termos correspondentes;
- verificar a pertinência dos termos escolhidos;
- traduzir os termos da linguagem natural para os termos correspondentes da linguagem documental, se for necessário;
- verificar a pertinência da descrição;
- formalizar a descrição, de acordo com a sintaxe da linguagem da indexação.

A indexação pode ser feita a partir do título, do resumo, do título e do resumo, do documento completo e a partir de determinadas partes do documento, tais como sumário, introdução, conclusão, tabelas, etc. A revisão pode ser feita comparando a indexação ao documento original ou aos termos extraídos originalmente, simulando algumas perguntas para ver se o documento pode ser recuperado, buscando informações sobre um dos assuntos de que ele trata, ou comparando, se for o caso, documentos semelhantes (GUINCHAT; MENO, 1994; ROBREDO; CUNHA, 1994).

Guinchat e Menou (1994) colocam que a qualidade da indexação pode ser julgada a partir de vários critérios:

- *exaustividade*: verificar se todos os temas, objetos e conceitos tratados pelo documento foram bem representados na indexação;
- *seletividade*: verificar se foram repetidas apenas as informações que interessam aos usuários;
- *especificidade*: verificar se a descrição traduz da forma mais precisa possível, o conteúdo do documento e se não utiliza descritores muito gerais ou muito específicos em relação às noções expressas no documento;
- *uniformidade ou consistência*: verificar se os indexadores descrevem um mesmo documento (ou documentos) sobre um mesmo assunto, da mesma forma.

Em relação a linguagem utilizada para representar os conceitos, Robredo e Cunha (1994) consideram a linguagem natural e a linguagem artificial.

A linguagem natural inclui:

- *linguagem natural livre*: indexação por palavras-chave extraídas do próprio documento;
- *linguagem natural controlada*: listas de descritores padronizados, com eliminação de sinônimos, listas de cabeçalhos de assuntos, tesouros, etc.

As linguagens natural e controlada serão tratadas mais a fundo no item 2.3.4 (página 49).

Por sua vez, a indexação em linguagens artificiais inclui:

- *linguagens artificiais não estruturadas*: esquemas de categorização codificados;
- *linguagens artificiais estruturadas*: esquemas de classificação hierárquica.

Robredo e Cunha (1994) argumentam que as linguagens de indexação são fundamentais para o processo de busca em bases de dados. As linguagens de indexação são consideradas por Rowley (2002) como sendo os termos (ou códigos) que podem ser utilizados como pontos de acesso num índice.

Portanto, para que uma recuperação tenha êxito, a linguagem de indexação e a linguagem de busca devem estar relacionadas.

As estratégias de busca são o meio de especificar as combinações dos termos para uma recuperação de informação bem sucedida (BERTHOLINO, 1999). As estratégias de busca serão tratadas no item recuperação de informações.

Diante do exposto sobre indexação, conclui-se que o estabelecimento de uma política de indexação é um fator extremamente importante no desenvolvimento de bases de dados. Com essa política, é possível definir o conteúdo do documento, determinando um ou mais termos para representá-lo. Os termos são então traduzidos para uma linguagem de indexação (listas de cabeçalhos de assuntos, tesouros, listas de autoridades, índices, etc.), para serem utilizados nas estratégias de busca e recuperação de informações. Assim, a indexação é um processo que requer uma análise cuidadosa do documento, a fim de identificar e selecionar os termos que melhor expressem o seu conteúdo, exigindo de quem realiza o processo, tempo e conhecimento adequados.

Nota-se ainda, que há uma grande carência de materiais relacionados a indexação de documentos na Internet. O que se percebe, é que as práticas utilizadas para indexar documentos impressos estão sendo aplicadas aos documentos eletrônicos. Porém, é importante considerar que estes documentos possuem características diferentes e diante da enorme quantidade de informações disponíveis, se faz necessário o desenvolvimento de técnicas que possibilitem indexar os documentos de maneira adequada, garantindo com isso, que os documentos sejam recuperados.

#### **2.3.4 Linguagem natural e linguagem controlada**

Lopes (2002, p. 42) conceitua a expressão linguagem natural como “sinônimo de discurso comum, isto é, a linguagem usada habitualmente na fala e na escrita sendo que, nas bases de dados, os termos do título e do resumo representam a linguagem natural.” Também é denominada vocabulário livre.

Lancaster (1993, p. 200) afirma que “a expressão normalmente se refere às palavras que ocorrem em textos impressos e, por isso, considera-se como seu sinônimo a expressão ‘texto livre’.” Nas bases de dados os campos de título e resumo registram os termos de linguagem

natural, enquanto os descritores, termos de indexação ou identificadores registram os termos da linguagem controlada.

No quadro 2, é possível verificar vantagens e desvantagens da linguagem natural.

<b>Vantagens</b>	<b>Desvantagens</b>
Permite o imediato registro da informação em uma base de dados, sem necessidade de consulta a uma linguagem de controle	Os usuários da informação, no processo de busca, precisam fazer um esforço intelectual maior para identificar os sinônimos, as grafias alternativas, os homônimos, etc.
Processo de busca é facilitado com a ausência de treinamentos específicos no uso de uma linguagem de controle	Haverá alta incidência de respostas negativas ou de relações incorretas entre os termos usados na busca (por ausência de padronização)
Termos de entrada de dados são extraídos diretamente dos documentos que vão constituir a base de dados	Custos de acesso tendem a aumentar com a entrada de termos de busca aleatórios
Temas específicos citados nos documentos podem ser encontrados	Uma estratégia de busca que arrole todos os principais conceitos e seus sinônimos deve ser elaborada para cada base de dados (ex: nomes comerciais de substâncias químicas não ocorrem no Chemical Abstracts)
Elimina os conflitos de comunicação entre os indexadores e os usuários, pois ambos terão acesso aos mesmos termos	Perda de confiança do usuário em uma possível resposta negativa

Quadro 2 - Vantagens e desvantagens da linguagem natural

Fonte: Lopes (2002, p. 48)

Por sua vez, a linguagem controlada, também denominada linguagem documentária ou vocabulário controlado, é definida por Lopes (2002, p. 42) como “um conjunto de termos organizados de forma hierarquizada e/ou alfabética, com o objetivo de possibilitar a recuperação de informações temáticas, reduzindo substancialmente a diversidade de terminologia.” Uma base de dados que utilize linguagem controlada possibilita a recuperação, no campo específico de descritor, apenas daquelas palavras-chave listadas no tesouro e/ou vocabulário controlado da base de dados. Na linguagem controlada das bases de dados, os termos que constituem o vocabulário são denominados palavras-chave, descritores, termos de indexação, códigos de classificação e outros.

Austin (1986 apud LOPES, 2002, p. 45) destaca que o uso de vocabulário controlado na indexação e o seu respectivo uso na recuperação de informações requerem o estabelecimento de certas regras terminológicas, tais como:

conceitos devem ser representados consistentemente para os propósitos de recuperação, por substantivos ou frases substantivadas; indexadores devem trabalhar com um vocabulário de termos preferidos, designando-se um dos sinônimos de um determinado conceito como o termo mais adequado para uso; a opção pelo singular ou plural dos conceitos e suas exceções devem ser registradas claramente nesse vocabulário, visando a consistência da indexação, sua fidedignidade e posterior uso na recuperação.

No quadro 3 são apresentadas vantagens e desvantagens da linguagem controlada.

Vantagens	Desvantagens
Controle total do vocabulário de indexação, minimizando os problemas de comunicação entre indexadores e usuários	Custos: a produção e manutenção da base de dados irá gerar despesas maiores com a equipe de indexadores. Será necessário ainda manter pessoas especializadas na atualização do tesouro
Com o uso de um tesouro e suas respectivas notas de escopo, os indexadores podem assinalar mais corretamente os conceitos dos documentos	O vocabulário controlado poderá não refletir adequadamente os objetivos do produtor da base, caso esteja desatualizado
Se bem constituído, o vocabulário controlado poderá oferecer alta recuperação e relevância e, também, ampliar a confiança do usuário diante de um possível resultado negativo	Um vocabulário controlado poderá se distanciar dos conceitos adequados para a representação das necessidades de informação dos usuários
As relações hierárquicas e as remissivas do vocabulário controlado auxiliam tanto o indexador, quanto o usuário na identificação de conceitos relacionados	Necessidade de treinamento no uso dos vocabulários controlados tanto para os intermediários, quanto para os usuários finais
Redução no tempo de consulta à base, pois a estratégia de busca será mais bem elaborada com o uso do tesouro	Desatualização do vocabulário controlado poderá conduzir a falsos resultados

Quadro 3 - Vantagens e desvantagens da linguagem controlada

Fonte: Lopes (2002, p. 47)

Lopes (2002) argumenta que muitos estudos sobre o uso da linguagem controlada e da linguagem natural na recuperação de informações têm se concentrado na utilização conjunta das duas linguagens nas estratégias de busca, comprovando que o uso simultâneo dessas linguagens proporciona melhor desempenho nos resultados.

Percebe-se então, a grande importância da utilização das linguagens natural e controlada no processo de indexação e recuperação de informações. É claro que tanto a linguagem natural quanto a linguagem controlada, apresentam vantagens e desvantagens. Assim, no momento de decidir pelo uso de uma dessas linguagens, devem ser levados em conta o tipo de usuário e as suas necessidades. E para obter bons resultados, uma alternativa interessante é tratar as duas linguagens em conjunto e não separadamente. Com isso, gera-se uma combinação ideal para ampliar os resultados da busca de informações.

### 2.3.5 Avaliação de bases de dados

Para avaliar bases de dados, visando sua seleção, aquisição e/ou implementação, Rowley (2002) apresenta os seguintes critérios:

- *cobertura temática*: assunto e tipo de material adequados, abrangência, duplicidade de outros serviços;
- *tipos de base de dados e tipo de informação que oferece*: podem ser referências, imagens, figuras, textos integrais, multimídia, acontecimentos, etc.;
- *atualidade*: período coberto pela base e/ou desde quando se acha disponível e frequência das atualizações;
- *facilidade de uso*: será influenciada por vários fatores, tais como *interface*, linguagem de indexação, estrutura da base e experiência do usuário;
- *tipo de saída*: se disponibiliza diversos formatos de impressão e permite importação de dados;
- *linguagem de indexação*: linguagem controlada ou natural, qual o grau de exaustividade da indexação;
- *custo*: custo inicial e adicional caso venha a ter, e relação custo benefício;
- *documentação e instrumentos auxiliares de busca*: manuais, *help*, informações sobre sistemas de classificação, tesouros, manuais de treinamento, etc.;
- *hospedeiro*: onde a base se acha disponível.

Por sua vez, Lancaster (1993) avalia as bases de dados em função de sua utilidade ao responder a várias necessidades de informações, de acordo com quatro critérios principais:

- *cobertura*: número de documentos sobre um determinado assunto, publicados num determinado período que estão incluídos na base de dados;
- *recuperabilidade*: número de documentos sobre um determinado assunto, incluídos na base de dados e que podem ser encontrados com a utilização de uma estratégia de busca razoável;
- *previsibilidade*: grau de eficiência com que o usuário pode comparar quais os itens serão e quais não serão úteis;
- *atualidade*: medida da velocidade com que novas publicações são incluídas num serviço de indexação/resumos.

Pereira et al. (1999) trata a questão da qualidade em bases de dados, enfatizando que o controle desta qualidade envolve todas as etapas do tratamento da informação, desde a sua criação até o uso final. A qualidade do produto da informação é influenciada pelo *software* de processamento e recuperação, pelo *hardware*, pelas telecomunicações, pelo suporte ao usuário e pela qualidade de conteúdo da base de dados.

Esta autora levantou, na literatura, duas vertentes principais para examinar a qualidade em bases de dados: a qualidade das estruturas de armazenamento e recuperação e a qualidade do conteúdo da base. Em relação a estruturas de armazenamento e recuperação, a participação dos usuários é fundamental na criação de projetos de *software* de alta qualidade. No tocante ao conteúdo da base, os mecanismos de coleta e registro são dimensões de suma importância, pois a funcionalidade destas *interfaces* será fundamental no processo de alimentação da base de dados. Uma base com erros de conteúdo, além de gerar prejuízo financeiro, pode recuperar informações incorretas. Outros fatores que merecem atenção são estratégias de crescimento, limpeza dos registros da base de dados e reorganização do conteúdo.

Complementando, Albrecht e Ohira (2000) argumentam que armazenagem eficiente de dados e fornecimento de mecanismos de recuperação são requisitos essenciais para uma boa base de dados.

Diante do exposto sobre os critérios de avaliação de uma base de dados, percebe-se que eles devem estar relacionados a todas as etapas que envolvem o tratamento das informações, ou seja, recuperação, armazenamento e disseminação. Outros aspectos que devem ser

considerados quando se avalia uma base de dados, dizem respeito ao tipo de documento, ao assunto e ao período de abrangência da base, além da frequência de atualização, da indexação, da *interface* de entrada e saída dos dados, dos custos, enfim, de uma série de fatores que vão garantir a escolha de uma base que possa atender as necessidades e os objetivos a que a mesma se propõe.

### 2.3.6 Bases de conhecimento

Apesar do termo base de conhecimento ter se originado na área de Inteligência Artificial e Sistemas Especialistas, ele é largamente utilizado para identificar uma coleção pesquisável de documentos e bancos de dados, onde estão reunidas informações sobre um domínio de conhecimento específico (BASES..., 2002).

Para Heinzle (1995, p. 12), base de conhecimento é “o local onde se armazenam fatos, heurísticas, crenças, etc., ou seja, é um depósito de conhecimentos acerca de um determinado assunto.”

Mendes (1997, p. 41) coloca que base de conhecimento “não é apenas uma coleção de informações. A tradicional base de dados, com dados, arquivos, registros e seus relacionamentos estáticos é substituída por uma base de regras e fatos e também heurísticas que correspondem ao conhecimento de um assunto específico.”

De acordo com Carvalho (2000, p. 41), base de conhecimento significa “tudo aquilo que a organização pode deter em termos de conhecimento explícito e conhecimento organizacional.” Nas bases de conhecimento, “encontram-se as bases de dados da organização e todos os arquivos operacionais, em todo e qualquer formato de mídia analógica, eletrônica.”

Para O’Dell e Grayson Júnior (2000) a principal função de uma base de conhecimento é compilar dados e informações. Ela pode colher informações de jornais, revistas, relatos de analistas, outros bancos de dados, relatórios de campo, Internet, apresentações e coisas do gênero, e armazenar os textos completos ou os resumos.

Uma base de conhecimento deve contemplar os seguintes recursos e facilidades:

- navegação simples e intuitiva;
- referências cruzadas através de *links* de hipertexto;
- pesquisa textual por palavras ou expressões;
- seleção do domínio da pesquisa;
- índice estruturados por assuntos;
- índices remissivos;
- indexação (categorização) dos documentos através de um vocabulário de assuntos ou palavras-chave;
- pesquisa em metadados ou em dados referenciais;
- FAQ – perguntas mais freqüentes com as respectivas respostas (BASES..., 2002)

É importante esclarecer que a presente dissertação não tem a pretensão de desenvolver uma base de conhecimento com a mesma conotação das bases de conhecimento que fazem parte da área de Inteligência Artificial. Por isso, não houve aprofundamento maior na revisão de literatura, e o que se buscou desta área, é a idéia de base de conhecimento como agrupamento de conhecimento específico.

Portanto, para a presente dissertação, base de conhecimento é concebida como um repositório de conhecimento, onde estará sistematizado e disponível, conhecimento nas áreas temáticas do Portal IGTL.

## **2.4 RECUPERAÇÃO DE INFORMAÇÕES**

A explosão de documentos disponíveis na Internet aumentou significativamente a dificuldade de recuperar informações úteis, com rapidez e eficiência. Há muita informação valiosa disponível, mas achá-la requer doses de paciência e conhecimento, além de tempo.

Isso era verdadeiro em 1994, na alvorada da *web* comercial, e ainda é realidade hoje em dia. Surgiram alguns catálogos virtuais, como Yahoo, HotBot e Google, mas eles cobrem apenas

uma fração dos recursos disponíveis na *web*. Muitos *softwares* desenhados para ajudar na busca de informações apareceram, porém, em muitos casos são lentos, imprecisos e mais difíceis de usar que os próprios *sites* de busca (TEIXEIRA FILHO, 1999).

Memorizar o conteúdo de um site é praticamente impossível, mas é necessário saber como encontrar esse conteúdo sem muita perda de tempo e de maneira satisfatória. Portanto, os mecanismos de busca e recuperação existentes são fundamentais para auxiliar no processo de recuperação de informações.

De acordo com Bueno e Vidotti (2000), a *interface* gráfica dos sistemas é um grande facilitador na recuperação automática de informações, pois nela estão embutidas as estratégias de busca em opções gráficas para o usuário, especialmente para o leigo, tornando a sua interação com o sistema automatizado mais agradável e menos complicada. As *interfaces* serão tratadas no item 2.4.3 (página 65).

Wurman (1989 apud DUARTE; SZOSTAK, 1999) cita algumas vantagens em relação a recuperação de informações:

- encontrar rapidamente informações úteis dentro de uma grande massa de dados e classificá-las por relevância, colaborando para o gerenciamento do conhecimento;
- oferecer valores imediatos às organizações, pois é possível pesquisar informações que já estão no formato eletrônico;
- os mecanismos, além de busca, podem disseminar as informações;
- os mecanismos de busca estão tirando vantagens da tecnologia *web*, oferecendo aplicativos baseados em *browsers*, permitindo a ampliação do espaço de pesquisa.

O processo de recuperação de informação consiste basicamente de três etapas: indexar, armazenar e recuperar (TEIXEIRA; SCHIEL, 1997).

E para que uma recuperação de informações apresente resultado satisfatório é importante conhecer os mecanismos e as estratégias de busca, que serão apresentados a seguir.

### 2.4.1 Mecanismos de busca

Na literatura analisada, é possível perceber que os mecanismos de busca também são denominados como ferramentas de busca, instrumentos de busca, sistemas de busca, serviços de busca, *search engines* ou máquinas de busca. Procurando homogeneizar o vocabulário apresentado, para esta dissertação adotou-se o termo mecanismo de busca.

De acordo com Cohen (1999), citado por Blattmann, Fachin e Rados (1999, p. 15) mecanismo de busca é “uma base de dados de arquivos da Internet coletados por um programa de computador (conhecidos como *wanderer*, *crawler*, *robot*, *worm*, *spider*).”

Bueno e Vidotti (2000, p. 7), definem os mecanismos de busca como

programas computacionais desenvolvidos com o objetivo de indexar informações descritivas e temáticas das páginas e/ou sites da Internet em bases de dados, com a finalidade de possibilitar a recuperação de documentos solicitados, pelos usuários da Internet, segundo as estratégias de busca e os critérios adotados.

Para Rowley (2002, p. 193) mecanismo de busca é “uma ferramenta de recuperação que realiza o trabalho básico de recuperação, a aceitação da consulta, o cotejo desta com cada um dos registros existentes na base de dados, bem como a apresentação, resultante disso, de um conjunto de itens recuperados.”

Pode-se afirmar então, que o principal objetivo de um mecanismo de busca é fornecer os resultados mais relevantes para uma determinada consulta no menor espaço de tempo possível.

Bueno e Vidotti (1999) salientam que os mecanismos de busca são apresentados através de uma *interface* gráfica amigável, onde as solicitações de busca são concretizadas em segundos e as respostas são apresentadas diretamente pelos *links* de acesso para as páginas e/ou por categorias de assuntos, ou ainda, segundo critérios de parametrização fornecidos pelos usuários, podendo apresentar ainda, roteiros de ajuda e exemplos de estratégias de busca diversificados.

De acordo com Araújo (1999/2000), Branski (2000), Bueno e Vidotti (1999), Cendón (2000), Cunha (2001), Teixeira e Schiel (1997) os mecanismos de busca podem ser classificados em:

- *diretórios (ou catálogos)*: são organizados a partir de *sites* que compõem sua base de dados em categorias, as quais podem conter subcategorias. Surgiram como a primeira tentativa de organizar e localizar informações na Internet. Os *sites* coletados passam pela seleção, na maioria das vezes, por seres humanos, ou ainda, pelo uso de robôs que coletam novas URLs (*Uniform Resource Locator*) e são compilados numa base de dados, contendo os endereços de recursos na Internet e suas descrições (ex. Yahoo, Cadê?, Surf, Lycos, Achei);
- *motores de busca (ou índices)*: não organizam hierarquicamente as páginas que buscam. Surgiram quando as informações disponíveis na Internet atingiram proporções que dificultavam sua coleta manual. Criam suas bases de dados automaticamente a partir de buscas realizadas por programas (denominados aranhas, robôs ou vermes) que realizam uma varredura procurando por páginas, que são analisadas para a extração de palavras que as descrevem. (ex. Alta Vista, Excite, Google, Radar Uol);
- *metamotores (ou metapesquisadores)*: buscam, ao mesmo tempo, em mais de um mecanismo (diretórios ou motores), trazendo resultados mais amplos do que a busca num único mecanismo. Não possuem nenhuma base de dados, utilizando exclusivamente dados de outros mecanismos de busca (ex. Webcrawler, Miner).

Branski (2002) coloca que já estão disponíveis também, mecanismos de busca especializados ou temáticos, que restringem-se a assuntos de um campo específico. A vantagem deste mecanismo sobre os demais, é que apresentam resultados mais relevantes num tempo de pesquisa menor. As bases de dados destes mecanismos são compiladas através de robôs ou por humanos.

Esta autora ainda coloca que uma mesma pesquisa em diferentes mecanismos trará resultados diferentes, pois os mecanismos não conseguem incluir em seus bancos de dados todos os *sites* existentes na Internet.

Logo, utilizar mais de um mecanismo de busca garante maior cobertura e conseqüentemente um resultado mais satisfatório.

As diferenças entre os mecanismos de busca podem estar relacionadas ao tamanho, especificidade, conteúdo de suas bases de dados, recursos disponíveis para a elaboração de estratégias de busca e resultado das buscas (BUENO; VIDOTTI, 2000; ROWLEY, 2002).

Avanços significativos em relação aos mecanismos de busca têm ocorrido ao longo dos últimos anos, disponibilizando aos usuários muitas opções de busca.

Cunha (2001), Duarte e Miranda (1999), Terra e Gordon (2002) destacam as seguintes opções de busca:

- *busca booleana*: permite a utilização dos operadores booleanos E (*AND*), NÃO (*NOT*), OU (*OR*) ou PRÓXIMO (*NEAR*) no sentido de incluir documentos ou arquivos que contenham determinadas palavras ou termos;
- *busca por palavra-chave e frase exata*: essas características estão disponíveis na maioria dos mecanismos de busca e requer que o resultado final contenha uma frase ou sentença exata ou específica, permitindo aos usuários direcionarem suas buscas para áreas específicas dos documentos (título, palavra-chave e autor). Atualmente tais mecanismos incluem URL e *links*;
- *busca por proximidade*: permite localizar documentos que contenham certas palavras perto de outras;
- *busca em texto integral*: permite procurar documentos do tipo texto a partir de palavras que se encontrem em qualquer posição dos mesmos;
- *busca difusa*: busca de grafias alternativas de palavras fazendo combinações mesmo quando as palavras estão grafadas de maneira errada;
- *busca por conceito*: busca de documentos que não contenham uma palavra específica, porém que esteja relacionada conceitualmente com essa palavra;
- *busca fonética*: permite que o texto informado para pesquisa seja procurado por seus fonemas e não pela maneira como é escrito;
- *busca com inferência bayesiana*: permite recuperar documentos baseados na frequência com que as palavras selecionadas aparecem neles;
- *busca por contexto*: permite buscar documentos em que as palavras podem possuir significados muito diferentes dependendo do contexto;

- *busca em linguagem natural*: possibilita ao usuário consultar utilizando inglês-corrente (ou qualquer outra língua). Entretanto, estão sendo desenvolvidas soluções mais sofisticadas que incluem agentes que simulam conversas humanas reais e inteligentes;
- *busca baseada em bases de conhecimento*: está estreitamente relacionada com os mecanismos de busca em linguagem natural, construindo automaticamente casos estruturados, permitindo aos usuários utilizar respostas prévias para questões semelhantes. No caso do mecanismo não encontrar a resposta apropriada ele poderá ser programado para direcionar o usuário para uma pessoa que tentará responder a questão;
- *busca baseada na popularidade*: muda automaticamente o *ranking* de *links* exibidos, baseando-se nos vários tipos de algoritmos que objetivam avaliar a perspectiva do próprio usuário sobre o documento ou *site*. Por exemplo, o mecanismo pode considerar o número de vezes que os usuários anteriores verificam o *link* ou a fonte de informação ou ainda quantos usuários referenciam aquele *site* específico (o google.com utiliza essa tecnologia);
- *busca com filtros colaborativos*: baseia-se na idéia de que os usuários que compartilham interesses semelhantes vão considerar documentos semelhantes como relevantes (as sugestões de livros da Amazon exemplificam esse tipo de mecanismo de busca);
- *busca por afinidade*: similar as buscas com filtros colaborativos. Porém, leva ainda mais longe a idéia de associar pessoas aos documentos. O mecanismo descobre muitos níveis de características comuns entre tipos diferentes de documentos, baseando-se na pessoa que o escreveu, leu ou o utilizou;
- *busca com mapeamento visual*: apresenta os resultados de duas buscas numa maneira mais gráfica, agrupando-os e associando-os de acordo com algoritmos de categorização que permitem aos usuários navegar facilmente entre blocos de informações relacionadas;
- *busca ponto-a-ponto (P2P)*: permite buscas dentro de uma comunidade auto-selecionada sem a necessidade de qualquer índice centralizado de documentos;
- *agentes personalizados*: possuem alta capacidade de aprendizado dinâmico baseado no padrão de buscas e escolhas de documentos de cada usuário. Os agentes avançados combinam conceitos de busca conceitual com busca por contexto para construir perfis relevantes dos usuários.

São muitas as opções de busca apresentadas pelos autores. E a escolha por uma delas, deverá levar em conta a finalidade da pesquisa e o conhecimentos dos mecanismos. Certamente, a utilização de mais de uma opção garantirá resultados mais relevantes.

Um fator de suma importância ao se utilizar um mecanismo de busca na Internet, é o conhecimento sobre os princípios dos operadores e das estratégias de busca de cada mecanismo (BLATTMANN; FACHIN; RADOS, 1999)

Os operadores são “elementos de uma estratégia de busca que relacionam os termos desta estratégia de modo a obter um todo que seja o reflexo da necessidade do usuário e que possa ser submetido ao computador.” (BARSOTTI, 1990, p. 80). Nos quadros 4 e 5, são apresentados operadores booleanos e sinais que podem ser utilizados para realizar as buscas.

<b>Operador</b>	<b>Função</b>
E ( <i>AND</i> )	relaciona dois ou mais termos, limitando a busca, apresentando resultados que contém todos os termos listados
OU ( <i>OR</i> )	relaciona dois termos e reúne todos os documentos que incluam pelo menos um deles ou ambos
NÃO ( <i>NOT</i> )	busca documentos que contém o termo de pesquisa que o precede, mas não o termo que o sucede

Quadro 4 – Operadores booleanos

Fonte: Barsotti (1990); Chagas et al. (2000); Nahuz (1999); Silva e Menezes (2001)

<b>Sinal</b>	<b>Função</b>
+ (mais)	o sinal de inclusão deve ser empregado antes de uma palavra ou frase para informar ao mecanismo de busca que ele deve selecionar os documentos que tenham obrigatoriamente todas as palavras precedidas do sinal +
- (menos)	o sinal de exclusão deve ser empregado antes de uma palavra ou frase para informar ao mecanismo de busca que ele não deve incluir os documentos que contenham aquela palavra ou frase
“ ” (aspas)	as aspas são empregadas para que o mecanismo de busca considere as palavras como sendo uma frase
* (asterisco)	É utilizado para solicitar ao mecanismo de busca que recupere todos os documentos que contenham a parte inicial da palavra (até o asterisco) com qualquer terminação
( ) (parênteses)	servem para elaborar pesquisas mais complexas, definindo operações menores dentro da expressão inteira. Nesse caso a busca funciona considerando os parênteses como se fossem termos isolados, e depois os combina

Quadro 5 – Sinais utilizados para realizar buscas

Fonte: Barsotti (1990); Chagas et al. (2000); Nahuz (1999); Silva e Menezes (2001)

O uso de sinais pode ser combinado, devendo ser utilizados de forma lógica e a primeira palavra deve ser sempre a de inclusão.

Ainda de acordo com estes autores destacam-se os operadores de proximidade, que também podem ser utilizados na realização de buscas:

- NEAR: recupera documentos contendo ambas as palavras especificadas ou frases contendo até dez palavras entre elas;
- ADJ: utilizado para termos que estão juntos. Este operador é muitas vezes reconhecido como a opção “frase exata” comuns nos mecanismos de busca na *web*.

Por fim, outro operador utilizado para recuperar informações é a truncagem, que consiste em suprimir uma parte da palavra, no início, meio ou fim, sendo o sistema capaz de recuperar todas as palavras que possuem a parte comum conservada na truncagem. Os símbolos utilizados na truncagem podem ser diferentes de uma base de dados para outra. São exemplos de símbolos de truncagem: ?, !, \*, #, \$ (BERTHOLINO, 1999; ROBREDO; CUNHA, 1994).

Geralmente, os mecanismos de busca trabalham com mais de um operador, combinando-os entre si, além de utilizarem também recursos adicionais como busca por linguagem natural, obrigatoriedade ou não da ocorrência dos termos, diferenciação entre maiúscula e minúscula, diferenciação de acentos e caracteres especiais, e outros recursos que necessitam de conhecimento prévio de conceitos e da sintaxe da busca, como busca por: data, domínio, URL, título, autor, palavras do resumo, idioma, tipo do documento, etc. (BUENO, VIDOTTI, 2000). Os autores ainda colocam que a utilização correta dos operadores e dos demais recursos vai interferir na qualidade das respostas obtidas com relação a sua pertinência e exaustividade.

Portanto, é importante conhecer e dominar os comandos e operadores de cada *site* ou base de dados, que são definidos de acordo com o *software* utilizado para sua elaboração e/ou com o sistema adotado pelo desenvolvedor do *site* e/ou produtor da base de dados (BERTHOLINO, 1999).

Lopes (2002) reforça a idéia colocando que a formulação exata de uma estratégia de busca é determinada pela natureza da base de dados a ser acessada e pela sua respectiva estrutura de

informação, ou seja, pela formatação de seus campos de identificação do documento e dos campos de identificação do conteúdo temático do mesmo. A identificação da estrutura de campos de busca implica no conhecimento da documentação básica fornecida pelos produtores das bases e pelos bancos de dados onde estão hospedadas.

Quanto ao nível de busca, em geral os mecanismos permitem mais de um nível, conhecidos por busca simples e busca avançada e podem ser utilizados conforme o conhecimento e experiência do usuário.

A busca simples pode ser feita na própria *homepage* e oferece a opção de uso de comandos mais gerais. A busca avançada é uma busca mais refinada, podendo ser feita apenas na *homepage* dos mecanismos de busca, abrindo uma janela especial, onde é possível utilizar os comandos mais específicos, para aproximar ao máximo o resultado da busca daquilo que se deseja encontrar (SILVA; MENEZES, 2001).

Bueno e Vidotti (2000) argumentam que uma mesma busca realizada nos níveis simples e avançado pode apresentar resultados diferentes, pois algumas opções podem estar disponíveis apenas em um nível de busca. A busca avançada oferece maiores opções e recursos para a elaboração de estratégias de busca.

#### **2.4.2 Estratégias de busca**

Para Lopes (2002, p. 41) a estratégia de busca pode ser definida como “uma técnica ou conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados.”

Barsotti (1990, p 71) afirma que estratégia de busca é “um conjunto de termos (descritores ou palavras da linguagem natural) relacionados entre si por meio de operadores.”

Segundo Rowley (1994, p. 129), a estratégia de busca é “o conjunto de decisões tomadas e de procedimentos adotados durante uma busca.” O objetivo das estratégias de busca é recuperar um número suficiente de registros relevantes e evitar que sejam recuperados registros irrelevantes e não pertinentes.

A definição da estratégia de busca é fundamental para refiná-la e conseguir obter resultados relevantes aos interesses do usuário. Na definição das palavras-chave é importante informar os termos sinônimos, correlacionados e equivalentes, além de suas respectivas definições no idioma inglês, adotado pela maioria das bases internacionais, e também consultar tesouros, vocabulários controlados, dicionários especializados e outras fontes (BERTHOLINO, 1999).

Grogan (1995) destaca cinco aspectos que servem de ponto de apoio para a estruturação da estratégia de busca: aspecto biográfico, bibliográfico, cronológico, de língua ou nacionalidade ou de forma. Acrescenta ainda que as características da questão para se traçar os passos a serem seguidos envolvem campo temático, tempo, espaço, língua, nível, quantidade e alguns outros fatores especiais, como competência e experiência do usuário, os quais contribuem nitidamente para as decisões a serem adotadas quanto a estratégia de busca.

Uma boa estratégia de busca assegura resultados mais relevantes. Logo, é importante conhecer as características dos mecanismos de busca, consultar o menu de ajuda, analisar os operadores e outros recursos disponíveis e observar os exemplos (BUENO; VIDOTTI, 1999).

Branski (2002) argumenta que o ideal para uma boa estratégia de busca é avaliar o que se conhece sobre o assunto e o que se quer saber sobre ele, definindo as melhores ferramentas e, a partir daí, realizar os seguintes passos:

- formular a questão da busca e sua abrangência;
- identificar os conceitos importantes dentro da questão;
- identificar os termos de busca que representem estes conceitos;
- considerar sinônimos ou variações destes termos;
- preparar a lógica de busca.

Considerando os aspectos abordados pelos autores e diante do grande volume de informações disponíveis na Internet, percebe-se que os mecanismos de busca, se utilizados de maneira adequada, permitem encontrar todo e qualquer tipo de informação.

No entanto, a escolha por um deles nem sempre é acertada. Logo, ao se elaborar uma estratégia de busca é preciso considerar o assunto desejado e as respostas esperadas, além de conhecer como os mecanismos indexam as informações e quais os recursos disponíveis para a

elaboração de estratégias de busca adequadas. Considerando a diversidade de informações disponíveis é interessante utilizar mais de um mecanismo para uma mesma pesquisa.

É importante também, acompanhar a evolução dos mecanismos de busca e das estratégias que os mesmos utilizam para recuperar informações, uma vez que cada tipo de mecanismo ou estratégia apresenta limitações, vantagens e desvantagens. Com isso, pode-se usufruir melhor das informações que a Internet oferece.

### **2.4.3 Interface**

Para Fachin (2002, p. 76) a *interface* é “a fronteira, o portão, a visão, ou seja, o acesso comum entre as tecnologias e os usuários.”

Benassi (1997, p. 70) afirma que a *interface* diz respeito “à maneira como o programa se apresenta ao usuário, à forma como suas funções são acessadas e à maneira como este programa fornece ao usuário o resultado de uma tarefa.”

Lewis e Rieman (1994 apud HEEMANN, 1997) e Fachin (2002) destacam que a *interface* é o ponto de acesso que permite a interação homem-máquina, através de elementos tais como menus, janelas, ícones, barras, teclado, *mouse* e *beeps*, que permitem a ligação entre os canais de comunicação e os usuários.

Por fim, Heemann (1997) conclui que é a *interface* que abre o sistema para o usuário, habilitando-o a extrair informações de dentro dele.

Assim, é o projeto da *interface*, mais do que qualquer outra característica, que habilita um sistema a ser de uso amigável.

A ISO 9241 é uma norma internacional que trata dos aspectos ergonômicos de *hardware* e *software* no uso de terminais de vídeo (BARBOSA; NUNES; SENA 2000). De acordo com esta norma, a interação entre o usuário e o sistema de recuperação da informação atingirá seu objetivo quando, ao projetar uma *interface*, forem utilizados os princípios ergonômicos gerais, que são:

- ser adequada para a tarefa para qual foi determinada;
- ser autodescritiva;
- ter controlabilidade;
- estar em conformidade/atender com as expectativas do usuário;
- ter tolerância de erro;
- atender às individualidades;
- propiciar o aprendizado.

Estes autores ainda colocam que para o usuário executar suas tarefas com eficácia, eficiência e satisfação, é importante que as características de clareza, discriminabilidade, concisão, consistência, detectabilidade, legibilidade e compreensibilidade sejam consideradas no projeto da *interface*.

A estas características, Bawden (1992 apud HEEMANN,1997) acrescenta funcionalidade, segurança, ajuda, explanação e mensagens de erro e padronização.

A *interface* deve apresentar atributos que facilitem o aprendizado, que pode ser medido num contexto específico, mediante comparação da usabilidade de um produto para um determinado usuário no decorrer de um tempo ou comparando a usabilidade do produto entre um usuário experiente e outro inexperiente (BARBOSA; NUNES; SENA, 2000).

Assim quando se projetam *interfaces*, é preciso considerar os seguintes elementos: *layout*, cores, fundos, textos, fontes e ícones. Estes elementos serão abordados a seguir.

Para Fachin (2002) e Cybis (1994), *layout* é o arranjo dos elementos, das informações, dos ícones e dos botões dentro de uma página *web*, de forma clara, objetiva e harmoniosa considerando o tamanho da tela do computador. As informações relevantes do *site*, ou página deverão estar dispostas no lado esquerdo da tela e acima, sendo que a mais importante deverá estar no canto superior esquerdo e as demais, logo abaixo, obedecendo a hierarquia. Assim, em se tratando de *layout*, deve-se estar atento a:

- organização hierárquica;
- sem excessos de imagens e figuras, ícones e botões;

- sem poluição sonora ou impactos fortes e piscantes;
- utilização em conjunto, de cores leves e/ou claras, a fim de identificar a ligação de assuntos que tenham correlação;
- informação objetiva, direta e clara, sem excesso de textos;
- ícones, quando acompanhados de texto, deverão estar dispostos no lado esquerdo e na vertical, e quando for composto de apenas uma ícone, de forma horizontal, na parte superior da tela;
- orientações, indicações, identificação de origem e data de atualização da página localizadas no rodapé das páginas *web*.

O *layout* tem um papel importante numa página *web*, influenciando na maneira como os usuários sentem e entendem as informações disponibilizadas, uma vez que a informação visual estabelece uma comunicação não-verbal, podendo incluir sinais emocionais, como motivação, direção, distração ou repugnância. Assim o modo como as informações são organizadas na página, pode tanto comunicar uma mensagem como deixar os usuários perplexos e/ou sobrecarregados (OLIVEIRA, 2001; PARIZZOTTO, 1997).

Em relação as cores utilizadas nas páginas *web*, muitos estudos têm sido realizados, e os resultados têm contribuído para a elaboração de diversas recomendações ergonômicas para os projetos de páginas na *web*.

O impacto das cores atinge diretamente o usuário, já que ele fica na frente da tela por pouco tempo ou por horas seguidas, podendo, sobre a influência delas, apresentar diversas reações, como tristeza, alegria, calma ou irritação. Destacam-se as seguintes recomendações em relação as cores:

- utilizar no máximo cinco e no mínimo duas cores;
- utilizar as cores de forma combinada, evitando deixar a página carregada, pesada;
- adaptar a resolução das cores ao tamanho da tela;
- combinar as cores de ícones, botões e figuras de forma suave, evitando impactos fortes;
- de preferência, utilizar o código das cores padrões, reconhecido regionalmente por todos;

- utilizar as mesmas cores para agrupar elementos relacionados;
- padronizar as cores, entre as páginas, a fim de fixar uma informação específica, ou formas, ou a direção e retorno;
- utilizar cores fortes para destacar um elemento, quando necessário;
- colorido supera o preto e branco, já que a memória grava mais fácil as cores, mas é preciso ter cuidado para não abusar no uso do colorido;
- permitir aos usuários, alterar as cores das páginas, de acordo com sua preferência, gravando um padrão que possa ser retornado (CYBIS, 1997; FACHIN, 2002; PARIZOTTO, 1997).

No tocante ao fundo de uma página *web*, Parizotto (1997) argumenta que o mesmo é composto por padrões de textura e/ou cores e que a escolha do fundo desempenha um papel importante no resultado final da página.

Cybis (1997), Fachin (2002) colocam que, normalmente, utiliza-se o branco, pois permite um trabalho de maior relevância e destaque, já que o fundo influencia na aparência, no equilíbrio e na localização dos demais itens que compõem a página, além de completar os espaços vazios, evitando que algum item da página fique flutuando. Os elementos utilizados como fundo de uma página *web* devem estar relacionados com o entendimento do conteúdo apresentado. Os autores ainda recomendam não utilizar elementos aleatórios, cuja finalidade seja a decorativa. O fundo de uma página não deve chamar mais atenção do que o conteúdo apresentado. No planejamento e definição do fundo, os usuários devem ser levados em conta, pois são eles que ficarão diante do site por algum tempo.

No que diz respeito a fonte de páginas *web*, a função principal está na organização das informações disponibilizadas, gerando maior capacidade de transmissão de dados, facilitando a interpretação por parte dos usuários. Elas definem o formato da apresentação textual, sendo que o formato da fonte pode ser alterado, dependendo da resolução do monitor (FACHIN, 2002).

Parizotto (1997) coloca que devido a natureza da tela do computador, as fontes geralmente são menos legíveis na tela do que quando impressas no papel. Logo, utilizar uma fonte como

um código auxiliador na compreensão dos elementos de um texto num *site* é de grande importância e deve ser realizada de modo consistente.

Sobre o texto de uma página *web*, Oliveira (2001, p. 60) coloca que ele é “o resultado do agrupamento de letras que formam as palavras, frases, parágrafos e páginas.”

Os textos devem ser curtos, objetivos e claros, podendo estar dispostos de forma simétrica ou assimétrica, por colunas e/ou pequenas colunas ou *menus* que se abrem quando solicitados. O texto deve sobrepor-se ao fundo de forma legível, além de estar de acordo com os demais recursos utilizados, procurando sempre o melhor contraste e legibilidade. A cor predominante é o preto (FACHIN, 2002).

É importante manter a mesma fonte e tamanho, do início ao final da informação. Os textos na *web* devem ser mais curtos e objetivos. Destacam-se ainda, as seguintes recomendações sobre os textos:

- utilizar a mesma fonte do início ao fim da página, alternando com outra para destacar título e informações relevantes;
- evitar o sublinhado, pois este recurso representa um elo a ser utilizado na Internet;
- evitar o hífen, pois interfere no carregamento da página e na velocidade de leitura da tela;
- evitar linhas longas, que requeiram o uso de barras de rolagem, preferindo utilizar colunas;
- utilizar os recursos de caixa alta, negrito e itálico com moderação, pois sobrecarregam a página e diminuem a velocidade de acesso;
- utilizar o vídeo-reverso (branco escrito em fundo preto) com muita cautela, evitando textos longos;
- evitar caracteres luminosos e piscantes, pois desviam a atenção do usuário (FACHIN, 2002; OLIVEIRA, 2001).

No tocante aos ícones, Fachin (2002) coloca que os mesmos são botões especialmente desenhados e/ou reutilizados que operam como guias funcionais e estéticos para a *interface*

homem-computador. São usados para representar tarefas, objetos e ações a serem executadas pelos usuários, facilitando a interação entre eles e o *site*.

Portanto, os ícones devem ser significativos, apropriados, coerentes, consistentes, claros, simples e definidos em pequenos grupos, não mais do que 20 por páginas e com tamanhos razoáveis, comparando-os sempre ao tamanho da tela (CYBIS, 1997).

Em relação a imagem, Molin et al. (2000 apud FACHIN, 2002) destaca que há várias formas de compreensão do conceito de imagem: verbais e não-verbais, visuais e auditivas, visuais-auditivas. A autora coloca que as imagens na *web* devem ser utilizadas com bastante cuidado, procurando centrar-se nas características dos seus usuários. Destacam-se alguns critérios para sua utilização:

- equipamentos (*hardware*): os vários tipos e modelos das máquinas;
- *software*: os vários tipos de recursos e formatos disponíveis para o uso na *web*;
- acessos: como o usuário acessa, as redes, a transferências de dados, os navegadores, a máquina do usuário;
- resoluções de cores;
- placa gráfica com resolução compatível com o monitor;
- resoluções dos vídeos mais atualizadas.

Por fim, Fachin (2002) enfatiza que na construção de *interfaces* é de vital importância que o projeto seja bem elaborado, considerando-se os elementos apresentados anteriormente, a identificação do público-alvo a quem as páginas se destinam e dosando-se os recursos tecnológicos utilizados (FACHIN, 2002).

Pode-se concluir então, que se o projeto de *interface* for bem elaborado, a interação com os usuários será satisfatória, permitindo recuperar informações relevantes sem maiores dificuldades e sem perda de tempo.

#### 2.4.4 Dificuldades do usuário na recuperação de informações

Antes de apresentar os problemas que os usuários enfrentam ao recuperar informações na Internet, é importante destacar que há vários tipos de usuários.

Chagas et al. (2000) e Rosenfeld e Morville (1999 apud BUENO; VIDOTTI, 2000) argumentam que usuários diversificados apresentam diferentes necessidades informacionais. Assim, devem ser levados em conta os seguintes fatores em relação as necessidades de informações destes usuários:

- alguns usuários têm claramente definido o tipo de informação que precisam, formulam suas perguntas com clareza e sabem onde a informação pode ser encontrada;
- alguns usuários sabem quais informações querem, mas não conseguem expressar adequadamente suas necessidades, além de não saberem exatamente onde as informações existem ou mesmo se existem;
- alguns usuários não sabem exatamente o que esperam encontrar, pois não sabem exatamente o que existe sobre o assunto;
- alguns usuários querem tudo sobre um assunto específico.

Mesmo utilizando os mais poderosos mecanismos de busca para recuperar informações, muitos usuários nem sempre obtêm um resultado satisfatório.

E vários são os fatores que contribuem para a frustração dos usuários: abrangência e profundidade da pesquisa, erros de digitação, péssimas descrições para limitar termos ou conceitos, tempo de resposta e custo, questões muito amplas ou específicas, falta do uso de sinônimos adequados. Além disso, é comum aparecerem dificuldades na utilização dos mecanismos, já que cada um trabalha de forma diferente. Isso acaba gerando dúvidas na escolha de um mecanismo adequado que permita ao usuário economizar tempo e dinamizar suas pesquisas (CHAGAS et al., 2000; CUNHA, 2001).

Cendón (2000) e Cuenca (1999) levantam outros problemas que o usuário enfrenta na recuperação da informação, que dizem respeito a limitações das *interfaces*, ao tempo de busca, aos campos disponíveis para a recuperação e a familiarização com o vocabulário de

indexação (livre ou controlado). Os autores colocam que, em muitos casos, as ferramentas de busca na *web* oferecem recursos menos sofisticados para pesquisa e recuperação que os sistemas computadorizados para acesso à base de dados. Por exemplo, em um sistema de bases de dados como o Dialog, existem opções como tesouros *on line* ou busca de palavras adjacentes que permitem um maior refinamento da pesquisa e que nem sempre estão disponíveis nas ferramentas de busca da *web*.

Reforçando os aspectos levantados por estes autores, Tomaél et al. (2001) acrescentam outras dificuldades em relação a apresentação da informação na Internet: uso de *frames*, uso de tecnologia inadequada devido a facilidade de aquisição, páginas muito longas, excesso de animações, URLs complexas, páginas soltas, falta de apoio para navegação, *links* sem padronização de cores, informação desatualizada e páginas que demoram muito tempo para carregar. Para os autores, estes problemas acabam dificultando o processo de recuperação de informações, deixando o usuário decepcionado com o resultado obtido.

Silva (2000) argumenta que a recuperação de informações na Internet ainda é um problema buscando solução. A autora coloca que a recuperação é mais complicada do que parece à primeira vista, pois envolve os aspectos de análise e indexação, que são processos que determinam se a informação que o usuário vai recuperar é relevante ou não.

Lynch (1997) citado por Silva (2000) argumenta que a análise e a indexação de informações nos mecanismos de busca da Internet são operações automáticas, realizadas por robôs, que utilizam a técnica de extração de palavras para representar o conteúdo do texto analisado. Esta técnica é diferente da técnica dos indexadores humanos, que é a da distribuição, onde o conteúdo é representado com termos significativos, mesmo que esses termos não apareçam no objeto analisado. Outra limitação dos indexadores automáticos é que reconhecem apenas texto, deixando de lado informações contidas em imagens, gráficos e vídeos. Alguns programas podem encontrar cores e padrões de imagem, mas não conseguem fazer deduções e relações de significado de uma imagem, que é uma atividade ainda de domínio somente dos humanos, pois depende de senso crítico.

Outro problema em relação a recuperação de informações, como relata Cendón (2000), é que nem sempre o usuário consegue saber se as informações disponíveis na Internet são recentes, pois é muito comum as páginas não serem atualizadas com uma frequência regular. Desta

forma, é importante verificar a data, pertinência e atualidade dos dados obtidos. A autora coloca que há ainda a questão do custo das informações disponíveis na Internet. Atualmente, muitas informações estão disponíveis sem custo, mas com a consolidação do caráter comercial da rede, alguns serviços já estão começando a exigir o pagamento de taxas e assinaturas para fornecer acesso às informações e a proporção dos serviços pagos tende a aumentar. Normalmente, esses serviços costumam oferecer alguns recursos gratuitos para os usuários e mais opções para quem tem assinaturas pagas. É comum também, o fornecimento de um período grátis para que o usuário experimente os serviços antes de decidir sobre a sua assinatura.

Ainda segundo a mesma autora, a confiabilidade da informação na Internet também representa um problema para o usuário. Muitas vezes, essa confiabilidade pode ser avaliada pela reputação da instituição que a fornece. Em alguns casos, um documento disponível na rede pode ser retirado a qualquer momento e, em outros, os documentos podem continuar disponíveis, porém alterados. Assim, na maioria das vezes, o usuário não pode comprovar facilmente a autenticidade das informações. Com a ausência de processos de seleção para a informação que é disponibilizada, a qualidade das fontes acaba variando de *site* para *site*.

De acordo com Cunha (2001), Teixeira e Schiel (1997), há centenas de mecanismos de busca na Internet e esse número tende a aumentar cada vez mais, oferecendo novas opções para realizar as buscas. Os autores consideram que os mecanismos de busca são, sem dúvida alguma, a melhor opção para efetuar pesquisas na Internet. Mas como nem sempre se é bem sucedido numa primeira busca, é interessante que o usuário reserve algum tempo para pesquisar, a fim de encontrar a informação que deseja, utilizando até mais de um mecanismo, levando em conta que os mesmos possuem critérios diferentes para organizar e apresentar as informações.

Assim, quando os usuários definem estratégias de busca, é importante que conheçam os mecanismos de busca que possam ser úteis e que permitam recuperar os documentos que mais se aplicam às suas necessidades.

Navegar na Internet em busca de informações consome tempo e requer paciência. Às vezes, os usuários se perdem pelo labirinto da rede, desviando-se do seu objetivo de pesquisa. Assim, para se extrair da Internet resultados efetivos, torna-se necessário uma capacitação

para a recuperação da informação e, para isso, tempo dedicado. O usuário deverá se tornar um pesquisador do ambiente da Internet, procurando descobrir o universo de informações que ela oferece, aliando a isso mecanismos customizados de busca, conforme as necessidades de informação definidas (SILVA, 2000).

Por fim, é importante salientar que usuários diferentes realizarão buscas de modos diferentes, de acordo com as características e funcionalidades do mecanismo de busca, o nível de experiência com mecanismos de busca, o conhecimento prévio do assunto específico e/ou assuntos correlatos, o tipo de inteligência predominante (intuitiva ou analítica, visual ou textual) e preferências por busca ou navegação (TERRA; GORDON, 2002).

Diante do exposto, percebe-se que uma série de fatores tem deixado o usuário frustrado na recuperação de informações na Internet, e o principal deles, é a sobrecarga de informações. Com isso, o usuário tem muita informação ao seu alcance, mas não tem condições de tratá-la, ou de encontrar o que realmente precisa.

Ao fator sobrecarga de informação, podem ser acrescentados outros fatores que podem deixar o usuário insatisfeito com o resultado de suas buscas, como confiabilidade, autenticidade, pertinência e atualização das informações, custos e tempo gasto na recuperação de informações. Além disso, muitas vezes o usuário não consegue utilizar de maneira satisfatória os mecanismos de busca, por não ter conhecimento suficiente sobre a sua funcionalidade.

Assim, é de suma importância que sejam desenvolvidas técnicas de recuperação de informação com o objetivo de recuperar com mais precisão e eficácia os conteúdos disponíveis na Internet, procurando satisfazer as necessidades informacionais da comunidade usuária.

## **CAPÍTULO 3 - MODELO PROPOSTO DE BASE DE CONHECIMENTO**

### **3.1 CONTEXTUALIZAÇÃO**

Considerando que a presente dissertação está inserida num projeto maior, que vem sendo desenvolvido pelo IGTI, se faz necessária uma breve descrição do IGTI e também do projeto do Portal. Em seguida, apresentam-se a estruturação, implementação, dinâmica de funcionamento e manutenção da base de conhecimento e a sua aplicação ao Portal IGTI. Por fim, apresentam-se os resultados esperados com o modelo proposto.

#### **3.1.1 Núcleo de Estudos em Inovação, Gestão e Tecnologia da Informação - IGTI**

O IGTI é um Núcleo de pesquisa voltado ao estudo da inovação e da tecnologia da informação, aplicadas à gestão de negócios. Foi criado em 1997, no Departamento e no Programa de Pós-Graduação em Engenharia de Produção, da Universidade Federal de Santa Catarina e é formado por uma equipe multidisciplinar de pesquisadores, mestrandos, doutorandos e bolsistas de iniciação científica.

O IGTI conta com dois sub-grupos:

- Inovação Tecnológica – INOVA - envolvendo P&D em redes empresariais, geração de idéias, difusão de tecnologia, transferência de tecnologia, competências para inovação e auditoria e planejamento tecnológico;
- Tecnologia da Informação – TI - envolvendo P&D nas áreas de comércio eletrônico, inteligência de negócios, gestão integrada de TI, estratégias de implantação e sistemas de gestão integrada.

Sendo um Núcleo de excelência, capacita e gera uma equipe de pesquisadores, mestrandos e doutorandos, com intuito de criar uma sinergia no planejamento e execução de projetos em

função das necessidades dos seus parceiros, identificando demandas de pesquisa nas áreas acadêmica e empresarial.

Pelo fato de dispor dessa equipe com formação e experiência multidisciplinar, as áreas de competência incluem:

- preparação de multiplicadores internos nas áreas temáticas do Núcleo;
- geração de idéias de negócios;
- comércio eletrônico;
- composição de portais corporativos e/ou de conhecimento (“*knowledge space*”);
- agentes inteligentes e recuperação de informação na Internet; inteligência competitiva;
- implantação de novas tecnologias;
- planejamento estratégico de informações;
- planejamento de sistemas de informações para executivos;
- modelagem de organizações baseada em suas competências essenciais;
- políticas e estratégias para gestão de ambientes inovadores;
- gestão do conhecimento e transferência de tecnologia em grupos de P&D.

A produção científica orientada à geração de ferramentas competitivas para as organizações é o foco de atuação do IGTI. A disponibilização de metodologias e modelos de soluções para busca de informações competitivas na *web*, metodologia para formação de redes empresariais de inovação para micro, pequenas e médias empresas, redes de informação e transferência de tecnologia, são alguns exemplos de trabalhos desenvolvidos pelo IGTI.

### **3.1.2 Projeto do Portal IGTI**

O projeto que vem sendo desenvolvido pelo IGTI é composto por uma equipe de alunos de mestrado e doutorado, tendo como coordenadora a professora Aline França de Abreu e como sub-coordenador o professor Pedro Felipe de Abreu. O projeto conta ainda com a participação de bolsistas de iniciação científica, que atuam nas atividades de pesquisa relacionadas ao mesmo.

O objetivo do projeto é desenvolver um Portal de Conhecimento (Portal IGTI) nas áreas de inovação e tecnologia da informação aplicada à gestão competitiva de negócios, para acesso a um ambiente virtual, onde comunidades de prática possam adquirir, criar, trocar, transferir, disponibilizar informação/conhecimento em forma de conteúdo disponível ou direcionamento a outras fontes de informação.

O desenvolvimento do Portal IGTI seguirá um modelo iterativo e incremental, partindo da atividade de pesquisa para a geração de conteúdo e definição estrutural do portal, evoluindo para o desenvolvimento prático de ferramentas da tecnologia da informação para suporte e implementação do mesmo e, posteriormente para uma etapa de transição dos resultados para o âmbito empresarial e acadêmico.

O desenvolvimento e implementação do Portal IGTI se dará no IGTI e contemplará os ambientes conhecimento, inteligência competitiva, redes de relacionamento e apoio à pesquisa, conforme mostra a figura 1.

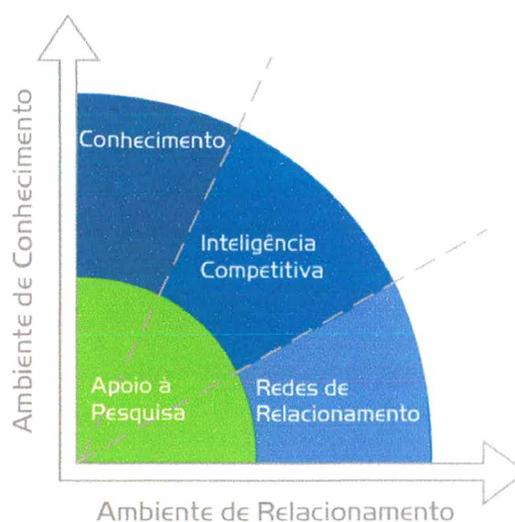


Figura 1- Portal IGTI  
Fonte: Adaptado de Abreu et al. (2002)

Os ambientes que compõe o Portal IGTI são descritos a seguir.

**Ambiente conhecimento** - compreenderá uma base de conhecimento, concebida como um repositório de conhecimento, composta por base bibliográfica, memória organizacional, *handbooks*, páginas temáticas com análise de tendências, mapas de competências e estado da arte na temática do Portal IGTI.

**Ambiente inteligência competitiva** - consistirá de ferramentas tecnológicas, centradas em agentes inteligentes, para busca e monitoramento de fontes de informações. Os agentes serão utilizados para o monitoramento de fontes conhecidas de informações, alertando para alguma referência às áreas temáticas do Portal IGTI. Também serão usados para buscar novas fontes, seguindo parâmetros e regras pré-estabelecidas e atuarão como uma primeira etapa no processo de aquisição, monitoramento e análise da informação. Os resultados das atividades dos agentes serão submetidos a filtros humanos, antes de serem considerados confiáveis. Com o tempo, espera-se poder adicionar mais funcionalidades e capacidades aos agentes, à medida que se obtém mais maturidade no seu desenvolvimento e uso.

**Ambiente redes de relacionamento** - possibilitará a interação entre os diversos membros desta e de outras comunidades, incorporará conceitos de *marketing* de relacionamento, redes de inovação, comunidades de prática. Permitirá, conseqüentemente, a criação e o compartilhamento do conhecimento, bem como a fidelidade das pessoas ao Portal IGTI.

**Ambiente apoio à pesquisa** - tem por finalidade disponibilizar um ambiente de trabalho virtual que promova a construção colaborativa de conhecimento. Este ambiente, integrado aos demais ambientes, permitirá alavancar a geração de novos conhecimentos contribuindo para a consolidação dos grupos de pesquisa existentes e para melhorar o processo de formação de novos pesquisadores, independente de tempo e de espaço.

Assim, esta dissertação está relacionada ao ambiente conhecimento e tem como objetivo estruturar uma base de conhecimento na temática do Portal IGTI, com ênfase na organização de conteúdo, visando seu armazenamento, sua recuperação e sua disponibilização na Internet.

Com o término do projeto do portal, serão gerados quatro trabalhos, ou seja, duas dissertações e duas teses. Assim, com a finalização desta dissertação, tem-se como resultado a parte teórica do ambiente conhecimento (desenho da base de conhecimento e sua funcionalidade). Os trabalhos relacionados aos demais ambientes estão em andamento.

Os quatro ambientes possuem forte relação entre si. O ambiente conhecimento influencia e é influenciado pelos outros ambientes, servindo como fonte de insumos e repositório do conhecimento por eles gerados. Essa dinâmica é que permitirá sua constante atualização.

### 3.2 PREMISSAS BÁSICAS

A revisão de literatura possibilitou identificar que nem sempre os usuários obtêm um resultado satisfatório quando buscam informações na Internet. Isso vem ao encontro da problemática levantada no início desta dissertação, que diz respeito às dificuldades de recuperação de informações e a necessidade de processos adequados de tratamento da informação, desde sua criação até seu uso final.

Entre os problemas que os usuários enfrentam quando buscam informações na Internet, destacam-se o grande volume de informações sem organização, a confiabilidade, a autenticidade, a pertinência e a atualidade das informações, os custos e o tempo gasto na recuperação de informações, as limitações das *interfaces* de busca, a falta de conhecimento sobre o vocabulário de indexação utilizado para realizar as buscas e, por fim, o desconhecimento sobre o funcionamento dos mecanismos de busca.

Buscando uma solução para os problemas levantados, propõe-se um modelo de base de conhecimento, visando o armazenamento, a recuperação e a disponibilização de conteúdo na Internet.

A base de conhecimento sistematizará e disponibilizará conteúdo de áreas específicas, oferecendo ao usuário uma noção sobre o universo a ser pesquisado, proporcionando um ponto de partida para que ele possa iniciar uma pesquisa. Consultando a base, o usuário poderá conhecer autores, conceitos, obras, eventos, especialistas, enfim, poderá acessar uma série de informações de maneira rápida e eficiente, a fim de satisfazer as suas necessidades informacionais.

A base de conhecimento contribuirá para solucionar os problemas relacionados a armazenamento, recuperação e disponibilização de informações relevantes na Internet, de forma eficiente e em tempo hábil. Com isso, o usuário poderá explorar melhor o grande potencial informacional da Internet.

### 3.3 MODELO PROPOSTO DE BASE DE CONHECIMENTO

Apresenta-se a seguir, um modelo de base de conhecimento, que é concebida como um repositório de conhecimento em áreas específicas. São apresentadas a estruturação, implementação e dinâmica de funcionamento e manutenção da base, bem como sua aplicação ao Portal IGTI.

A base de conhecimento será composta pelas seções base bibliográfica, memória organizacional, estado da arte, *handbooks*, páginas temáticas com análise de tendências e mapas de competências, onde estarão disponíveis os conteúdos, em forma de referência ou na íntegra, conforme mostra a figura 2.

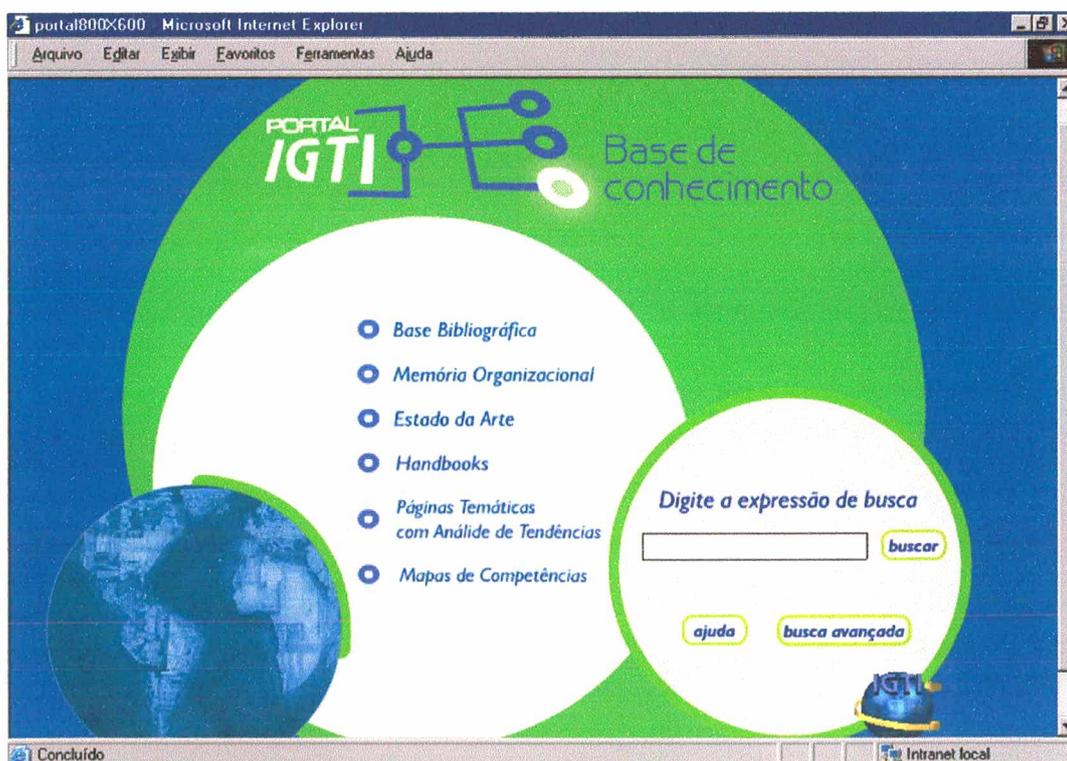


Figura 2 – Modelo proposto de base de conhecimento  
Fonte: Dados da pesquisa

A seguir, são descritas cada seção que compõe o modelo.

## **Base bibliográfica**

*O que é* - uma base composta por:

- informações bibliográficas sobre livros, teses, dissertações, projetos de tese, trabalho de conclusão de curso (TCC), trabalhos de disciplinas, artigos de revistas, de periódicos e de anais, vídeos, *cd-rooms*;
- textos integrais produzidos e publicados pelos integrantes do IGTI;
- *links* para *sites* relacionados às áreas de atuação do Portal IGTI.

*Objetivo* - armazenar, num só local, conteúdo de interesse do IGTI, em qualquer tipo de suporte e de qualquer fonte e permitir acesso ao mesmo via Internet.

*Resultado esperado* - sistematização de conteúdo nas áreas de abrangência do Portal IGTI, consolidação do conteúdo produzido pelos membros do IGTI e disponibilização de informações relevantes e de maneira adequada à comunidade usuária.

*Funcionamento* - a seleção dos documentos a serem indexados na base será feita por uma pessoa capacitada para tal. Os documentos selecionados serão cadastrados na base por bolsistas. A atualização se dará na medida em que forem surgindo documentos considerados relevantes e de interesse para o IGTI.

É importante salientar que os textos integrais produzidos pelos integrantes do IGTI, somente serão disponibilizados com a autorização dos autores.

## **Memória organizacional**

*O que é* - todo e qualquer tipo de documento que registre o conhecimento considerado relevante para a memória de uma organização.

*Objetivo* - reunir, num mesmo local, todo e qualquer tipo de documento que tenha importância e contribua para a memória organizacional do IGTI. Documentos tais como relatórios, atas, memorandos, estatuto, palestras, cursos e projetos desenvolvidos pelos integrantes do IGTI.

*Resultado esperado* - contribuir para a organização, preservação e disponibilização do acervo de conhecimento do IGTI.

*Funcionamento* - os documentos produzidos pelo IGTI serão organizados e disponibilizados nesta seção, na medida em que forem sendo gerados. Esta atividade será realizada por bolsistas, orientados por uma pessoa capacitada para tal.

Cabe destacar que esta seção terá o acesso restrito ao IGTI.

### ***Handbooks***

*O que é* – espécie de manual que consolida literatura sobre um determinado tema, dividido em sub-seções específicas, organizadas com a participação e colaboração de especialistas nos temas envolvidos.

*Objetivo* – consolidar literatura nas áreas temáticas do Portal IGTI.

*Resultado esperado* – consolidação e atualização de temas específicos que possam servir aos interesses de profissionais e estudantes, como ponto de partida para um estudo ou para atualização de conhecimento sobre uma área específica.

*Funcionamento* – os textos serão extraídos das revisões de literatura das dissertações e teses. O *handbook* será composto por sub-seções específicas das áreas temáticas do Portal IGTI. No primeiro *handbook* têm-se a consolidação dos temas envolvidos e nos próximos, faz-se apenas a atualização, que poderá ser anual ou bienal.

### **Estado da arte**

*O que é* - relatório sobre o estado do assunto em um domínio do conhecimento, feito para avaliar a literatura em um período determinado.

*Objetivo* - oferecer ao usuário conceitos básicos, principais autores, eventos, obras e periódicos nas áreas temáticas do Portal IGTI.

*Resultados esperados* - contribuir para o processo de pesquisa, tanto para os integrantes do IGTI, quanto para outros usuários que tenham interesse pelas áreas envolvidas.

*Funcionamento* – os principais conceitos, autores e obras serão extraídos das revisões de literatura das dissertações e teses. Os principais eventos e periódicos serão monitorados pela Internet.

### **Páginas temáticas com análise de tendências**

*O que é* – páginas compostas de *links* de *sites* sobre temas específicos.

*Objetivo* – reunir num só local, informações sobre quais direções as áreas temáticas do Portal IGTI estão seguindo.

*Resultado esperado* – facilitar a localização e o acesso a informações de ponta nas áreas temáticas do Portal IGTI.

*Funcionamento* – além dos *links* para os *sites*, haverá notas explicativas sobre o que o usuário vai encontrar, podendo avaliar se o *link* é do seu interesse ou não.

### **Mapas de competências**

*O que é* - espécie de guia que serve para identificar e localizar pessoas, organizações, etc., com conhecimento relevante em áreas específicas.

*Objetivo* - permitir ao usuário identificar de maneira rápida, as principais competências de pessoas (integrantes do IGTI ou não), empresas e outros grupos atuantes nas áreas de abrangência do Portal IGTI.

*Resultado esperado* – agilizar o contato com pessoas, organizações ou outros grupos que tenham interesse pelas áreas de abrangência do Portal IGTI e facilitar o processo de compartilhamento de conhecimento.

*Funcionamento* - o guia listará as pessoas (do IGTI ou não), os grupos e as organizações que atuam na área de abrangência do Portal IGTI, fornecendo uma breve descrição dos mesmos, o telefone, o *e-mail* ou o endereço do *site*.

Todas estas seções serão acessíveis através de *links* na página principal da base de conhecimento. Passando o *mouse* sobre a seção, aparece o significado da mesma, conforme figura 3. Com isso, o usuário não precisa ficar clicando para entrar na seção e verificar qual é o seu conteúdo, tendo condições de decidir se aquela seção lhe interessa ou não.



Figura 3 – Interface de descrição e acesso ao conteúdo das seções  
Fonte: Dados da pesquisa

Na figura 4 é possível visualizar a *interface* da seção estado da arte. As demais seções seguirão o mesmo padrão.



Figura 4 – *Interface* da seção estado da arte  
Fonte: Dados da pesquisa

A base de conhecimento terá como público alvo todos os integrantes do IGTI: graduandos, mestrandos e doutorandos, presenciais ou à distância, além de outros grupos de pesquisa, instituições e/ou empresas que atuem na área compreendida pelo IGTI.

### 3.3.1 Recuperação de informações na base de conhecimento

A busca na base de conhecimento poderá ser realizada de duas formas: busca simples e busca avançada.

A busca simples estará disponível na página principal da base de conhecimento (figura 2), onde deverá ser digitada a expressão de busca e em seguida clicado o botão buscar (ou tecla enter). Este tipo de busca trará como resultado documentos de todos os tipos e de todas as seções disponíveis na base de conhecimento.

Para realizar uma busca mais refinada, deve-se escolher a opção busca avançada, que levará a uma nova janela, conforme mostra a figura 5.

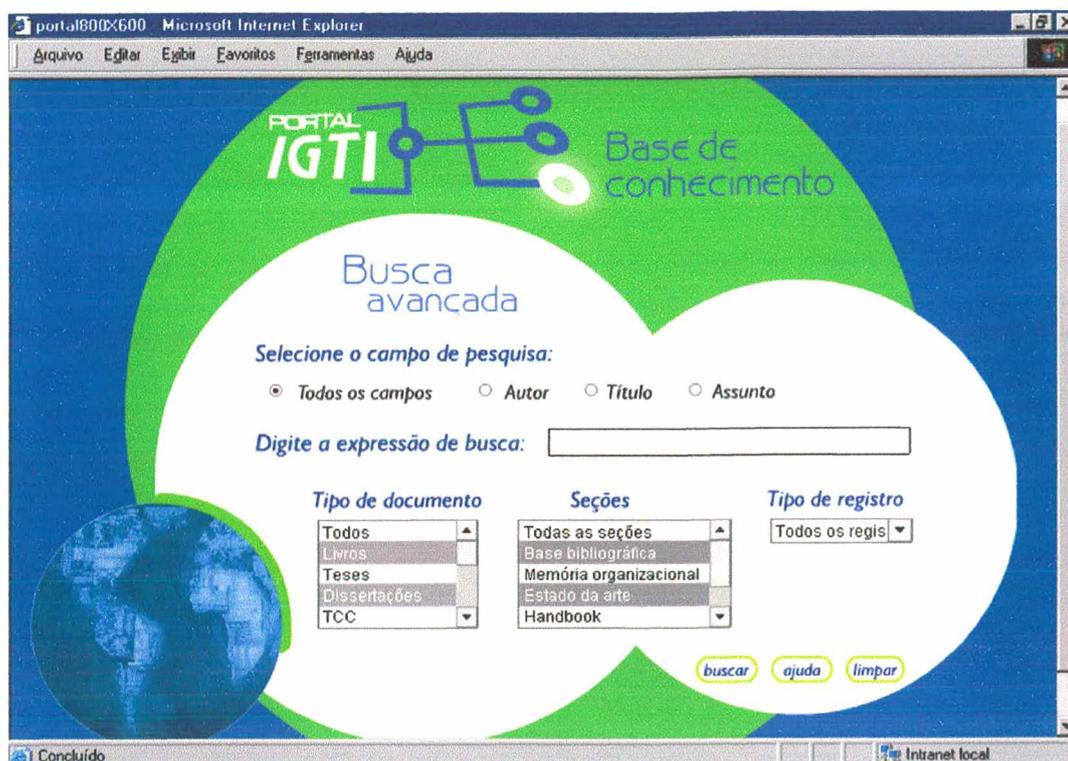


Figura 5 – Busca avançada na base de conhecimento  
Fonte: Dados da pesquisa

Pode-se pesquisar em todos os campos do registro, ou apenas em autor, título ou assunto.

Em relação ao tipo de documento, pode-se optar por todos, apenas um, ou mais de um, selecionando o tipo desejado. O mesmo se aplica às seções que compõem a base.

Quanto ao tipo de registro, pode-se optar por todos os registros, registros sem resumo, registros com resumo ou registros com texto completo.

### 3.3.1.1 Tipos de busca

Na opção ajuda, abrirá uma nova janela onde estão disponíveis os tipos de busca que podem ser realizadas, conforme mostra a figura 6.



Figura 6- Tipos de busca  
Fonte: Dados da pesquisa

Automaticamente a busca é do tipo “E” lógico, para todos os termos da busca. Ex.: uma busca com a expressão “engenharia mecânica “E” acústica, trará como resultado documentos que contenham obrigatoriamente todas as palavras ligadas por esse operador.

O uso do operador “NÃO” busca documentos que contenham a expressão que o precede, mas não a expressão que o sucede. Ex.: uma busca com a expressão engenharia mecânica “E NÃO” acústica, trará como resultado documentos que contenham a expressão engenharia mecânica, mas que não contenham a expressão acústica.

O asterisco ( \* ) pode ser utilizado para solicitar uma busca que recupere todos os documentos que contenham a parte inicial da palavra (até o asterisco), com qualquer terminação. Ex.: uma busca com a expressão produ\*, trará como resultado documentos que contenham as expressões produção, produtivo, produtividade, produto.

As aspas ( “ ” ) podem ser utilizadas para buscar expressões ou frases exatas. Ex.: uma busca contendo a expressão “ciência da informação” trará como resultado documentos que contenham exatamente essa expressão.

As buscas não diferenciam maiúsculas e minúsculas. Ex.: uma busca com a expressão Informação ou informação, trará como resultado documentos que contenham as duas expressões.

As buscas ignoram a acentuação das palavras, assim como as cedilhas. Ex.: numa busca a expressão informação, informacao e informação são consideradas a mesma palavra.

### 3.3.1.2 Resultado da busca na base de conhecimento

Os registros recuperados serão apresentados no formato de referência bibliográfica (de acordo com a norma da ABNT, NBR 6023), conforme mostra a figura 7.



Figura 7 – Resultado da busca na base de conhecimento

Fonte: Dados da pesquisa

Na resposta da busca aparecerá o termo pesquisado, o número de registros recuperados e as opções: primeira página, próxima página, página anterior, última página e nova busca. Com isso, tem-se uma noção sobre a quantidade de documentos recuperados, sendo possível analisar se o resultado da busca foi satisfatório ou não. Ao lado dos documentos recuperados

aparece o *link* detalhe, que levará para outra página, onde aparecerá o detalhamento do registro, conforme mostra a figura 8.



Figura 8 – Detalhamento do registro  
Fonte: Dados da pesquisa

No detalhamento do registro, aparecerão todos os campos existentes no mesmo, sendo possível retornar à página anterior e escolher outro registro recuperado ou realizar nova busca. No caso do registro recuperado ser um documento na íntegra ou um *site*, haverá um *link* que dará acesso ao mesmo.

### 3.4 RESULTADOS ESPERADOS

Espera-se, como resultado, que o modelo proposto possibilite sistematizar o conhecimento nas áreas temáticas do Portal IGTI, contribuindo para o desenvolvimento do portal, para a gestão do conhecimento do IGTI e para a comunidade em geral, oferecendo informações relevantes, de forma eficiente e em tempo hábil, dando ao usuário uma noção de quais informações estão disponíveis, como estão organizadas e como podem ser recuperadas.

Desse modo, espera-se que o modelo proposto permita o aumento do escopo de acesso à informação nas áreas de abrangência do Portal IGTI, agilize o acesso à informações úteis, confiáveis e relevantes, disponibilize grande quantidade de conteúdo digital, gere conhecimento de ponta nas áreas temáticas do Portal, enfim, traga para o IGTI, uma repercussão positiva tanto na área acadêmica quanto na área de negócios.

Cabe ressaltar que, embora o modelo seja proposto ao Portal IGTI, trata-se de um modelo genérico, podendo ser aplicado em qualquer organização, unidade administrativa ou grupo de pesquisa que tenham na informação tecnológica sua base de competitividade, permitindo organizar conteúdo em áreas específicas e facilitando o acesso ao mesmo.

## **CAPÍTULO 4 - CONSIDERAÇÕES FINAIS**

Neste capítulo, são apresentadas as conclusões, as contribuições do trabalho e as recomendações para trabalhos futuros.

### **4.1 CONCLUSÕES**

A problemática levantada no início desta dissertação, de sobrecarga de informações e da necessidade de processos adequados de recuperação, armazenamento e disponibilização de conteúdo na Internet, tem profunda relação com os temas desenvolvidos na revisão de literatura.

A revisão de literatura possibilitou chegar a várias conclusões, apresentadas a seguir.

A informação é um fator indispensável para impulsionar o desenvolvimento da sociedade, constituindo-se em um insumo essencial para a geração de conhecimento, possibilitando com isso, a satisfação das diversas demandas informacionais dos usuários.

Os portais, as bases de dados e os mecanismos de busca se constituem em ferramentas eficientes para o tratamento das informações, desde sua criação até seu uso final. Contudo, é necessário que estas ferramentas sejam utilizadas de maneira adequada, levando o usuário a identificar quais informações estão disponíveis, a forma como elas estão organizadas e como podem ser recuperadas.

Atualmente, estas ferramentas estão sendo utilizadas também para iniciativas de gestão do conhecimento, onde o foco principal é a captura, a organização, a utilização e o compartilhamento do conhecimento, tanto interno quanto externo, visando atingir plenamente os objetivos da organização.

Ainda em relação a gestão do conhecimento, percebe-se que o conhecimento é um recurso essencial para as organizações obterem vantagem competitiva. Mas para que isso aconteça, é necessário que o conhecimento esteja associado a uma ação, sendo que a sua gestão só fará sentido, se estiver voltada para a melhoria do desempenho da organização.

Quanto aos portais, os mesmos podem ser considerados como uma porta de entrada para o conteúdo *on line*, onde os usuários podem encontrar informações sobre inúmeros assuntos, além de uma variada opção de produtos e serviços que atendam as suas necessidades.

No tocante a base de dados, conclui-se que é o local onde se armazenam dados, inter relacionados entre si e gerenciados de forma a permitir sua posterior recuperação, com acesso via terminais locais ou remotos.

No que diz respeito a recuperação de informações, foi possível notar que muitos fatores têm deixado o usuário decepcionado quando busca informações na Internet. Entre estes fatores, destacam-se: sobrecarga de informações, confiabilidade, autenticidade, pertinência, atualização, custos, *interfaces* mal elaboradas, falta de conhecimento sobre o funcionamento dos mecanismos de busca e tempo gasto para recuperar as informações.

Deste modo, para recuperar informações de maneira adequada, algumas medidas se fazem necessárias, dentre elas, acompanhar a evolução dos mecanismos de busca, conhecer como eles indexam as informações e quais as estratégias de busca são mais adequadas para realizar as buscas. Com isso, pode-se usufruir melhor das informações que a Internet oferece.

A revisão de literatura permitiu observar ainda, que há uma grande carência de materiais relacionados a indexação de documentos na Internet. O que se percebe, é que as práticas utilizadas para indexar documentos impressos estão sendo transferidas para os documentos disponíveis na Internet. Porém, é importante considerar que estes documentos possuem características diferentes dos documentos impressos. E diante da enorme quantidade de informações disponíveis, se faz necessário o desenvolvimento de políticas e métodos padronizados, para revisar, selecionar e indexar os documentos disponíveis de maneira adequada, garantindo com isso, que os documentos sejam recuperados.

Portanto, foi possível constatar o grande potencial da Internet e da tecnologia da informação na proposição de estruturas e serviços para o armazenamento, a recuperação e a disponibilização de informações, visando atender às necessidades dos usuários.

Além de consolidar o conhecimento sobre os temas mencionados anteriormente e desenvolvidos na dissertação, a revisão de literatura permitiu levantar os subsídios necessários para a estruturação do modelo de uma base de conhecimento, concebida como um repositório de conhecimento específico.

O modelo foi proposto ao Portal IGTI, com a finalidade de sistematizar e disponibilizar conteúdo, tanto do IGTI quanto de outras fontes que atuem nas áreas de interesse do mesmo.

Com isso, considera-se que o objetivo principal da dissertação, de estruturar uma base de conhecimento nas áreas de inovação e tecnologia da informação, com ênfase na organização de conteúdo, visando seu armazenamento, sua recuperação e sua disponibilização na Internet, foi alcançado, contribuindo para solucionar os problemas relacionados a armazenamento, recuperação e disponibilização de informações relevantes na Internet, de forma eficiente e em tempo hábil.

O modelo proposto ao Portal IGTI, compreende as seções base bibliográfica, estado da arte, *handbooks*, memória organizacional, páginas temáticas com análise de tendências e mapas de competências, onde estarão disponíveis os conteúdos, em forma de referência ou na íntegra. O conteúdo poderá ser recuperado pela busca (simples ou avançada), disponível na base de conhecimento. Com isso, considera-se que os objetivos específicos também foram alcançados.

Cabe ressaltar que, embora o modelo tenha sido proposto ao Portal IGTI, trata-se de um modelo genérico, aplicável a qualquer organização, unidade administrativa ou grupo de pesquisa que tenham na informação tecnológica sua base de competitividade.

É importante destacar ainda, que a base de conhecimento não foi implementada. Portanto, quando da sua implementação, é possível que alguns ajustes se façam necessários, visando seu melhor desempenho e garantindo um melhor atendimento às necessidades informacionais dos usuários.

## 4.2 CONTRIBUIÇÕES

As contribuições desta dissertação são tanto para o IGTI, como para os usuários em geral.

Para o IGTI, no sentido de colaborar para o desenvolvimento do Portal IGTI, onde a base de conhecimento estará inserida. Com a implementação do portal, a base servirá como repositório de informações/conhecimento oriundos dos outros ambientes (inteligência competitiva, redes de relacionamento e apoio à pesquisa), além de interagir com eles oferecendo suporte de informação.

A presente dissertação pretende contribuir também para a gestão do conhecimento do IGTI, já que as seções que fazem parte da base de conhecimento conterão conhecimento gerado pelos integrantes do mesmo. Assim, através das seções da base de conhecimento, este conhecimento será recuperado, armazenado e disponibilizado aos integrantes do IGTI, presenciais ou à distância, e contribuirá para a preservação do acervo de conhecimento do IGTI.

Para os usuários que têm interesse por informações nas áreas temáticas do IGTI, a organização das seções na base de conhecimento dará uma noção sobre o universo que poderá ser pesquisado, identificando o conteúdo existente, a forma como está disponível e como poderá ser recuperado. A base de conhecimento servirá como ponto de partida para que o usuário possa iniciar uma pesquisa, permitindo que ele conheça autores, conceitos, obras, eventos, especialistas, enfim, uma série de informações que o levarão a explorar a base, buscando sempre melhores resultados para satisfazer as suas necessidades.

Por fim, acredita-se que esta dissertação trará ao IGTI uma repercussão positiva nas áreas acadêmica e de negócios, permitindo o aumento do escopo de acesso à informação na temática do Portal IGTI, agilizando o acesso à informações úteis, confiáveis e relevantes, disponibilizando grande quantidade de conteúdo digital, em forma de referência ou texto integral, bem como gerando conhecimento de ponta numa ação colaborativa e de compartilhamento com parceiros e disponível à sociedade via Internet.

### **4.3 LIMITAÇÕES**

Esta dissertação apresenta como limitação o fato da base de conhecimento não ter sido implementada, com a conseqüente avaliação da satisfação dos usuários e teste do desempenho dela. Com isso, é possível que, para otimizar o desempenho da base de conhecimento, alguns ajustes se façam necessários, procurando-se manter a funcionalidade da mesma, apresentada no modelo proposto.

### **4.4 RECOMENDAÇÕES**

Com base nas conclusões obtidas com a revisão de literatura, pode-se considerar que em relação a bases de conhecimento, portais e recuperação da informação, existem abordagens relativas a Ciência da Informação e a Ciência da Computação, e que ambas possibilitam inúmeras formas de organização e recuperação de conteúdo.

Considerando os avanços tecnológicos, trabalhos futuros deverão dar atenção especial ao processo de indexação de documentos disponíveis na Internet, pois ele é fundamental para a organização de informação, determinando a forma como o conteúdo será armazenado, para que possa ser recuperado de forma a atender às expectativas do usuário, considerando relevância, pertinência e tempo de busca.

Outra sugestão de trabalho futuro, é a implementação da base de conhecimento para avaliar a satisfação dos usuários na recuperação de documentos e testar o desempenho da mesma.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, Aline F. et al. Gestão do conhecimento em grupos de pesquisa: a experiência de composição de uma base de referência na Internet. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 21., 2001, Salvador. **Anais...** Porto Alegre: ABEPRO, 2001.

ABREU, A. F. et al. Portal IGTI: ambiente virtual para comunidades de prática nas áreas de inovação e tecnologia da informação aplicadas à gestão de negócios: uma iniciativa que atende a perspectiva da sociedade da informação no Brasil. In: WORKSHOP BRASILEIRO DE INTELIGÊNCIA COMPETITIVA E GESTÃO DO CONHECIMENTO, 3.; CONGRESSO ANUAL DA SOCIEDADE BRASILEIRA DE GESTÃO DO CONHECIMENTO, 2002, São Paulo. **Anais...** São Paulo: [s.n], 2002.

ALBRECHT, Rogéria F.; OHIRA, Maria Lourdes B. Bases de dados: metodologia para seleção e coleta de documentos. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 5, n. 5, p. 131-144, 2000.

ANGULO, Marcelo J.; ALBERTIN, Alberto Luiz. Portais ou labirintos? In: ENCONTRO ANUAL DA ANPAD, 24., Florianópolis, 2000. **Anais...** Florianópolis: [s.n.], 2000.

ARAÚJO, Wagner J. Ferramentas para promoção em *web sites* de unidades de informação. **Revista de Biblioteconomia de Brasília**, Brasília, v. 23/24, n. 1, p. 89-107, 1999/2000. (especial)

ARCE, Maria V. S.; PÉREZ, Tomás S. Las comunidades virtuales y los portales como escenarios de gestión documental y difusión de información. **Anales de Documentación**, n. 4, 2001. Disponível em: <[www.um.es/fccd/anales](http://www.um.es/fccd/anales)>. Acesso em 19 ago. 2002.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 10520**: informação e documentação: citações em documentos: apresentação. Rio de Janeiro, 2002.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14724**: informação e documentação: trabalhos acadêmicos: apresentação. Rio de Janeiro, 2002.

ATAÍDE, Maria E. M. O lado perverso da globalização na sociedade da informação. **Ciência da Informação**, Brasília, v. 26, n. 3, p. 268-270, set./dez. 1997.

BARBOSA, Elza, M. F.; NUNES, Eny M. A.; SENA, Nathália K. *Web sites* governamentais, uma esplanada à parte. **Ciência da Informação**, Brasília, v. 29, n. 1, p. 118-125, jan./abr. 2000.

BARROSO, Antônio C. O.; GOMES, Elizabeth B. P. Tentando entender a gestão do conhecimento. **Revista de Administração Pública**, v. 32, n. 2, p. 147-170, mar./abr. 1999.

BARSOTTI, Roberto. **A informática na Biblioteconomia e na documentação**. São Paulo: Polis; APB, 1990. (Palavra-chave, 2).

BASES de conhecimento. Disponível em: <[www.ecj.com.br/basesdeconhecimento.htm](http://www.ecj.com.br/basesdeconhecimento.htm)>. Acesso em: 19 ago. 2002.

BENASSI, Marcos T. O acesso à base Eric via Internet: o usuário diante de duas *interfaces* distintas. **Transinformação**, v. 9, n. 2, p. 69-92, maio/ago. 1997.

BERTHOLINO, Maria L. F. Buscas em bases de dados. In: RAMOS, Maria Etelvina M. (Org.). **Tecnologia e novas formas de gestão em bibliotecas universitárias**. Ponta Grossa: UEPG, 1999. p. 145-155.

BLATTMANN, Ursula; FACHIN, Gleisy R. B.; RADOS, Gregório J. V. Recuperar a informação eletrônica pela Internet. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 4, n. 4, p. 9-27, 1999.

BORGES, Mônica E. N.; CARVALHO, Natalia G. M. Produtos e serviços de informação para negócios no Brasil. **Ciência da Informação**, Brasília, v. 27, n. 1, p. 76-81, jan./abr. 1998.

BRANSKI, Regina M. Localização de informações na *web*. 2002. Artigo não publicado.

BRANSKI, Regina M. Localização de informações na Internet: características e formas de funcionamento dos mecanismos de busca. **Transinformação**, v. 12, n. 1, p. 11-19, jan./jun. 2000.

BUENO, Márcia C.; VIDOTTI, Silvana A. B. G. Ferramentas de busca na Internet: para quê, por quê e como utilizá-las? In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 11., 2000, Florianópolis. **Anais...** Florianópolis: UFSC/BU/CIN, 2000. 1CD.

BUENO, Márcia C.; VIDOTTI, Silvana A. B. G. Uso de ferramentas de busca da Internet. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECONOMIA, 3., 1999, Marília. **Anais...** Marília: Faculdade de Filosofia e Ciências da UNESP, 1999. p. 39-49.

CARVALHO, Fábio C. A. **Gestão do conhecimento: o caso de uma empresa de alta tecnologia**. 2000. 115 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2000.

CARVALHO, Rodrigo B.; FERREIRA, Marta A. T. Acelerando a espiral do conhecimento com a tecnologia da informação. In: SIMPÓSIO INTERNACIONAL DE GESTÃO DO CONHECIMENTO/GESTÃO DE DOCUMENTOS, 4., 2001, Curitiba. **Anais...** Curitiba: [s.n.], 2001.

CENDÓN, Beatriz V. Bases de dados de informação para negócios. **Ciência da Informação**, Brasília, v. 31, n. 2, p. 30-43, maio/ago. 2002.

CENDÓN, Beatriz V. A internet. In: CAMPELLO, Bernadete S.; CENDÓN, Beatriz V.; KREMER, Jeannette M. (Org.). **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: Ed. UFMG, 2000. p. 275-305.

CERVO, Amado L.; BERVIAN, Pedro A. **Metodologia científica**. 4. ed. São Paulo: Makron Books, 1996.

CHAGAS, Joseane et al. Interação do usuário na busca de informações. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 11., 2000, Florianópolis. **Anais...** Florianópolis: UFSC/BU/CIN, 2000. 1CD.

CIACONI, Regina B. Banco de dados de acesso público. **Ciência da Informação**, Brasília, v. 16, n. 1, p. 53-59, jan./jun. 1987.

COSTA, Marília D. et al. Gestão de informações ou gestão de conhecimento? In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 19., 2000, Porto Alegre. **Anais...** Porto Alegre: ARB/PUCRS, 2000.

CUENCA, Angela M. B. O usuário final da busca informatizada: avaliação da capacitação no acesso a bases de dados em biblioteca acadêmica. **Ciência da Informação**, Brasília, v. 28, n. 3, p. 293-301, set./dez. 1999.

CUNHA, Murilo B. As tecnologias de informação e a integração das bibliotecas brasileiras. **Ciência da Informação**, Brasília, v. 23, n. 2, p. 182-189, maio/ago. 1994.

CUNHA, Murilo B. **Para sabe mais: fontes de informação em ciência e tecnologia**. Brasília: Briquet de Lemos/Livros, 2001.

CYBIS, Walter A. **A identificação dos objetos de interfaces homem-computador e de seus atributos ergonômicos**. 1994. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 1994.

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

DAVENPORT, Thomas H.; PRUSAK, Laurence. **Conhecimento empresarial: como as empresas gerenciam o capital intelectual**. São Paulo: Campus, 1998.

DAVYDOV, Mark M. EIP: the second wave. **Intelligent Enterprise**, v. 3, n. 4, Mar. 2000. Disponível em: <[www.intelligententerprise.com/000301/supplychain.shtml](http://www.intelligententerprise.com/000301/supplychain.shtml)>. Acesso em: 20 ago. 2001.

DIAS, Cláudia A. Portal corporativo: conceitos e características. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 50-60, jan./abr. 2001.

DUARTE, Denio; MIRANDA, Paulo. Sistemas gerenciadores de documentos, uma base para knowledge management. In: SIMPÓSIO INTERNACIONAL DE GESTÃO DO CONHECIMENTO/GESTÃO DE DOCUMENTOS, 1999, Curitiba. **Proceedings...** Curitiba: Ed. Universitária Champagnat, 1999.

DUARTE, Denio; SZOSTAK, Ronaldo. Aplicando a tecnologia *wap* na gestão do conhecimento. In: WORKSHOP DO PROJETO IDOC- GESTÃO DE CONHECIMENTO BASEADO EM DOCUMENTOS, 1999, Curitiba. **Anais...** Curitiba: CEFET/PR, 1999.

ECKERSON, Wayne. 15 rules for enterprise portals. **Oracle Magazine**, July/Aug. 1999. Disponível em: <<http://www.oracle.com/oramag/oracle/99-Jul/index.html?49ind.html>>. Acesso em: 16 jul. 2002.

EIP, ou portal da empresa. **Informationweek**, 18 mar. 2000. Disponível em: <[http://www.actis.pt/Documentos/EIP\\_portal\\_da\\_empresa\\_IW.htm](http://www.actis.pt/Documentos/EIP_portal_da_empresa_IW.htm)>. Acesso em: 25 mar. 2002.

ENTERPRISE information portals. **Computerworld**, ano 15, n. 462, 21 jun. 2001. Disponível em: <<http://www.computerworld.iol.pt/index.asp?edicao=62>>. Acesso em: 25 mar. 2002.

EVANS, Thomas E.; YOKOYAMA, Alexandre. It is more than just knowledge management, it is portals. In: WORKSHOP DO PROJETO IDOC - GESTÃO DE CONHECIMENTO BASEADO EM DOCUMENTOS, 1999, Curitiba. **Anais...** Curitiba: CEFET/PR, 1999.

FACHIN, Gleisy R. B. **Modelo de avaliação para periódicos científicos *on line*: proposta de indicadores bibliográficos e telemáticos**. 2002. 210 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

FREITAS, Maria do Carmo D. et al. Portal cerâmica Brasil: o espaço virtual do setor cerâmico. **Revista Cerâmica Informação**, Florianópolis, v. 13, p. 68-73, nov. 2000.

FURRIE, Betty. **O MARC bibliográfico: um guia introdutório, catalogação legível por computador**. Brasília: Thesaurus, 2000.

GROGAN, Denis. **A prática do serviço de referência**. Brasília: Briquet de Lemos/Livros, 1995.

GUINCHAT, Claire; MENU, Michel. **Introdução geral às ciências e técnicas da informação e documentação**. Brasília: IBICT, 1994.

HEEMANN, Vivian. **Avaliação ergonômica de interfaces de dados por meio de *checklist* especializado**. 1997. 96 f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 1997.

HEINZLE, Roberto. **Protótipo de uma ferramenta para criação de sistemas especialistas baseados em regras de produção**. 1995. 145 f. Dissertação (Mestrado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 1995.

KLEINSCHMIDT, Cory; GOODMAN, Andrew. Frequently asked questions about portals (FAQs). 24 mar. 2002. Disponível em: <<http://www.traffick.com/article.asp?aID=9#what>>. Acesso em: 22 maio 2002.

- LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos/Livros, 1993.
- LOH, Stanley. **Descoberta de conhecimento em bases de dados textuais**. Maio, 1997. Disponível em: <<http://mozart.ulbra.tche.br/~apostilas/dc-texto.htm>>. Acesso em: 24 set. 2002.
- LOPES, Ilza Leite. **Uso das linguagens controlada e natural em bases de dados: revisão da literatura**. **Ciência da Informação**, Brasília, v. 31, n. 1, p. 41-52, jan./abr. 2002.
- MARCONI, Marina A.; LAKATOS, Eva M. **Técnicas de pesquisa**. 2. ed. São Paulo: Atlas, 1990.
- MENDES, Raquel D. **Inteligência artificial: sistemas especialistas no gerenciamento da informação**. **Ciência da Informação**, Brasília, v. 26, n. 1, p. 39-45, jan./abr. 1997.
- MOREIRA, Wálter. **Eficácia dos mecanismos de busca na Internet**. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 19., 2000, Porto Alegre. **Anais...** Porto Alegre: FEBAB: ARB: PUCRS, 2000.
- MURRAY, Gerry. **The portal is the desktop**. **Group Computing Magazine**, May 1999. Disponível em: <<http://www.e-promag.com/eparchive/index.cfm?fuseaction=viewarticle&ContentID=166&websiteid=>>>. Acesso em: 16 jul. 2002.
- NAHUZ, Fernanda. **World wide web: aspectos teóricos dos mecanismos de busca**. **Informação e Sociedade: Estudos**, v. 9, n. 2, 1999. Disponível em: <[www.informacaoesociedade.ufpb.br](http://www.informacaoesociedade.ufpb.br)>. Acesso em: 30 jul. 2002.
- NONAKA, Ikujiro; TAKEUCHI, Hirotaka. **Criação de conhecimento na empresa**. Rio de Janeiro: Campus, 1997.
- O'DELL, Carla; GRAYSON JÚNIOR, C. Jackson. **Ah... se soubéssemos antes o que sabemos agora**. São Paulo: Futura, 2000.
- OLIVEIRA, Elaine R. **Avaliação ergonômica de interfaces da SciELO – Scientific Electronic Library On line**. 2001. 113 f. Dissertação (Mestrado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2001.
- PARIZOTTO, Rosamelia. **Elaboração de um guia de estilos para serviços de informação em ciência e tecnologia via web**. 1997. 90 f. Dissertação (Mestrado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 1997.
- PEREIRA, Maria de Nazaré Freitas et al. **Bases de dados na economia do conhecimento: a questão da qualidade**. **Ciência da Informação**, Brasília, v. 28, n. 2, p. 215-223, maio/ago. 1999.

REYNOLDS, Hadley; KOULOPOULOS, Tom. Enterprise knowledge has a face. **Intelligent Enterprise**, v. 2, n. 5, Mar. 1999. Disponível em: <[http://www.intelligententerprise.com/db\\_area/archives/1999/993003/feat1.shtml](http://www.intelligententerprise.com/db_area/archives/1999/993003/feat1.shtml)>. Acesso em: 20 ago. 2002.

ROBREDO, Jaime; CUNHA, Murilo B. **Documentação de hoje e de amanhã: uma abordagem informatizada da biblioteconomia e de sistemas de informação**. São Paulo: Global, 1994.

ROMANI, Claudia; BORSZCZ, Iraci. Banco de talentos: ferramenta para mapear o conhecimento nas organizações. **Revista de Ciências da Administração**, Florianópolis, ano 3, n. 6, p. 21-30, set. 2001.

ROWLEY, Jennifer. **A biblioteca eletrônica**. Brasília: Briquet de Lemos/Livros, 2002.

ROWLEY, Jennifer. **Informática para bibliotecas**. Brasília: Briquet de Lemos/Livros, 1994.

SILVA, Edna L.; MENEZES, Estera M. **Metodologia da pesquisa e elaboração de dissertação**. 3. ed. rev. atual. Florianópolis: Laboratório de Ensino a Distância, 2001.

SILVA, Helena P. **Inteligência competitiva na Internet: proposta de um modelo**. 2000. 188 f. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2000.

STRAUHS, Faimara R.; ABREU, Aline F.; RENAUX, Douglas P. B. Uma proposta para gestão de conhecimento em laboratório acadêmico. In: SIMPÓSIO INTERNACIONAL DE GESTÃO DO CONHECIMENTO/GESTÃO DE DOCUMENTOS, 2000, Curitiba. **Anais...** Curitiba: Ed. Universitária Champagnat, 2000. p. 17-36.

TARGINO, Maria das Graças; GARCIA, Joana C. R. Ciência brasileira na base de dados do *Institute for Scientific Information* (ISI). **Ciência da Informação**, Brasília, v. 29, n. 1, p. 103-117, jan./abr. 2000.

TEIXEIRA, Cenivalda M. S.; SCHIEL, Ulrich. A Internet e seu impacto nos processos de recuperação da informação. **Ciência da Informação**, Brasília, v. 26, n. 1, p. 65-71, jan./abr. 1997.

TEIXEIRA FILHO, Jayme. A onda dos portais: comunidades virtuais de consumidores. **Insight Informal**, n. 31, 1999. Disponível em: <<http://www.informal.com.br/insight/insight.htm>>. Acesso em: 23 out. 2000.

TEIXEIRA FILHO, Jayme. **Gerenciando conhecimento: como a empresa pode usar a memória organizacional e a inteligência competitiva no desenvolvimento dos negócios**. Rio de Janeiro: Ed. SENAC, 2000.

TEIXEIRA FILHO, Jayme. Tecnologia da informação para a gestão do conhecimento. **Insight Informal**, n.10, 1998. Disponível em: <<http://www.informal.com.br/insight/insight10.html>>. Acesso em: 8 nov. 1999.

TERRA, José Cláudio C.; GORDON, Cindy. **Portais corporativos: a revolução na gestão do conhecimento**. São Paulo: Negócio Editora, 2002.

THIVES JÚNIOR, Juarez J.; ANGELONI, Maria T. Workflow, uma tecnologia para transformação do conhecimento nas organizações – estudo de caso no Conselho Estadual de Educação de Santa Catarina – CEE/SC. In: SIMPÓSIO INTERNACIONAL DE GESTÃO DO CONHECIMENTO/GESTÃO DE DOCUMENTOS, 2000, Curitiba. **Anais...** Curitiba: Ed. Universitária Champagnat, 2000. p. 317-332.

TOMAÉL, Maria I. et al. Avaliação de fontes de informação na Internet: critérios de qualidade. **Informação & Sociedade: Estudos**, v. 11, n. 2, 2001. Disponível em: <[www.informacaoesociedade.ufpb.br](http://www.informacaoesociedade.ufpb.br)>. Acesso em: 30 jul. 2002.