

**UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE CIÊNCIAS HUMANAS E LETRAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM SOCIEDADE E  
CULTURA NA AMAZÔNIA**

**MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO:  
*AValiação Estrutural dos Tesouros em  
BIOTECNOLOGIA***

SUELY OLIVEIRA MORAES MÁRQUEZ

MANAUS

2007

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

**UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE CIÊNCIAS HUMANAS E LETRAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM SOCIEDADE E  
CULTURA NA AMAZÔNIA**

SUELY OLIVEIRA MORAES MÁRQUEZ

**MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO:  
*AValiação Estrutural dos Tesouros em  
Biotecnologia***

Dissertação apresentada ao Programa de Pós-Graduação em Sociedade e Cultura na Amazônia da Universidade Federal do Amazonas como parte dos requisitos para obtenção do título de Mestre em Sociedade e Cultura na Amazônia.

Orientadora: Profª. Drª Célia Regina Simonetti Barbalho

MANAUS

2007

M827d Márquez, Suely Oliveira Moraes

Modelos de representação do conhecimento: avaliação estrutural dos  
tesauros em biotecnologia – 2007.

133 p.; 27cm; Ilust.

Dissertação (Mestrado) – Instituto de Ciências Humanas e Letras da  
Universidade Federal do Amazonas - UFAM.

Orientadora: Célia Regina Simonetti Barbalho, Dra.

1. Tesouro 2. Linguagem documentária 3. Biotecnologia I. Título

CDU: 004.823 :025.43

SUELY OLIVEIRA MORAES MÁRQUEZ

**MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO:  
AVALIAÇÃO ESTRUTURAL DOS TESAUROS EM  
BIOTECNOLOGIA**

Dissertação apresentada ao Programa de Pós-Graduação em Sociedade e Cultura na Amazônia da Universidade Federal do Amazonas como parte dos requisitos para obtenção do título de Mestre em Sociedade e Cultura na Amazônia, área de concentração Linguagem, cultura e comunicação na Amazônia.

Aprovado em \_\_\_\_\_ de \_\_\_\_\_ de 2007.

**BANCA EXAMINADORA**

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Célia Regina Simonetti Barbalho, Presidente  
Universidade Federal do Amazonas

Prof. Dr<sup>ª</sup>. Ierecê Barbosa Monteiro, Membro  
Universidade Estadual do Amazonas

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Márcia Eliane de Souza e Mello, Membro  
Universidade Federal do Amazonas

***DEDICATÓRIA***

*A meu esposo Gabriel, pela compreensão.*

*Ao meu filho querido Daniel e*

*A minha mais preciosa jóia Danielly Beatriz.*

## AGRADECIMENTOS

*Meus agradecimentos,*

*A DEUS, que com sua imensa grandeza me deu forças para terminar.*

*Aos meus filhos Daniel e Danielly Beatriz, que sempre me proporcionaram momentos de alegria e ternura.*

*Ao meu esposo Gabriel pelo incentivo, paciência e amor.*

*A minha orientadora Profa. Célia Regina Simonetti Barbalho pela honra do convívio acadêmico, pela dedicação ao ensino e pela atenção e rigor na orientação deste trabalho.*

*A minha mãe, que me incentivou a continuar.*

*Aos companheiros do Departamento de Biblioteconomia, pela amizade.*

*Aos professores do Programa de Pós-Graduação Sociedade e Cultura, em especial Narciso Júlio Freire Lobo, Gabriel Arcanjo Albuquerque, Marcos Frederico Krüger Aleixo, Nelson Matos de Noronha, Heloísa Helena Corrêa da Silva, Sérgio Ivan Gil Braga, Antônio Carlos Witikoski, Maria do Perpétuo Socorro R. Chaves, Heloisa Lara Campos da Costa e Ernesto Renan Melo Freitas Pinto pelos valiosos conhecimentos e experiências transmitidas para minha formação.*

*Aos colegas da turma pelo companheirismo e prazer de convívio.*

*A todos, que assim como eu, acreditam em um futuro melhor...*

*Eu pedi forças... E DEUS me deu dificuldades para me fazer forte.  
Eu pedi sabedoria... E DEUS me deu problemas para resolver.  
Eu pedi prosperidade... E DEUS me deu cérebro e músculos para trabalhar.  
Eu pedi coragem... E DEUS me deu perigo para superar.  
Eu pedi amor... E DEUS me deu pessoas com problemas para ajudar.  
Eu pedi favores... E DEUS me deu oportunidades.  
Eu não recebi nada do que pedi... Mas recebi tudo que precisava.*

***Ao Senhor, Mestre Supremo,  
meu eterno agradecimento.***

## ***RESUMO***

Discorre sobre a importância do tesouro como um instrumento de representação do conhecimento em Sistemas de Recuperação da Informação - SRI. A identificação de conceitos relacionada ao processo de representação de uma determinada área do conhecimento na linguagem do tesouro, garante uma terminologia atualizada para os SRI, o que facilita ao pesquisador/usuário no processo de busca, seleção e recuperação da informação e do documento, uma linguagem padronizada e por ordem de relevância. Com a atual diversidade de termos técnicos e científicos, bem como, pela demanda por maior rapidez e facilidade na recuperação da informação, se faz sentir uma necessidade crescente no desenvolvimento de instrumentos mais avançados e efetivos de organização de informações, onde controle da terminologia é feito por área do conhecimento. Enfim, em todas as formas de recuperação da informação o que mais importa é que o usuário identifique os documentos que vão ao encontro de suas necessidades, pois melhorar a qualidade dos serviços tem sido uma preocupação freqüente de bibliotecários e outros profissionais que trabalham, administram ou simplesmente usam esses serviços. A pesquisa caracterizou-se como um estudo do tipo qualitativo, baseada no método bibliográfico documental e exploratório-descritivo, realizada através da análise dos tesouros selecionados na área da biotecnologia, tendo como subárea produtos alimentícios e fármacos. Analisa tesouros na área da biotecnologia do ponto de vista da sua organização estrutural, tendo em vista a criação de um voltado para a biodiversidade amazônica, levando em conta não somente a inexistência de um, como também, a dificuldade da indexação, recuperação e disseminação da informação nessa área de conhecimento. Baseado na análise dos três tesouros selecionados apresenta, indicadores que podem contribuir para a elaboração e manutenção do tesouro em biotecnologia da Amazônia.

**Palavras-chave:** Tesouro; Linguagem documentária; Biotecnologia.

## ***ABSTRACT***

The importance of tesauro as an instrument of representation of the knowledge in Systems of Recovery of Information - SRI. The identification of concepts related to the process of representation of one determined area of the knowledge in the language of tesauro, guarantees a terminology brought up to date for the SRI, what it facilitates to the researcher/user in the process of search, election and recovery of the information and the document, a standardized language and for relevance order. With the current diversity of scientific terms technician and, as well as, for the demand for bigger rapidity and easiness in the recovery of the information, if it makes to feel an increasing necessity in the development of more advanced and effective instruments of organization of information, where control of the terminology is made by area of the knowledge. At last, in all the forms of recovery of the information what more it matters are that the user identifies the documents that go to the meeting of its necessities, therefore to improve the quality of the services has been a frequent concern of librarians and other professionals that work, manage simply or use these services. The research was characterized as a study of the qualitative type, based in the documentary bibliographical method and exploration-description, carried through through the analysis of the tesauros selected in the area of the biotechnology, having as sub-area nourishing and pharmacon products. He analyzes tesauros in the area of the biotechnology of the point of view of its structural organization, in view of the creation of one directed toward Amazonian biodiversity, taking in account not only the inexistence of one, as also, the difficulty of the indexation, recovery and dissemination of the information in this area of knowledge. Based in the analysis of the three selected tesauros it presents, pointers that can contribute for the elaboration and maintenance of tesauro in biotechnology of the Amazonian.

**Word-key:** Tesauro; Documentary language; Biotechnology.

## LISTA DE FIGURAS

Figura 1 - Modelo de sistema de comunicação .....	39
Figura 2 – Modelo de processo de comunicação.....	40
Figura 3 – Processo de comunicação num serviço de informação.....	41
Figura 4 – Modelo de processo de comunicação.....	42
Figura 5 – Tradicional modelo de comunicação .....	43
Figura 6 - Esquema de comunicação proposto por Coelho Netto.....	44
Figura 7 – Esquema de comunicação documentária estruturado.....	46
Figura 8 – Triângulo de Ogden e Richards.....	56
Figura 9 – Página principal do tesauros HONselect.....	93
Figura 10– Página mostrando a área de Tecnologia, alimentos e bebidas.....	95
Figura 11 – Lista de favoritos.....	96
Figura 12 – Lista de doenças raras .....	97
Figura 13– Página mostrando as áreas trabalhadas pelo HONselect.....	98
Figura 14– Página mostrando a área de alimentos e bebidas.....	99
Figura 15– Página mostrando o assunto de bebidas.....	100
Figura 16– Página mostrando o assunto de bebidas.....	101
Figura 17– Página mostrando lista de favoritos em ordem alfabética.....	102
Figura 18– Página contendo introdução.....	103
Figura 19– Página mostrando a quantidade de termos.....	104
Figura 20– Página mostrando a relação associativa.....	105
Figura 21– Página mostrando a relação equivalência.....	106
Figura 22– Página mostrando a relação hierárquica.....	107
Figura 23– Página mostrando nota esclarecedora.....	108
Figura 24– Página mostrando a morfologia das palavras.....	109
Figura 25 - Página mostrando os idiomas.....	110
Figura 26– Forma de apresentação do TCA.....	111
Figura 27 - Forma de apresentação do TCA.....	112
Figura 28 – Índice do Tesauro Cadeia Alimentícia.....	113
Figura 29 – Índice alfabético do Tesauro Cadeia Alimentícia.....	114
Figura 30 - Página mostrando os tipos de relacionamentos.....	115
Figura 31 – Notas esclarecedoras.....	116
Figura 32 – Morfologia das palavras.....	117
Figura 33 – Página mostrando o idioma do TCA.....	118
Figura 34 – Página principal do glossário de biotecnologia em alimentação e agricultura.....	119
Figura 35 – Página mostrando a lista de entrada em ordem alfabética.....	121
Figura 36 – Página mostrando a lista de entrada em ordem alfabética – letra B.....	122
Figura 37 – Página mostrando tipos de relacionamentos.....	124
Figura 38 – Página mostrando tipos de relacionamentos.....	125
Figura 39 – Página mostrando notas esclarecedoras.....	126
Figura 40– Página mostrando morfologia das palavras.....	127
Figura 41– Página mostrando morfologia das palavras.....	128
Figura 42– Página mostrando forma de busca.....	129
Figura 43– Página mostrando os idiomas.....	130

## LISTA DE QUADROS

Quadro 1 - Linguagem natural: vantagens e desvantagens .....	72
Quadro 2 – Vocabulário controlado: vantagens e desvantagens .....	74
Quadro 3 – Tesouros avaliados com suas variáveis .....	132

## SUMÁRIO

INTRODUÇÃO.....	23
SISTEMA DE RECUPERAÇÃO DA INFORMAÇÃO.....	34
1.1 EFICÁCIA DO SRI.....	36
1.2 COMUNICAÇÃO DOCUMENTÁRIA.....	38
1.3 TERMINOLOGIA.....	47
1.4 TERMINOLOGIA COMO INSTRUMENTO DA LD.....	51
1.5 SIGNO LINGUÍSTICO.....	53
1.6 SIGNO DOCUMENTÁRIO.....	57
INDEXAÇÃO.....	61
2.1 PROCESSO EVOLUTIVO.....	62
2.2 INDEXAÇÃO NO CICLO DOCUMENTÁRIO.....	64
2.3 LINGUAGENS DOCUMENTÁRIAS.....	66
2.4 POLÍTICA DE INDEXAÇÃO.....	67
2.5 LINGUAGEM NATURAL.....	70
2.6 LINGUAGEM CONTROLADA.....	73
2.7 TESAURO.....	75
2.7.1 Considerações Gerais.....	76
2.7.2 Função do Tesouro.....	79
2.7.3 Tipos de Tesouros.....	80
2.7.4 Tipos de Relações nos Tesouros.....	80
2.7.5 Metodologia de elaboração dos Tesouros.....	82
2.8 TESAUROS E A BIOTECNOLOGIA.....	84
2.9 APLICAÇÕES DA BIOTECNOLOGIA.....	86
PERCURSO METODOLÓGICO.....	89
3.1 UNIVERSO E AMOSTRA.....	89
3.2 VARIÁVEIS DE ANÁLISE.....	90
3.3 PROCEDIMENTOS DE COLETA DE DADOS.....	91
ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	93
4.1 ANÁLISE INDIVIDUAL DOS TESAUROS.....	93
4.2 ANÁLISE GLOBAL.....	131
CONCLUSÃO.....	138
REFERÊNCIAS.....	141

## INTRODUÇÃO

A intensificação das relações sociais, culturais, políticas e econômicas entre os países, conseqüência do crescente desenvolvimento científico e tecnológico, conduz a uma demanda cada vez maior por informações produzidas por esta nova sociedade constituída a partir da chamada globalização. Esse processo tem sido consolidado pela intensa revolução nas tecnologias de informação e comunicação que se apresentam em forma de produtos como os telefones, televisão, computadores, bases e banco de dados sofisticados, etc.

Tal contexto tem favorecido a oferta em grande quantidade de informação sendo que, tal abundância não é garantia de fácil acesso e absorção dos saberes disponíveis quer seja pelos problemas ocasionados pelo analfabetismo tecnológico e digital, quer seja pela ausência de mecanismos que assegurem a sua rápida e eficaz recuperação.

Para melhor compreensão, neste estudo, informação deve ser entendida como “[...] conhecimento<sup>1</sup> inscrito (gravado) sob a forma escrita (impressa ou numérica), oral ou audiovisual” (LE COADIC, 1996, p. 05), independente de qualquer suporte físico, que segundo Miranda (1996, p.310) “[...] não depende de registro material para existir, e, por esse motivo requer novas abordagens teóricas e metodológicas, novas práticas e novas tecnologias para seu ciclo de vida e transformação”.

Diante do exposto, um dos meios possíveis de tornar mais democrático o uso da informação é o desenvolvimento de investigações sistemáticas sobre sua transferência e relação com a aprendizagem, ou seja, constituir mecanismos de busca de informação mais eficazes e interativas que impliquem no estudo de vocabulários inter e multidisciplinares e em bases multilingüísticas que permitam a sua fácil recuperação.

---

<sup>1</sup>Resultado do ato de conhecer; ato pelo qual o espírito apreende um objeto. Conhecer é ser capaz de formar a idéia de alguma coisa.

Gonzalez de Gomez (1993) destaca que o efeito globalizador dos mercados e as tecnologias de informação pressupõem a vigência de uma premissa epistemológica de máximo alcance, que assegure condições de uniformidade e padronização lingüística no uso da informação, ou seja, para a autora assegurar o acesso à informação disponibilizada nos mais variados meios e suportes significa olhar atentamente para as questões que envolvem o modo como ela é armazenada, criando padrões seguros para permitir uma eficaz recuperação.

Desta forma, pode-se inferir que uma das qualidades necessárias para que a globalização se efetive como um processo de desenvolvimento democrático é a criação de condições para que a geração, uso e transferência da informação seja também um processo democrático, e isto só se tornará possível quando todos os envolvidos dominarem o código utilizado para a representação da recuperação da informação, ou seja, em uma situação ideal de inexistência de ruídos no processo de comunicação.

Ferneda (2003) afirma que no contexto da Ciência da Informação, o termo “recuperação da informação” significa, para uns, a operação pela qual se seleciona documentos. Para outros, consiste no fornecimento dos elementos de informação documentária correspondente. O termo pode ainda ser empregado para designar a operação que fornece uma resposta mais ou menos elaborada a uma demanda, sendo esta convertida em um produto cujo formato é acordado com o usuário. Há ainda autores que conceituam a recuperação da informação de forma muito mais ampla, ao subordinar à mesma o tratamento da informação, através das atividades de catalogação, indexação e classificação.

Independente do sentido, o processo de recuperação da informação deverá atender às necessidades da clientela que faz uso de um determinado sistema de recuperação, motivo pelo qual eles deverão promover a rápida seleção das informações desejadas, permitindo à localização dos documentos selecionados.

O processo de indexação busca descrever ou identificar cada documento através de seu conteúdo. Tal representação geralmente é realizada através do processo de indexação, onde são extraídos conceitos do documento através da análise de seu conteúdo e traduzidos em termos de uma linguagem de indexação, tais como cabeçalhos de assunto, tesouros e outros. Esta representação identifica o documento e define seus pontos de acesso para a busca e recuperação dos mesmos.

A análise de um documento pode envolver uma interpretação de seu conteúdo com a finalidade de agregar assuntos que não estão diretamente explicitados em sua superfície textual, mas que podem ser facilmente abstraídos por um indexador humano. A indexação de um documento é efetuada tendo em vista a sua recuperação. Nesse caso a análise do documento é feita com a preocupação de tornar o seu conteúdo visível para os usuários de um sistema de informação.

Assim, para o usuário obter êxito no processo de busca, faz-se necessário que os descritores registrados nos sistemas de recuperação, através da linguagem natural ou artificial, que serão vista nos itens 2.5 e 2.6, estejam devidamente estruturados, pois poderão se constituir em facilitadores ou dificultadores para a busca, já que permitirão a recuperação de informações.

Nesse sentido, Nascimento (2004) destaca que o processo de busca da informação, desde sua criação, enfrenta o problema da recuperação relevante, em função das falhas junto às atividades de indexação. Baeza-Yates e Ribeiro-Neto (1999, p. 31) afirmam que:

[...] os principais objetivos da Recuperação de Informações são construir índices eficientes, processar consultas de usuários com bom desempenho e desenvolver algoritmos de ordenação (*ranking*) que aumentem a qualidade das respostas.

Em vista dos motivos expostos pelos autores é que se destaca a importância da linguagem documentária para o processo rápido e preciso de recuperação da informação.

Com base em Guinchat e Menou, (1994), é possível afirmar que a linguagem documentária é concebida como um instrumento utilizado por um centro ou unidade de informação para a representação do conteúdo dos documentos. Contudo, o uso incorreto das linguagens naturais, ou seja aquela que não é padronizada, prejudicará o tratamento da informação e, conseqüentemente, sua recuperação.

Por esta razão, as linguagens documentárias possuem um conjunto de regras. Por ser linguagem, ela é composta de unidades sógnicas. O signo lingüístico é constituído de um conceito (significado) e de uma imagem acústica (significante).

Isso implica afirmar que as linguagens documentárias devem ter os signos com significados recortados e fixados para evitar a polissemia. Devem, além disso, ter os significados partilhados com a comunidade de usuários do sistema informacional. Com efeito, os descritores devem ter significado invariante para representar informações de maneira previsível. Estabelecer o significado nas linguagens documentárias é, desse modo, uma atividade complexa, pois reconceituar ou delimitar termos oriundos da linguagem natural, com sua pluralidade de significados, implica em uma escolha entre várias possibilidades, que pode ser influenciada por questões relacionadas a grupos sociais, regiões e países.

Por ser uma linguagem constituída de parâmetros estáveis e predeterminados cultural e socialmente, vinculada a características institucionais, uma linguagem documentária interfere nas formas de organização e disseminação da informação. Para isso, as linguagens documentárias, ao serem estabelecidas, deverão ir ao encontro das expectativas e necessidades dos usuários (LANCASTER, 1987).

Ademais, a atualização das linguagens documentárias deve ser uma constante, devido às mudanças correntes na linguagem natural e na área de especialidade. Guinchat e Menou (1994) sugerem controlar o emprego da linguagem registrando-se as dificuldades e soluções

encontradas, bem como a ocorrência dos descritores. Deve-se, além disso, registrar os termos que não aparecem na linguagem para futura decisão sobre eventual inclusão.

Desse modo, afirma-se que somente com um vocabulário construído a partir de uma rede de significações lógico-semânticas, baseadas em um dicionário terminológico, com descritores definidos de forma clara é possível proporcionar o rigor necessário para a melhoria dos procedimentos de indexação (TÁLAMO et. al., 1997).

Na perspectiva mais pragmática, a informação deixa de ser um dado e passa a ser construção. Portanto, no âmbito documentário, representar conteúdos para constituir informação significa recortar, segmentar. Com efeito, para que se possa falar de transmissão de informação documentária é necessário formular hipóteses (e explicitá-las) que sejam razoavelmente compartilhadas pela comunidade de usuários do sistema que se propõe, já que um mesmo universo documental pode ser organizado e representado de diversas maneiras, pautando-se nos distintos propósitos a que se destinam.

Mediante o exposto, resta afirmar que a construção de hipóteses de organização dos sistemas de informação ou, mais especificamente, das Linguagens Documentárias, não pode ser um processo aleatório. Embora se possa afirmar que as propostas de organização têm ingredientes empíricos, isso só pode ser entendido como fundado em necessidades pragmáticas. Dito de outro modo, uma linguagem documentária que será utilizada para "representar conteúdos" não pode ser formulada a partir de opiniões ou referências particulares e isoladas. Ao contrário, se o objetivo dos sistemas documentários é transmitir informação, tais linguagens devem ser organizadas a partir de hipóteses que façam sentido para determinados objetivos.

Todavia, embora o uso da terminologia seja crucial para referendar o processo de construção de linguagens documentárias – e, com isso, possibilitar melhor representação e recuperação da informação -, ela não garante por si só o sucesso da comunicação

documentária. Isso significa dizer que as hipóteses de organização de tais linguagens são mais complexas do que poderia parecer inicialmente, uma vez que a atividade de informar requer, além dos parâmetros de um sistema conceitual, a observação das necessidades de uso ou recepção da informação.

A linguagem documentária, portanto, não é a simples reprodução da terminologia de uma determinada área porque ambas se propõem as tarefas distintas. Se a terminologia visa estudar a estrutura, a formação, o desenvolvimento, uso e gestão em diferentes domínios, a Ciência da Informação, quando constrói linguagens documentárias, tem como objetivo assegurar a organização e a transferência da informação.

Uma vez que a informação é uma construção, a linguagem documentária, embora se aproprie das terminologias e dos sistemas conceituais por ela mapeados, se propõe como modo de organização do conhecimento, para o que articula temas, subtemas e termos orientados para atender a conjuntos documentários e para transferir informação para determinados públicos segundo objetivos específicos. Com efeito, ela segmenta, a seu modo, o universo enfocado segundo perspectivas pragmáticas. A partir dessa segmentação que se apresenta como proposta organizadora, uma linguagem documentária configura-se como um instrumento facilitador da comunicação em contextos documentários específicos.

Ao contrário da função de dispor e localizar o documento na sua forma física integral, as linguagens documentárias – listas de cabeçalhos de assuntos e os tesouros – orientam a organização intelectual para a desconstrução do texto em língua natural e sua representação codificada e sintética, constituindo-se em instrumento de análise semântica e sintática. Visam, portanto, obter, por ocasião das respostas aos pedidos de informação, um máximo de indicadores relevantes ou pertinentes e, para tal, de coincidência entre as descrições de autoria do indexador e as dos usuários.

Os tesouros possuem uma característica singular no âmbito das Linguagens Documentárias que é a organização do seu campo terminológico. Segundo Leska (1979), o desenvolvimento de terminologias é direcionado à melhor designação para o significado dos conceitos. Este é, precisamente, o campo onde os tesouros se identificam com os sistemas conceituais, ou seja, para cada conceito só pode existir uma representação simbólica, designada por termo ou descritor.

A construção e, conseqüente, avaliação de um tesouro não é tarefa simples, porém necessária ao bom desenvolvimento da linguagem controlada. Cada linguagem construída é única, especial para um domínio do conhecimento e, portanto, passível de ser modificada na mesma medida em que as línguas naturais evoluem.

A avaliação dos tesouros deve ser uma constante, pois velhos conceitos em desuso podem ser retomados. Os usuários do sistema de informação para o qual a linguagem foi construída podem ter novas necessidades de informação, sejam elas disciplinares ou interdisciplinares. Não importa se o tesouro vai ser gerado manual ou automaticamente, pois, com certeza, todas essas hipóteses vão estar refletidas na literatura, que é o ponto de partida para a construção de novas linguagens.

A análise do domínio para o qual será desenhado o tesouro, a identificação dos conceitos necessários à sua compreensão e o processo de relacionamento conceitual, etapas necessárias à construção de qualquer linguagem artificial, são processos que, tanto dedutiva quanto indutivamente, vão criar categorias que possibilitarão a desconstrução do documento e sua reconstrução simbólica, sintética e representacional, em uma linguagem documentária capaz de fazer coincidir as necessidades de informação do usuário com os registros disponíveis na memória documentária.

No contexto atual há uma sobrecarga de informação sem precedentes. Seja pela página impressa, pelos meios e suportes eletrônicos, pelo rádio, pela televisão, ou qualquer outro, o mundo está saturado de informação. Shenk (1997, p.38) afirma que “[...] a sobrecarga de informação emergiu como ameaça genuína [...]. Enfrentamos agora a perspectiva da obesidade de informação”.

A palavra informação, originada do termo latino *informare*, é permeada da concepção de moldar um material como o oleiro faz com o barro, ou seja, pela própria morfologia o termo enfatiza que ela é particular e individualmente assimilada, de modo diferente, por cada um que com ela entra em contato. Com efeito, no contexto atual, transformar informação em conhecimento implica, dentre outras exigências, na necessidade de compreender e lidar com os mecanismos de busca e recuperação existentes não só nos suportes eletrônicos como também nos demais que, através do emprego da linguagem documentária, reúnem o conteúdo extraído de vários suportes com a finalidade de permitir a recuperação da informação.

A linguagem documentária, enquanto codificadora e decodificadora, opera no sentido de construir representações de conteúdos na expectativa de organizar a informação para garantir sua posterior recuperação.

Lara (2001) afirma que o processo de representação em documentação se desenvolve a partir da seleção de algumas qualidades (propriedade) que são mais perceptíveis pelo indivíduo que está a indexar, fruto da segmentação de conteúdos e Eco (1988, p.55) complementa afirmando que diante um fenômeno desconhecido “[...] reagimos por aproximação, procuramos aquele recorte de conteúdo, já presente na nossa enciclopédia, que bem ou mal parece prestar contas do novo fato”.

Neste sentido, o processo de representação da informação documentária é permeado por uma linguagem intermediária que opera pelo recorte dos conteúdos de um dado texto visando organizar para transferir, transferir para permitir a apropriação da informação. É,

portanto, portadora de significados que demandam por estudos sistematizados de modo a constituir análises que viabilizem a compreensão de sua construção.

Assim, elegendo a área de biotecnologia pela relevância que ela possui no contexto regional, este estudo busca examinar os aspectos estruturais dos tesouros dessa área, de modo a compreender a eficácia da linguagem documentária adotada no processo de comunicação que ela desencadeia quando do manuseio pelo usuário, que demanda por dominar um código subjetivo estabelecido pelo indexador.

Deste modo, a comprovação da capacidade de organização do conhecimento e conseqüente recuperação da informação por estes tesouros poderá contribuir para a criação e consolidação de mecanismos desta natureza, com uma metodologia de avaliação clara, pautada em modelos já disseminados internacionalmente, de forma a beneficiar a geração de elementos que permitam a disseminação de saberes, através de sua ágil recuperação com fim último de potencializar o desenvolvimento da região e do uso da biota amazônica.

A biotecnologia é um processo tecnológico que permite a utilização de material biológico (plantas e animais) para fins industriais. Por ser multidisciplinar, envolve a aplicação em inúmeros bens e serviços obtidos através dela que favorece o incremento de áreas como:

Agricultura – adubo composto, pesticidas, silagem, mudas de plantas, plantas transgênicas, etc.

Alimentação – pães, queijos, pickles, cerveja, vinho, proteínas unicelular, aditivos, etc.

Química – butanol, acetona, glicerol, ácidos, enzimas, metais, etc.

Eletrônica – biosensores.

Energia – etanol, biogás.

Meio ambiente – recuperação de petróleo, tratamento do lixo, purificação da água.

Pecuária – embriões.

Fármacos – antibióticos, hormônios e outros produtos farmacêuticos, vacinas, reagentes e testes para diagnósticos, etc.

Dentre as áreas acima relacionadas, este projeto se debruçará sobre as de alimentos e fármacos tendo em vista serem estes campos os que vêm apresentando maior aplicabilidade quanto ao desenvolvimento de produtos e serviços oriundos da biodiversidade amazônica.

Diante do exposto, o objetivo geral desta pesquisa é examinar a composição das linguagens documentárias presentes nos tesouros em biotecnologia e os objetivos específicos são:

- Dissertar sobre a importância da linguagem documentária para o processo de Serviço de Recuperação da Informação (SRI);

- Identificar os tesouros existentes na área de biotecnologia, tendo como subárea alimentos e fármacos;

- Analisar a composição estrutural da linguagem controlada empregada em três tesouros da área;

- Investigar as peculiaridades estruturais de três tesouros em biotecnologia, tendo como subárea produtos alimentícios e fármacos;

- Comparar a composição estrutural dos tesouros analisados;

- Apontar elementos que favoreçam a geração de um tesouros em biotecnologia para a Amazônia.

Visando atender aos objetivos propostos, esta dissertação apresenta a seguinte estrutura:

O primeiro capítulo destaca teoricamente o Sistema de Recuperação da Informação (SRI), mostrando sua eficácia na comunicação documentaria através das terminologias e dos signos lingüísticos.

No segundo capítulo é explorada a temática da indexação, desde seu processo evolutivo até sua importância na linguagem controlada, mostrando, dessa forma, o tesouro como resultado dessa técnica.

No terceiro capítulo são registrados os percursos metodológicos empregados para efetivar o estudo, de modo a possibilitar a compreensão do universo e composição da amostra da pesquisa, bem como das variáveis compostas para análise dos tesouros selecionados.

No quarto capítulo apresenta-se a análise dos tesouros selecionados, considerando inicialmente o contexto individualizado a partir das variáveis compostas para exame proposto e, posteriormente, cotejando as características particulares no âmbito do conjunto do corpus selecionado.

## CAPÍTULO I

Este capítulo discute o Sistema de Recuperação da Informação (SRI), que entre outras funções, objetiva oferecer acesso às informações contidas em documentos, facilitando o processo de comunicação através de sua eficácia e proporcionando mais satisfação aos usuários. Deste modo, teoricamente o capítulo reflete sobre a eficácia dos SRI, a comunicação documentária, a terminologia e sua importância enquanto elemento de uma linguagem documentária, o signo lingüístico e o signo documentário, na expectativa de propiciar uma ampla compreensão sobre a temática em estudo.

### SISTEMA DE RECUPERAÇÃO DA INFORMAÇÃO

Sistemas de informação serão considerados, no contexto deste estudo, como sinônimo de Sistemas Recuperação da Informação (SRI) que, dentre outras funções tendem a recuperação de modo a atender às necessidades dos usuários. Recuperar é tornar acessível os estoques informacionais organizados e armazenados nesses sistemas para garantir a eficácia do processo de busca e recuperação da informação pelo usuário.

Cesarino (1985, p.157) define Sistema de Recuperação da Informação (SRI) como “[...] um conjunto de operações consecutivas executadas para localizar, dentro da totalidade de informações disponíveis, aquelas realmente relevantes”. Salienta ainda, que a interação usuário e sistema é fundamental, evidenciando que o ato de recuperar informações se constitui como um sistema de comunicação.

Buckland (1991, apud LIMA, 1998, p.12) define sistemas de informação como “[...] quaisquer unidades que colem, tratem, organizem e disponibilizem ‘coisas’ potencialmente informativas”. Já Cohen (1995, p.14), os define como “[...] canais formais e informais de comunicação da informação dentro de uma organização ou de uma comunidade [...]”, sendo

que cada sistema de informação é organizado de acordo com as necessidades de seus usuários e sempre como apoio a uma atividade principal, seja ela ligada ao setor produtivo ou não.

De acordo com estas definições, um sistema de informação realiza um conjunto de ações que permitem a organização e comunicação do conhecimento. Em outras palavras, cabe a ele possibilitar a transferência de informações aos usuários através de procedimentos seletivos que regulam sua geração, distribuição e uso. Conforme Gonzalez de Gomez (1993), esta transferência ocorre através de um processo de comunicação entre os estoques de saberes e os usuários. Os SRI têm por princípio básico possibilitar ao usuário o acesso à informação de forma rápida e precisa e, para que isso aconteça, vários são, atualmente, os mecanismos utilizados para representar o conhecimento de uma determinada área do saber.

Os sistemas de recuperação da informação são consolidados através das seguintes etapas:

1. Indexação – é a operação pela qual escolhe-se os termos mais apropriados para descrever o conteúdo de um documento, que é expresso pelo vocabulário da linguagem documental escolhida pelo sistema e os termos são ordenados para constituir índices que servirão à pesquisa bibliográfica.

2. Armazenamento – processo que consiste em guardar documentos em meios impressos e eletrônicos, em áreas que lhes são próprias.

3. Recuperação – processo onde se localiza os artefatos da informação.

Nesse sentido, a eficácia da recuperação da informação depende de todo o processo de indexação, ou seja, da análise temática dos documentos, da correta determinação de termos pertinentes para serem indexados e do vocabulário controlado utilizado para representar os conteúdos expostos. Os conceitos e termos são, portanto, elementos primordiais de qualquer esquema de classificação e dos tesouros, que são tipos de instrumento de controle terminológico, objeto de estudo desta pesquisa que será amplamente discutido no Capítulo 2,

deste trabalho dissertativo. Deste modo, no próximo tópico, abordar-se-á a eficácia do SRI com objetivo de promover um amplo entendimento sobre suas características.

### **1.1 EFICÁCIA DO SRI**

A rápida evolução dos campos científicos e tecnológicos, o surgimento de novas áreas de conhecimento, o trabalho científico interdisciplinar, os sistemas eletrônicos para o processamento de dados e a disponibilidade dos catálogos automatizados em redes de acesso remoto estão exigindo mudanças nos sistemas tradicionais de tratamento da informação, por eles serem rígidos e incapazes de se adequarem ao ritmo de desenvolvimento das distintas áreas científicas e tecnológicas.

O uso de uma linguagem controlada, como os tesouros, um conjunto controlado de termos entre os quais foram estabelecidas relações hierárquicas e de associação, e as listas de cabeçalhos de assunto, facilitam o controle do vocabulário indexado e possibilitam uma maior precisão na hora da recuperação da informação, mas pode não ser suficiente para que a comunicação entre o usuário e o sistema seja satisfatória e garanta o acesso à informação buscada.

O grau de eficácia de um SRI, entendido como a capacidade de recuperar o que se busca e de evitar o que não é desejado, pode ser obtido através das chamadas medidas de eficácia, sendo as principais denominadas de índice de precisão ou pertinência e índice de revocação.

O índice de precisão ou pertinência para, Lancaster (1993, p.4), “[...] é a capacidade de evitar documentos inúteis”, ou seja, a recuperação dos que não fazem parte da pesquisa do usuário. Já o termo revocação designa, dentro de um SRI, a recuperação de itens que serão úteis ou que terão serventia para o pesquisador.

Guinchat e Menou (1994, p.325) afirmam que: “O índice de revocação é a proporção de informações relevantes ou pertinentes recuperados em relação ao total de documentos existentes no sistema”.

O índice de precisão, é a extensão com que os itens recuperados durante uma busca são considerados relevantes ou pertinentes. Uma busca que alcance uma precisão alta será aquela em que a maioria dos itens recuperados, se não todos, forem considerados relevantes ou pertinentes.

Foskett (1973, p.11) afirma que: “[...] o sistema deve ser planejado de modo a viabilizar, ao máximo, a coincidência entre as necessidades informacionais do usuário e aquilo que o autor oferece”. Obviamente, deve-se ter consciência que tal coincidência nem sempre será exata. Um bom planejamento irá minimizar a possibilidade de erro. Os erros repercutirão sobre a relevância, pois serão obtidas respostas erradas, também prejudicando a revocação, uma vez que escaparão itens de interesse do usuário. “É preciso, portanto, que nos asseguremos de que o sistema utilizado não possui uma tendência intrínseca a aumentar o erro humano” (FOSKETT, 1973, p.12).

Alguns fatores influem no desempenho e eficácia de um SRI e estão diretamente ligados a indexação, como a sua política, onde as decisões quanto a ela serão tomadas pelos administradores do serviço de informação. A principal deliberação diz respeito à exaustividade da indexação, que corresponde ao número máximo de termos atribuídos em média por um documento. A indexação exaustiva implica no emprego de termos em número suficiente para abranger o conteúdo temático do documento de modo bastante completo.

Intervêm também na qualidade do SRI os recursos tecnológicos disponíveis, uma vez que estes irão determinar desde o armazenamento até a estratégia de busca que será utilizada pelo usuário na hora da recuperação da informação, entre outros aspectos importantes.

Destaca-se, ainda, os atributos necessários ao bom indexador que são, dentre outros, o conhecimento do assunto, das necessidades dos usuários, experiência, concentração, capacidade e compreensão de leitura, domínio dos recursos de informática disponíveis e do vocabulário utilizado pelos usuários em suas buscas.

Os aspectos acima arrolados serão tratados com maior profundidade no Capítulo 2 deste trabalho.

A eficácia de um SRI está diretamente ligada a vários aspectos como exposto, porém faz-se necessário examinar preliminarmente, como a comunicação documentária faz parte desse processo.

## **1.2 COMUNICAÇÃO DOCUMENTÁRIA**

A Teoria da Informação, no seu significado abrangente, compreende também a comunicação, uma vez que não há informação fora de um sistema qualquer de sinais e de um veículo ou meio apto a transmiti-los. Comunicação significa partilha de elementos ou modos de vida e comportamento, por virtude da existência de um conjunto de normas.

Em suas origens, a Teoria da Informação surge como um ensinamento estatístico e matemático, tendo-se originado nos campos da telegrafia e da telefonia, especialmente com os trabalhos de Shanon e Weaver.

Na Teoria da Informação proposta por Shanon e Weaver, a comunicação é compreendida por um sistema (fonte) que influencia outro sistema (receptor) enviando signos alternados, transmitidos por um canal que os une (COYAUD, 1966 apud LIMA, 1998), isto é, o meio torna possível à comunicação e, conseqüentemente, a transmissão da informação, conforme pode ser observado na figura 1.

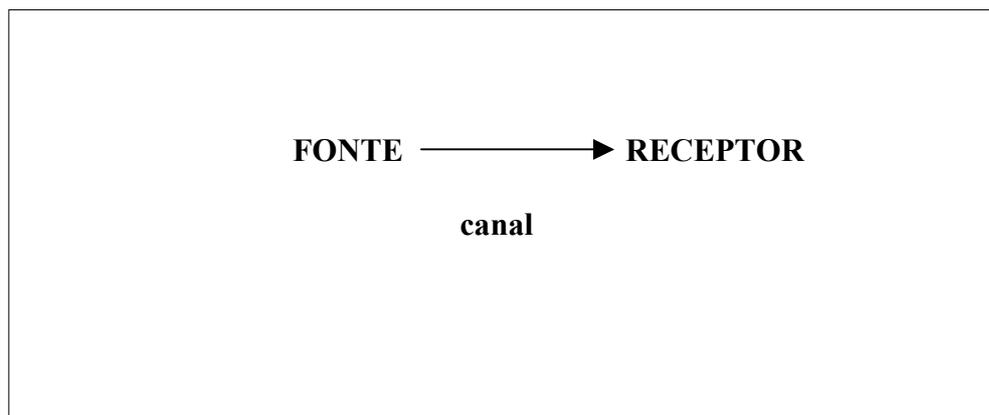


Figura 1 - Modelo de sistema de comunicação  
FONTE: Lima, 1998.

De acordo com Lopes (1996), para os teóricos da comunicação, a informação ou mensagem, deve ser entendida como conseqüência de sinais que correspondem a regras de combinação precisas, sendo que eles são transmitidos por um canal que atua entre emissores e receptores.

Os sinais enviados pela fonte ou emissor para o receptor são articulados através de um código, que deve ser comum tanto à fonte quanto ao receptor. Estes sinais constituem a mensagem através da qual o receptor recebe as informações produzidas pelo emissor, estabelecendo a comunicação entre os mesmos e possibilitando a transferência da informação.

A Teoria da Informação utiliza a frequência da ocorrência de sinais como critério para determinar o valor informacional da mensagem. Quantifica-se a informação tornando-a mais objetiva, entretanto, este critério, fácil para avaliar e processar, falha ao refletir a realidade, pois não responde ao critério de pertinência.

Eco (1984, p.11), também define o processo de comunicação como “[...] aquele onde um emissor (E) manda uma mensagem (M) através de um canal para um destinatário (D) utilizando-se de um código que deve ser comum, ao emissor e ao receptor [...]”, mas enfatiza na questão do código a ser utilizado, conforme mostra a figura 2.

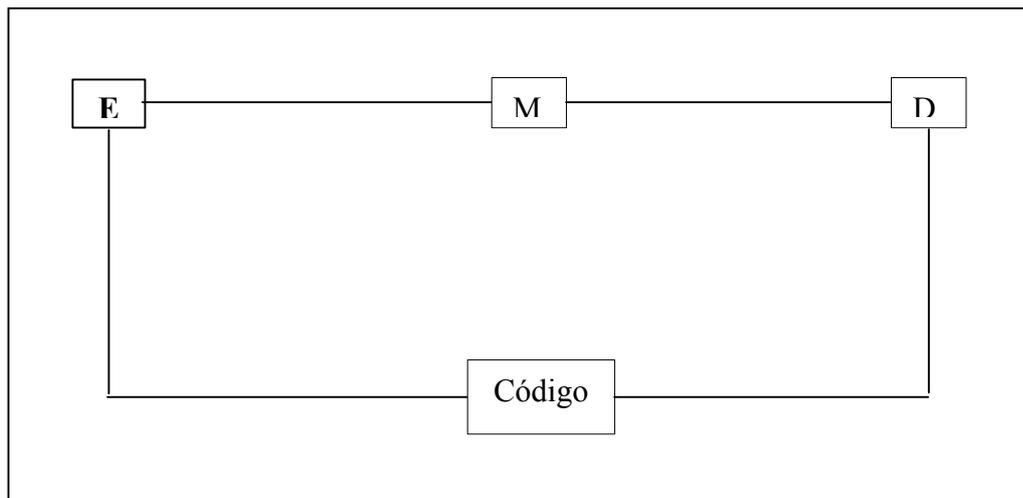


Figura 2 – Modelo de processo de comunicação  
 FONTE: Eco, 1984.

Na figura acima, o código é indicativo de um sistema de significação, o qual é construído e analisado independentemente de ser posteriormente utilizado em um processo de comunicação, ainda que não possa existir este sem aquele. Em outras palavras, a existência de um sistema de significação (código) é condição essencial para a efetivação do processo de comunicação.

Encontra-se na literatura muitos modelos que abordam o processo de transferência da informação entretanto, desde a proposta implementada por Shanon e Weaver até os dias atuais, houveram mudanças significativas, no modo de entendimento da questão.

Robredo (2003) aceita o princípio de que a informação é primeiramente mantida em uma fonte (emissor) e, mais tarde, alcança o entendimento do receptor (usuário) após passar por uma série de processos, conforme pode-se observar na figura 3.

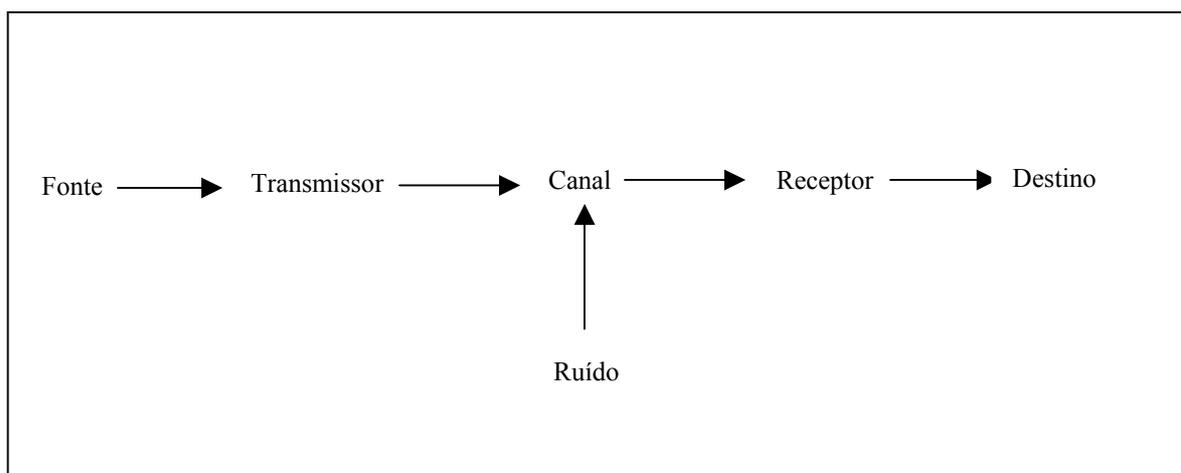


Figura 3 – Processo de comunicação num serviço de informação.  
FONTE: Robredo, 2003.

O modelo acima ilustra como a informação, ao ser comunicada, possui uma fonte e um destino distintos no tempo e no espaço, onde se origina a cadeia que os une e que constitui o canal de comunicação. Para que a mensagem transite por esse canal, necessário se torna reduzi-la a sinais aptos a essa transmissão, sendo esta operação chamada de codificação e quem ou o que a realiza é o transmissor. No ponto de destino, um receptor reconverte a informação à sua forma original, decodificando-a com vistas ao seu destinatário.

O processo de comunicação pode sofrer, em cada uma das etapas intermediárias, a interferência de ‘ruídos’, sinais irrelevantes e obstrutivos. Poucas mensagens são totalmente auto-suficientes, de maneira que parte da informação comum passa de fonte para o usuário por meio de outros canais.

Assim, para entender o processo de comunicação como também o que é signo e significação, é necessário fazer referência à linguagem por meio da qual o homem se comunica um com os outros no plano social e no interior da qual objetos e processos materiais podem, sob circunstâncias definidas, funcionar como signos, isto é, adquirir significados (SCHAFF, 1968). No processo tradicional de comunicação, a linguagem utilizada é a natural, isto é, aquela comum, que comporta uma grande quantidade de unidade, que é adaptada à comunicação oral ou escrita (GUINCHART; MENO, 1994).

A comunicação documentária se desenvolve em um sistema de informação, o qual geralmente está ou é desenvolvido por uma instituição com objetivo de atender aos usuários. Portanto, na procura de um modelo de comunicação documentária não se pode desprezar a avaliação de como se processa a comunicação especializada, isto é, como ela ocorre entre os especialistas de um domínio.

Segundo Cabré (1993, p.103) “[...] a comunicação especializada, processo de comunicação entre especialistas, contempla dois elementos que não estão explícitos no esquema de Shanon e Weaver”, que pode ser observado na figura 4.

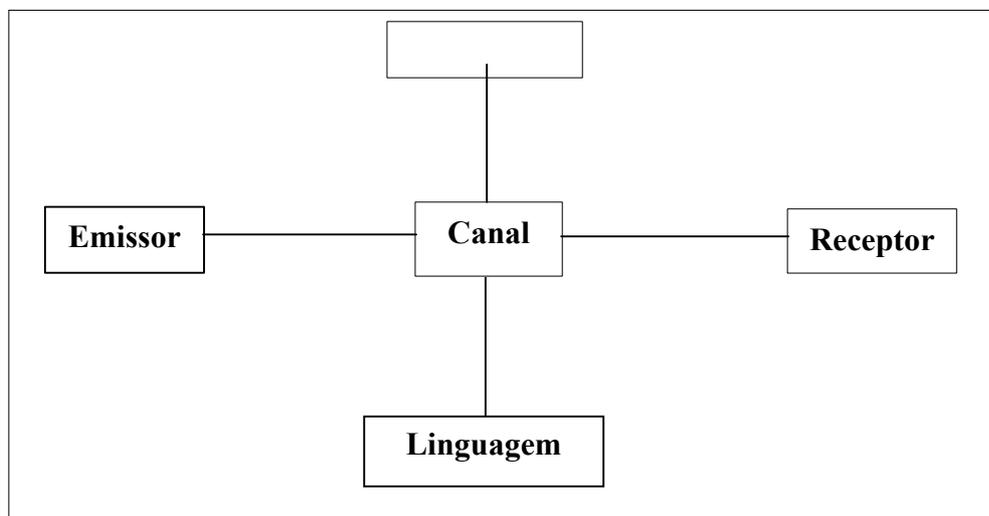


Figura 4 – Modelo de processo de comunicação  
FONTE: Cabré, 1993.

Neste esquema, observa-se a presença não só dos elementos tradicionais como a Fonte (Emissor), o Canal e o Receptor, mas também a Linguagem (Código) e a Realidade, porque na comunicação especializada a linguagem utilizada não é somente a natural, mas uma linguagem específica, a qual, ainda que seja parte da primeira, é constituída principalmente pela terminologia do domínio específico em questão.

Em contraste com o esquema tradicional, na comunicação especializada os interlocutores são especialistas, em maior ou menor grau de conhecimento sobre o assunto, e

o mundo de referência de sua comunicação se limita ao campo específico da especialidade, que se encontra mais formalmente conceituado (CABRÉ, 1993).

A terminologia, veículo de comunicação formal entre os especialistas de um domínio, segundo Cabré (1993, p.106), é o que diferencia a comunicação especializada da geral,

[...] pois os termos que a constituem designam conceitos próprios de cada disciplina específica, os quais são monoreferenciais, isto é, cada termo apresenta um conceito específico, que se relaciona com os demais através de sua definição, contribuindo para a precisão e normalização desta linguagem e conferindo univocidade à comunidade especializada.

A circulação da informação e do conhecimento dentro de um domínio específico é um dos atos comunicacionais da esfera humana. A comunicação documentária, no campo dos sistemas de informação, lida com atos de comunicação materialmente explícitos, isto é, o documento, sua representação, a pergunta do usuário e a ação comunicativa do profissional da informação.

Em 1978, Coelho Netto relatava sua insatisfação mediante ao tradicional modelo de comunicação que era retratado conforme se observa na figura 5.

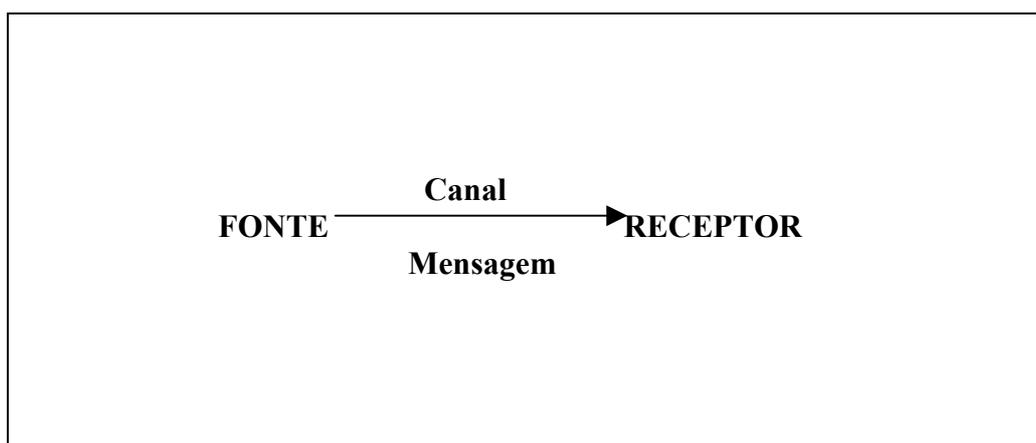


Figura 5 – Tradicional modelo de comunicação  
FONTE: Coelho Netto, 1978.

O tradicional modelo de comunicação, para Coelho Netto (1978), era inadequado devido a sua natureza paternalista, onde todo processo era orientado a partir da fonte para o receptor, ou seja, a mensagem era produzida pela fonte e para o receptor. A fonte surgia como

ponto de partida de todo o processo, produzindo informação a sua vontade e levando-a ao receptor, cujo único papel era o de receber essas informações produzidas longe de seu campo de ação, sem uma ação dialógica, promovendo a manipulação, pois o que era recebido da fonte não produzia *feedback*.

Ainda segundo Coelho Netto (1978), o modelo citado teria de ser desconstruído e reelaborado, fazendo com que o receptor não mais fosse manipulado pela fonte, mas como interlocutor que se desperta para uma certa necessidade de informação, manipulando a fonte de acordo com suas próprias necessidades.

Westley e MacLean (1972) citado por Coelho Netto (1978), propôs a reforma do modelo tradicional para um modelo conceitual da comunicação que já retirava da fonte a possibilidade das iniciativas.

No modelo proposto em um sistema de informação a comunicação entre o acervo (fonte), o usuário (receptor) e o profissional da informação (canal), pode ocorrer de dois modos na comunicação documentária, ou seja, direta e indiretamente, conforme se observa na figura 6.

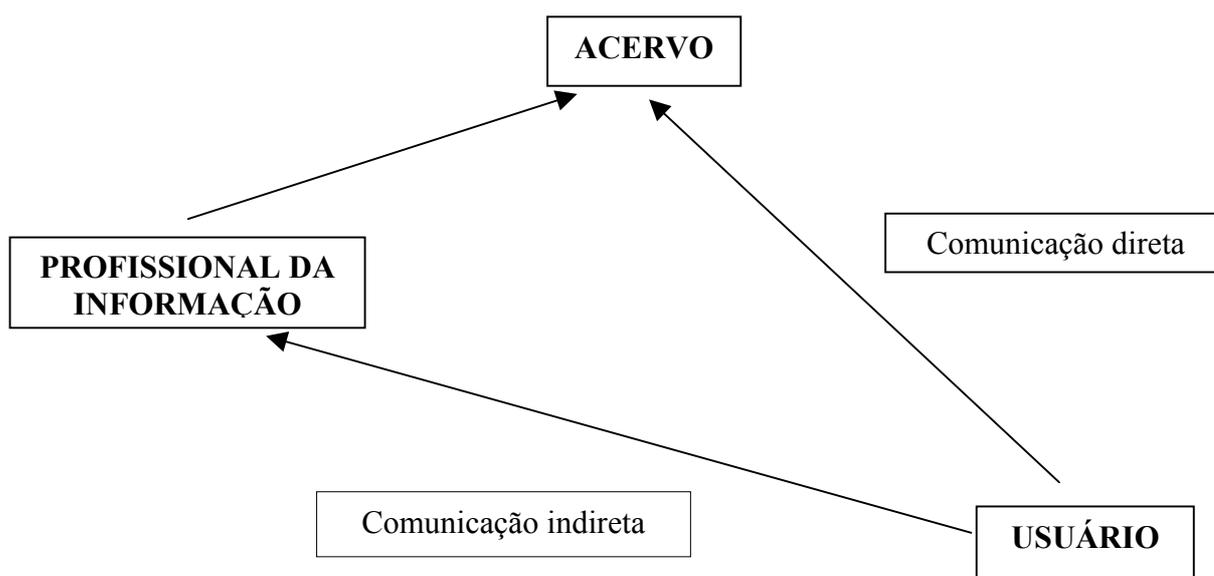


Figura 6 - Esquema de comunicação proposto por Coelho Netto.  
FONTE: Lima, 1998.

A figura acima mostra que ao procurar uma determinada informação no acervo (fonte), sem solicitar o auxílio do profissional da informação (canal), o usuário (receptor) está coletando diretamente da fonte aquilo que necessita mas, por outro lado, quando solicita a informação ao profissional (canal), o usuário (receptor) está acionando o acervo (fonte) indiretamente.

Diante dos modelos anteriormente discutidos, observa-se que, enquanto o esquema tradicional de comunicação, considera a interseção entre os códigos da fonte e do receptor, na abordagem proposta por Lima (1998) esta questão não é mencionada, pois a autora aponta o profissional da informação como canal na comunicação indireta e não inclui o código utilizado para representar, ou seja, codificar, o conteúdo informacional dos documentos e mediar a comunicação documentária possibilitando a transferência da informação, seja no acesso direto ou indireto à fonte.

Tanto na comunicação direta quanto indireta, o usuário ou o profissional da informação, no momento da busca por informação, se deparam com a linguagem documentária (LD), que é instrumento construído pelo sistema informacional com a função de representar a informação contida nos documentos de forma padronizada.

Portanto a LD, ao codificar a informação existente em um documento permiti que esta seja decodificada pelo usuário e se apresenta como um código, que será capaz de viabilizar o processo discutido por Lara (1993c, p.4) como “[...] comunicação documentária que se estabelece entre o acervo de um sistema de informação e o seu usuário”. Através de uma linguagem específica, a LD, os conteúdos informacionais de um documento são codificados pelo indexador no tratamento e decodificado pelo usuário na recuperação da informação. Isto se torna possível devido ao fato dela ser construída com a função de compatibilizar a linguagem dos documentos com a dos usuários.

Analisando o esquema de comunicação, exposto na figura 6, verifica-se que o mesmo, com o intuito de facilitar a recuperação da informação por parte do usuário, deveria ser complementado ou reestruturado com a inclusão da linguagem documentária como se observa na figura 7.

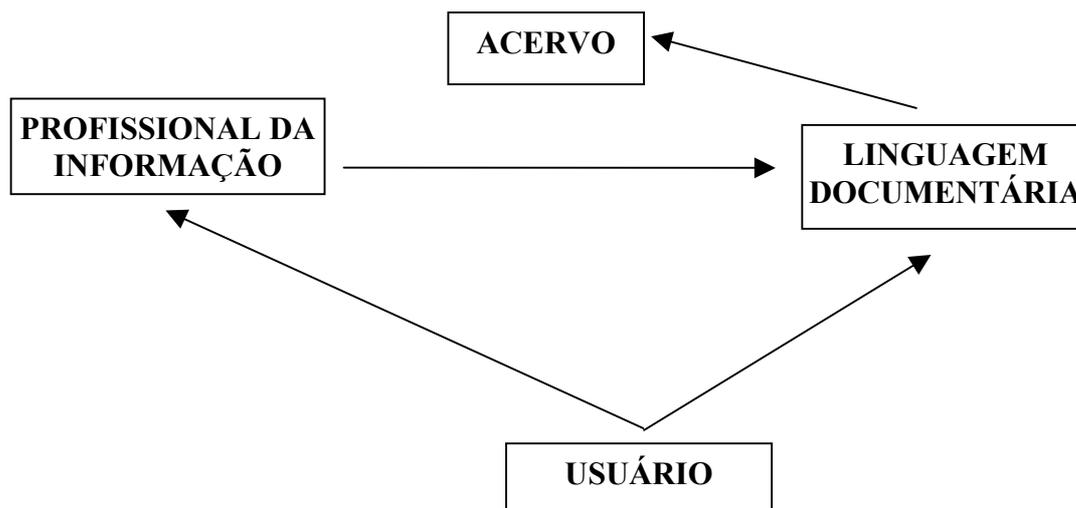


Figura 7 – Esquema de comunicação documentária estruturado.  
FONTE: Lima, 1998.

No esquema acima, considera-se como fonte o acervo ou sua representação, de forma organizada pelo sistema, que utiliza a LD para recuperar a informação pretendida pelo usuário. O código que possibilita a comunicação documentária é a LD que, após ser transformada em linguagem padronizada, vai ser acionada direta ou indiretamente, através do profissional da informação ou do usuário. Assim, verifica-se que sempre existe uma intermediação entre o usuário e o acervo quer seja somente pela LD ou pela associação do profissional da informação com a LD.

Como se observa, a competência do sujeito para integrar o circuito da produção de conhecimento só será efetivada a partir de uma ação comunicativa entre a fonte (estoques de informação) e o receptor (usuários), pois a transferência da informação é condição essencial para a produção do conhecimento e, para que ela efetivamente aconteça, é primordial que as informações sejam reunidas, armazenadas e representadas de acordo com uma organização

que possibilite a sua recuperação de maneira eficiente por aqueles que delas necessitam, sendo este o papel do sistema de informação.

Assim, para cumprir com o seu objetivo, um sistema de informação deve dispor de uma LD, pois sem este instrumento de representação, não se efetiva a comunicação documentária. É neste contexto que discute-se, a seguir, a terminologia, um dos elementos envolvidos no processo da comunicação documentária.

### 1.3 TERMINOLOGIA

A terminologia nasce sob a influência de duas tendências: a interdisciplinaridade das ciências e as suas micro-especializações. A primeira surge em consequência da necessidade de normalização terminológica pois a univocidade, ou seja, a designação dos termos, é condição essencial para a comunicação entre especialistas, seja no interior de uma mesma disciplina ou entre disciplinas diferentes. Já a segunda acentua a criação de neologismos, isto é, termos novos ou antigos com nova significação como por exemplo, o termo vírus, originalmente da área médica e que foi assimilado por empréstimo pela área da informática (RONDEAU, 1984, apud LIMA, 1998).

Os primeiros trabalhos terminológicos surgem em 1906 através da *Comission Electrotechnique Internationale*, mas é sob a influência de Eugene Wuster, fundador da Teoria Geral da Terminologia – TGT, eles começam a ser sistematizados, graças a definição de postulados fundamentados no desenvolvimento de métodos de trabalho. A terminologia de Wuster é uma ferramenta capaz de eliminar as ambigüidades dentro das comunicações científicas e técnicas.

O surgimento da *International Federation National Standardization Association - ISA* ocorreu pela busca de desenvolver a normalização dos produtos e procedimentos. Sob a influência de Wuster, a ISA constituiu um comitê, em 1936, para elaborar os princípios

metodológicos destinados a assegurar a harmonia das terminologias e dos modos de preparação e apresentação. Em 1946 surgem a *International Organization for Standardization - ISO*, responsável, até hoje, por rever todas as normas de produção e, por conseqüência, colocar em uso os mecanismos de normalização terminológica, definindo terminologia como qualquer atividade relacionada com a sistematização e apresentação de conceitos ou de termos baseados em princípios e métodos estabelecidos e ainda como um conjunto de termos que constituem um sistema de conceitos de uma determinada área.

O que se observa na literatura, é que existe uma diversidade grande em relação a definição de terminologia. Cabré (1991) aponta, como motivo para essa variedade, seus enfoques e suas aplicações práticas, além da conhecida polissemia do termo, que pode remeter a três noções: a) uma disciplina, b) uma prática; c) produto gerado por essa prática. Como disciplina, é uma matéria que se ocupa dos termos especializados; como prática, é um conjunto de métodos e atividades voltado para coleta, descrição, processamento e apresentação de termos; como produto, é um conjunto de termos de uma determinada especialidade.

Até mesmo quando tratada apenas sob uma dessas perspectivas, a terminologia apresenta significados diversos. Se tomada como objeto, isto é, como um conjunto de termos de uma especialidade, nota-se que cada área do conhecimento aborda seus termos de forma diversa.

Cabré (1991) ressalta três diferentes concepções sobre os termos. Para a lingüística, os termos são conjuntos de signos lingüísticos que constituem um subconjunto dentro do componente léxico da gramática, são uma forma de saber. Para a filosofia, as terminologias são unidades cognitivas que representam o conhecimento especializado. E, por fim, para as diferentes disciplinas técnico-científicas, são unidades de expressão e comunicação que permitem transferir o pensamento especializado.

Na concepção da terminologia como disciplina, também existem correntes divergentes apontadas na literatura. Alguns autores, adeptos da Teoria Geral da Terminologia, defendem que essa matéria é autônoma e auto-suficiente, dotada de fundamentos próprios, apesar de conectada historicamente a outras disciplinas.

A terminologia é definida como um conjunto de práticas que evolui no contexto da criação de termos, de sua coleta, explicação e apresentação em diferentes meios impressos e eletrônicos.

Assim como a lingüística, a semântica ou a semiótica, a terminologia estuda os signos que se manifestam através da linguagem natural e se ocupa dos conjuntos estruturados de termos agrupados em classes por critérios que estão presentes nas suas definições.

Neste trabalho, interessa entender a terminologia enquanto conjunto de termos de uma área específica do conhecimento, isto é, como objeto, onde eles estão relacionados e definidos rigorosamente para designar os conceitos que lhes são úteis. A definição terminológica deve exprimir os traços pertinentes do termo, refletir o conceito e só pode ser aplicada a um termo puro, perfeitamente traduzível em todas as línguas e a um sistema terminológico perfeitamente coerente onde cada conceito é construído de maneira unívoca e explícita. O objetivo da terminologia, segundo Cintra et al (1994, p.27), é: “[...] identificar em uma área do conhecimento, primeiramente os conceitos, depois lhes atribuir um termo controlando, ao mesmo tempo, suas relações de sinonímia e homonímia e dar a cada conceito uma definição rigorosa”, ou seja, depois de identificar o conceito relacionado ao conteúdo do documento, traduzir para uma linguagem controlada que seja equivalente.

Ao trabalhar com expressões de classes de objetos tomados em um universo determinado, isto é, de acordo com os pontos de vista próprios de cada área, a terminologia vai constituir predicados relacionados, ou seja, termos referenciais e unívocos, pois uma

palavra designa um determinado objeto, na medida em que o insere em uma classe particular dentro de um contexto (LARA, 1993c).

A terminologia difere do léxico comum porque designa os conceitos próprios das disciplinas e das atividades de especialidade. Os textos especializados se diferenciam dos gerais por utilizarem vocabulários relativos às áreas de especialidade. A terminologia confere a eles características particulares: os termos pertencem a domínios concretos e são monoreferenciais, o que oferece uma precisão e univocidade a comunicação (CABRÉ, 1994). A terminologia é encontrada na documentação especializada em estado natural e depois de codificada aparece nos dicionários de área. Para Cabré (1994), o objetivo prioritário a que se propõe a terminologia é a normalização, dos termos próprios de um domínio especializado, com a finalidade de fixar as unidades terminológicas com formas normalizadas, estabelecendo as referências e separando as outras variantes para denominar o mesmo conceito.

Em sua função de representação, a terminologia serve a documentação pois é um elemento chave para representar o conteúdo dos documentos e para chegar até eles, pois a capacidade de uma LD em recuperar as informações presente nos textos, está condicionada à sua habilidade de referir. Para que essa referência exista concretamente, a construção das linguagens documentárias deve ser feita com base na terminologia de área, em campos bastante delimitados, pois como elo remete a sistemas de significação de área, se constitui em base legal para a construção das LDs. A terminologia pode tornar a linguagem documentária interpretável, possibilitando não apenas a restituição de informações dos documentos do sistema, mas o estabelecimento de novas relações; não só correlações lexicais, mas instruções para o desencadeamento dos processos de inferência contextual.

Como o usuário, na maioria das vezes, não conta com uma LD que possua um sistema de referência explícito, isto é, da terminologia de área, a interpretação fica praticamente livre

e indeterminada, impossibilitando a univocidade conceitual, o que induz à subjetividade e à ambigüidade comprometendo, conseqüentemente, a circulação de informações (LARA, 1993c).

Para que ocorra transferência de informação, é necessário que o sistema de significação utilizado seja claro e explícito. A ausência de definição nas LDs, de delimitação do contexto a que se refere, isto é, do referente, mais a artificialidade do sistema acabam por dificultar esta transferência, comprometendo a comunicação documentária. Trabalhar-se-á a seguir as questões inerentes a terminologia sob o aspecto da LD.

#### **1.4 TERMINOLOGIA COMO INSTRUMENTO DA LD**

Como visto, a LD, que tem por função representar o conteúdo informacional de um documento através de um código utilizado no processo de comunicação documentária para a transferência da informação, necessita de um instrumento de referência para sua construção e uso, tendo em vista atender as reais necessidades do usuário.

Após o estabelecimento do modelo de comunicação documentária, exposto na figura 7 e, analisando o código da LD bem como a questão da representação documentária, assim como verificado os problemas decorrentes da não existência de um referente quando da construção e uso de uma ou várias LDs, busca-se, a partir deste momento, construir um modelo teórico para definir o lugar e a função da terminologia no processo de representação documentária, demonstrando como ele pode ser operacionalizado para solucionar a diversidade de representações para efetivar, assim, a comunicação documentária e, conseqüentemente, a transmissão da informação e a geração de novos saberes.

Da mesma forma que no processo tradicional de comunicação, observa-se que na comunicação documentária o receptor recebe informações da fonte através da utilização de um código comum. A fonte é o acervo, o receptor, o usuário e o código a linguagem

documentária, sendo que a transferência da informação só se efetiva quando o receptor conhece o código utilizado pela fonte, entra em sintonia com ele e assimila aquilo que lhe foi repassado.

Um sistema de comunicação não existe sem um repertório de signos e um corpo de regras que definam como eles selecionam e se combinam para formar as mensagens transmissíveis, portanto a comunicação documentária só se realiza através de signos.

No caso de uma linguagem natural Véron (1970, p.172) afirma que,

[...] o código é aquilo que tradicionalmente se chamava 'gramática', isto é, as regras que todo usuário competente dessa linguagem conhece e aplica quando emite uma mensagem verbal qualquer. Nas ciências, a linguagem natural se especializa como universo de discurso referido a certo setor da realidade e o uso especializado de uma linguagem natural, isto é, o uso científico, exige a introdução de regras de construção adicionais, tanto sintáticas quanto semânticas em um esforço para reduzir na medida do possível o alto grau de ambigüidade característico do uso espontâneo da língua.

De fato, este é o papel das terminologias de área. Assim, se é na linguagem natural que a linguagem documentária vai buscar as relações semânticas para adaptá-las às suas necessidades, é na terminologia que ela deve procurar a produção do sentido referente a um determinado contexto.

Ademais, a comunicação documentária só ocorre quando os signos utilizados do acervo (fonte) para representar a informação contida em seus documentos, são de conhecimento de seu usuário (receptor). Contudo, qual seria a natureza destes signos que, conforme Lara (1993c), podem também ser denominados de signos documentários? Antes de oferecer um maior aprofundamento neste assunto, no próximo item abordar-se-á o signo lingüístico, visando ampliar o entendimento da temática em estudo e permitir uma compreensão dinâmica da relação que se estabelece entre ambos.

## 1.5 SIGNO LINGUÍSTICO

O principal pilar da teoria Semiótica é a noção de signo como entidade que permeia toda a existência humana, de forma a estabelecer um elo comunicacional entre as coisas do mundo, extrapolando assim a esfera exclusiva da comunicação.

A palavra signo, através do latim *signum*, vem do étimo grego *secnom*, raiz do verbo “cortar”, “extrair uma parte de”, de onde derivou em português, por exemplo, secção, seccionar, sectário, seita, etc. Do derivado latino, são numerosas e expressivas as palavras que se compuseram, a partir dele na língua portuguesa, como sinal, senha, sineta, etc.

De qualquer forma, convém reter a idéia de signo enquanto alguma coisa que substitui outra, de modo a desencadear um complexo análogo de reações. Ademais, pode-se, para adotar a definição do fundador da Semiótica Charles Peirce (1839-1914), compreender que o signo, ou *representamen* é toda coisa que substitui outra, representando-a para alguém, sob certos aspectos e em certa medida.

Quando se fala sobre a natureza dos signos, muitas são as teorias que discutem tal questão. A semiótica (do grego *semeiotiké*, (arte) dos sinais, sintomas) é a ciência dos signos ou seja, do processo de significação na natureza e na cultura. Essa ciência é relativamente nova e teve como maiores expoentes Charles S. Peirce e o suíço Ferdinand de Saussure. Na Semiótica de Peirce, o signo tem como característica básica a de poder representar as coisas ou objetos, isto é, um signo ou *representamen*, é o sustentáculo dele próprio ou aquilo que funciona como signo, enviando-o para um interpretante. É através dele que o signo se remete por alguma causa (seja a semelhança, indicação ou convenção) a um objeto. Já Saussure define o signo como a união do sentido e da imagem acústica. O que ele chama de “sentido” é a mesma coisa que *conceito* ou *idéia*, isto é, a representação mental de um objeto ou da realidade social em que o homem se situa. Em outras palavras, para Saussure, conceito é sinônimo de significado (plano das idéias), em oposição ao significante (plano da expressão).

Dessa forma, um signo consiste em:

- um conceito – ou seja, o *significado* (*signifié*)
- uma imagem acústica – ou seja, o *significante* (*signifiant*)
- um referente – a coisa a que se refere ou designa.

Para que um signo exista é preciso que haja, no mínimo, o reconhecimento da possibilidade de construção de um significado. Para Azevedo Netto (2005, p.3), “[...] esta entidade (signo), antes de estar no lugar de alguma coisa ou mesmo de representá-la, é um processo de encadeamento, que leva à construção de um outro signo que se relaciona com o primeiro, e assim por diante em uma constante semiose”. O procedimento citado pelo autor deve-se à identificação de uma das entidades que compõe o signo, o “interpretante”, que possui em si o atributo de produção de significado, pela semiose e, com isto, de novos signos. O autor esclarece que a entidade “signo” é composta pelo “[...] signo-objeto, signo-veículo e signo-interpretante”, que compõem a entidade signo, estas, ao mesmo tempo, possibilitam um tratamento diferenciado e independente de cada uma delas. Assim, as três esferas que compõem o signo podem e, em alguns casos devem, ser tratadas como entidades em si.

Embora a expressão peirceana *interpretant* seja usualmente traduzida por “intérprete”, convém esclarecer que interpretante não designa tão somente o intérprete ou usuário do signo, mas antes uma espécie de supersigno ou supercódigo, individual ou coletivo, que reelabora constantemente o seu repertório de signos em confronto com a experiência, conferindo aos signos, em última instância, o seu significado real, prático. O interpretante, assim, não é uma “coisa”, mas antes o processo relacional pelo qual os signos são absorvidos, utilizados e criados.

De fato, o signo é a reunião de várias coisas, em que ocorre uma relação triádica entre o objeto, veículo e interpretante, dentro da construção da significação e do processo de comunicação. O signo-objeto se origina da percepção, que remete a um signo-veículo, fruto

da transmissão de uma idéia, que se instala em um signo-interpretante resultante das semioses de construção do significado e que remete ao seu objeto, criando assim uma cadeia de relações entre as esferas do signo, ou seja, o mesmo objeto produz determinado signo no transmissor, emitido em forma de um segundo até o receptor, que produz um terceiro, com referência ao mesmo objeto.

O signo não é, portanto, uma entidade física e palpável, mas abstrata, existente na consciência daqueles que compartilham uma mesma cultura, já que prescinde de uma característica relacional entre os transmissores e receptores destas entidades. Ao mesmo tempo em que é necessária a relação entre os envolvidos em um processo de comunicação para a existência do signo, é, igualmente, necessário que as instâncias de transmissão e significação sejam relacionadas.

Tanto o é que, em uma relação triádica genuína, não só o signo, mas também o objeto, assim como os interpretantes são todos de natureza sgnica. Isto implica afirma que todos os três correlatos são signos, sendo que aquilo que os diferencia é o papel lógico desempenhado por todos eles.

Então o signo pode ser entendido como entidade que permeia a vida humana, seja no processo de comunicação, na construção do conhecimento, em uma relação triádica entre suas três esferas (“signo-objeto”; “signo-veículo”; e “signo-interpretante”).

Como exposto, o signo tem como peculiaridade a capacidade de se reproduzir ao infinito em novos signos estando, portanto, ligado a um processo de comunicação pela sua própria necessidade de reprodução através da interlocução, na qual um signo produz outros, pela semiose, resultando no processo que Eco (1980) chama de semiose ilimitada. Então a interlocução, que é característica da comunicação, é condição primordial para a existência do signo, já que este somente existe enquanto age sobre o interpretante.

A representação gráfica do signo é feita pelo conhecido triângulo de Ogden e Richards.

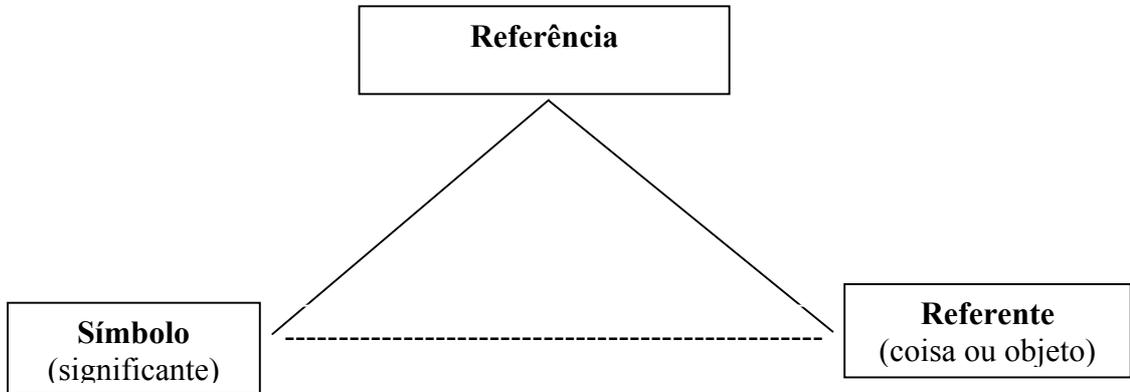


Figura 8 – Triângulo de Ogden e Richards  
 FONTE: Pignatari, 2002.

O esquema acima, constitui uma simplificação da relação triádica estabelecida por Pierce para os signos em geral: a idéia de “referência” ou “pensamento” absorve as idéias de “interpretante” e “base”. Ogden e Richards o estabeleceram principalmente tendo em vista o problema do significado nos signos verbais: a linha pontilhada indica que não há ligação direta entre signo e o referente, ou melhor, que a relação é apenas convencional e que só adquire significado em função do intérprete.

Blikstein (1983) apud Lima (1998, p. 94), efetua uma análise profunda dos elementos que constituem o signo lingüístico que podem se estender ao signo documentário. Segundo este autor, “[...] o signo representaria a realidade extralingüística e em princípio é por meio dele que podemos conhecê-la”. Ao procurar compreender os mecanismos de transformação do referente é que Blikstein desenvolve o seu modelo de signo lingüístico.

Pode-se, desta forma, inferir que o indivíduo utiliza a linguagem para representar a realidade através de um referente fabricado de acordo com a práxis social de que participa e que, sem ela, não há significação. É possível utilizar este mesmo raciocínio para observar a questão da função da terminologia na representação documentária, pois o sistema de informação se utiliza da LD para representar a informação através da terminologia a qual é

fabricada de acordo com a práxis social dos membros de uma comunidade específica especializada.

A seguir aborda-se o signo documentário, mais especificamente voltado para a Análise Documentária, uma vez que a exposição anterior examinou e ofereceu suporte para as discursões que serão efetuadas.

## **1.6 SIGNO DOCUMENTÁRIO**

Como visto anteriormente, a representação da informação resulta de um conjunto de procedimentos denominado Análise Documentária (AD), cuja metodologia permite expressar o conteúdo dos documentos sob formas distintas, facilitando assim sua recuperação. Um dos instrumentos utilizados pela AD é a Linguagem Documentária (LD) que tem por função, a normalização das unidades significantes ou conceituais presentes no texto original como meio de viabilizar sua comunicação. Ao transformar a informação presente nos textos em linguagem natural e/ou especializada para uma LD controlada, a AD opera com a significação, isto é, procura representar adequadamente as informações sem comprometer o seu significado.

Segundo Lara (1993c, p.37),

[...] a LD confere às expressões de representação documentária o caráter de signos documentários. A LD seria então um conjunto de signos documentários utilizados para representar a informação contida nos documentos tendo em vista a sua recuperação pelo usuário do sistema de informação através do processo de comunicação documentária.

Da mesma forma que na linguagem natural, as palavras isoladas não significam nada ou, ao mesmo tempo, podem significar tudo dependendo da referência a determinados contextos os quais lhes confere significado. Na LD os signos documentários também devem se referir a um contexto determinado para que se estabeleçam as articulações necessárias ao

engendramento dos significados. Conforme Lara (1993a, p.223) diversos fatores contribuem para isso, tais como:

- a concepção de representação suposta na construção da LD;
- a correspondência entre o sistema de significação do domínio em questão (a terminologia) e a LD;
- os princípios estruturais que embasam a construção da LD;
- a adequação de seu uso por parte dos documentalistas;
- a adequação da LD ao perfil do grupo de usuários em questão.

Assim, pode-se dizer que ao representar-se a informação contida nos documentos, isto é, ao transformar os saberes registrados em informação, estar-se criando a informação documentária, que é construída a partir de signos documentários. Este signo documentário, no sentido semiótico, deve permitir a recuperação não apenas das semelhanças e equivalências entre palavras, mas estabelecer o elo de relações possíveis desencadeadas a partir de determinada palavra, em uma dada área do conhecimento, em um determinado contexto e circunstância (LARA, 1993c).

A semiótica objetiva definir o signo e distinguir os diferentes elementos presentes quando qualquer coisa funciona como signo, analisando as relações entre estes elementos e oferecendo uma classificação e uma tipologia dos signos.

Desta maneira, segundo Lara (1993a, p.225), a análise documentária se apropria deste conjunto de reflexões para pensar, teórica e operacionalmente, a questão da construção e recepção do signo documentário, orientando a sua delimitação e definição enquanto um referencial e funcionalidade.

Se de Saussure a AD retira o conceito da arbitrariedade do signo, o qual permite compreender que as LDs são essencialmente fundamentadas institucionalmente e refletem as convenções da língua e do paradigma ideológico que sustenta as suas definições e a sua organização, de Peirce ela se apropria do conceito de semiose, processo pelo qual de interpretante em interpretante, enriquece-se cada vez mais de propriedades o objeto. Isto é determinante da natureza referencial da semiose documentária, a qual não pode funcionar

aleatoriamente, devendo estar fundamentada em referenciais terminológicos os quais, por sua vez, remetem à estruturas conceituais dos domínios de especialidades.

A pragmática peirciana caracteriza também o funcionamento do signo documentário como necessariamente vinculado a contextos determinados, prevendo, portanto, significados funcionais. Isto é, o signo documentário é de controle do significado que só pode funcionar como elemento de representação da informação e possibilitar a semiose documentária desde que o contexto no qual se insere seja considerado.

Para Peirce, a interpretação do signo deve ser determinada pelo seu objeto através de uma observação paralela totalmente independente da ação do signo, caso contrário ela não será determinada a pensar nesse objeto.

Para a AD, essa experiência paralela pode ser observada em dois momentos distintos: o da elaboração da LD e o de sua utilização por parte do usuário, mas é insuficiente para consolidar a veiculação da significação nos diversos domínios de especialidade quando da elaboração da LD e dependente da explicitação do contexto e da existência de uma fundamentação no campo conceitual do domínio específico a ser considerado (LARA, 1993a).

Essa fundamentação, segundo Lara (1993a, p.226), “[...] só pode estar nas terminologias de especialidade, e é por essa razão que elas devem ser integradas à construção de tais instrumentos mediadores”. A terminologia funciona, neste contexto, como fonte de significação uma vez que pode fornecer o “conhecimento paralelo” correspondente do domínio de especialidade.

Assim, a representação documentária que procura refletir o conteúdo e a organização de um texto deve, necessariamente, prever como o usuário procura a mensagem e, para isto, deve buscar referenciais externos como as terminologias das áreas de especialidade, além de se reportar ao sistema de significação que lhe dá origem.

Pode-se dizer que a representação documentária denominada LD, a qual é constituída por um conjunto de descritores, é a reprodução da informação contida nos documentos através de seu conceito/significado utilizando como referente/objeto o conjunto de termos da terminologia da área do conhecimento.

Blikstein (1983, apud LIMA, 1998, p.105) “[...] considera que é na prática social ou práxis que reside o mecanismo gerador do sistema perceptual que a seu turno vai fabricar o referente”. Assim, no processo de representação documentária, dentro de determinado contexto, a prática social vai estabelecer uma determinada terminologia (termos), portanto um referente para os significados (conceitos) e significantes (descritores), os quais constituem o signo documentário.

É necessário destacar que o processo de SRI é permeado por elementos específicos, como a LD, que demandam pela compreensão do que seja a indexação, o que será abordado no próximo capítulo.

## **CAPÍTULO II**

Este capítulo busca aprofundar a questão do processo de indexação, partindo do seu processo evolutivo e sua inserção no ciclo documentário, bem como discute a importância da existência de uma política organizacional que favoreça o trabalho dos indexadores quanto ao trato das linguagens naturais e artificiais, visando a construção de instrumentos de controle terminológicos como os tesouros, de forma segura e dinâmica. Contrapõe a linguagem natural com a controlada de modo a favorecer a compreensão dos aspectos conceituais dos tesouros e, sobretudo, daqueles relacionados a biotecnologia.

### **INDEXAÇÃO**

A abundante geração de saberes se consolida pela disponibilidade de uma quantidade elevada de informação, recurso que tem sido considerado como o capital mais importante da sociedade contemporânea. No meio em que se vive, a informação tem um papel fundamental e lugar importante na vida cotidiana dos cidadãos, que precisam estar bem informados para que possam exercer dignamente a sua cidadania.

A indexação é a atividade que se ocupa em estabelecer a representação do conhecimento registrado e compõe um sistema global: o SRI, o qual é constituído por um conjunto de atividades que contemplam desde o processo de seleção e aquisição até a recuperação da informação.

O objetivo da indexação é estabelecer mecanismos para expressar, de maneira o mais fiel possível, a representação dos elementos que pertencem ao conteúdo de um documento, independente do suporte, material ou eletrônico, em que se encontre – a fim de que o mesmo possa ser recuperado posteriormente.

Entretanto, mesmo com a existência das tecnologias para auxiliar o ser humano na recuperação da informação, o usuário necessita estar atualizado para não ficar a mercê das

mudanças que se processam a todo momento e se por a frente das dificuldades naturalmente encontradas para acessar tais informações. Kuramoto (1999) apud Pinto (2001, p. 223-234) afirma que: “[...] para navegar sobre o espaço do saber é preciso conhecer as estratégias de busca e recuperação da informação, ou ainda, ter a ajuda de um especialista neste domínio para que as necessidades de informação dos usuários possam ser atendidas de maneira eficaz”. Nesta perspectiva, torna-se imprescindível saber o que fazer, como utilizar as novas tecnologias e quais as conseqüências de sua má utilização, o que pode, certamente, comprometer o que se busca (produto/serviço), quanto seus possíveis usos.

Além dessas dificuldades, deve-se considerar, ainda, a ausência e/ou deficiência no processo de representação em nível de indexação. Esse fato é notório onde quer que se busque informação, pois quase sempre tal busca resulta na recuperação de uma ampla quantidade de dados que, na maioria das vezes, não condizem com o que foi demandado, frustrando àquele que à procura. Para uma melhor compreensão da temática indexação, a seguir será destacado seu processo evolutivo visando permitir uma melhor contextualização no estudo.

## **2.1 PROCESSO EVOLUTIVO**

Estudiosos e filósofos, desde a Antiguidade, procuraram dividir e inter-relacionar idéias representadas por documentos de forma a permitir sua fácil localização através da identificação de seus conteúdos, representação temática, como critério para validação de sua introdução no campo do conhecimento. O essencial é examinar como os mesmos textos podem ser diversamente apreendidos, manejados e compreendidos.

A indexação, ao longo da história, para o mundo ocidental a partir da Idade Média, está irremediavelmente ligada a história do livro e das bibliotecas. Chartier (1998) refere-se a este contexto como a busca pelo ordenamento do mundo escrito.

Não é papel desta pesquisa relatar a história da indexação mas, em alguns momentos, pontuá-la no tempo e no espaço para traçar considerações importantes sobre seus métodos, sua função e funcionalidade, o que vem a ser uma prática de ordenamento e de sua permanência através dos tempos, como essencial para a formulação do conhecimento e compreensão de sua evolução.

A indexação trata de uma questão fundamental, cada vez mais necessária no contexto informacional. Muito além do tradicional ponto de vista de que o volume de informações aumentou descontroladamente, mas sim da significativa questão de que indexar se tornou uma chave para participar do conhecimento, portanto de sua produção, a indexação assume o papel principal no processo de disponibilização de saberes.

Compreende-se então que a indexação, em uma acepção ampla, é o processo de identificação e organização dos itens necessários à posterior recuperação da informação contida em um documento, destacando-se que os principais pontos de acesso para tal recuperação da informação são o autor, título e assunto, de um determinado registro.

Neste estudo considerar-se-á a indexação como o processo de determinação de assunto de um documento e a sua tradução para uma linguagem de indexação, um vez que o objeto desta pesquisa são as unidades indexadoras utilizadas para representar/recuperar as informações contidas nos documentos.

Partindo dessa concepção, entende-se que indexação é o processo pelo qual é determinado o tema principal ou assunto, e os subtemas ou assuntos secundários, tratados em um documento e, posteriormente, traduzidos para uma linguagem de indexação, que é um sistema de signos estruturados, cuja finalidade é representar e recuperar as informações registradas nos documentos.

Conforme Gutiérrez (1989, p.147), “La indexación es um ejercicio hermenêutico inseparable de la capacidad humana, de la cognición e inserto en la dialéctica que da vida al

binômio, documentalmente interativo, homem-cultura”<sup>2</sup>. Ainda Gutiérrez (1989, p. 152) segue afirmando que: “En efecto, no debemos olvidar que la comunicación documental es una comunicación mediatizada y convencionalizada tanto en el sentido tecnológico como ideológico”<sup>3</sup>. Na realidade, o autor vem reforçar o papel do indexador, que é aquele que ativa a reação e a resposta em um processo geral de comunicação documentária.

Tendo em vista o exposto, faz-se necessário compreender o modo como a indexação se coloca no ciclo documentário.

## **2.2 INDEXAÇÃO NO CICLO DOCUMENTÁRIO**

Como exposto anteriormente, um SRI compõe-se de um conjunto de fatores que, através de operações consecutivas, busca localizar, na totalidade das informações, àquelas realmente relevantes para o usuário.

Os estudos de determinadas cadeias informacionais referem-se à expressão do conhecimento científico, registrado em algum tipo de suporte, que denomina-se documento. A proposta de um SRI, incluso em uma cadeia de sucessivos relacionamentos, é recuperar informações através de pontos demonstrados como comuns, em uma ordem de possíveis fatores estabilizados que, nas escolhas oferecidas, propiciam uma indicação física de localização espacial, virtual ou mesmo temporal do documento propriamente dito.

Independente das definições de tipologia, ideologia, filosofia, ou mesmo concepções tecnológica, existem operações básicas, comuns, que teoricamente constituem-se na delimitação desta abordagem, permitindo aos sistemas responderem as necessidades informacionais posta em questionamento. Este processo, onde as fases e operações ganham identidade, denomina-se de ciclo documentário.

---

<sup>2</sup> A indexação é um exercício hermenêutico inseparável da capacidade humana, da cognição e inserta na dialética que dá vida ao binômio, documentalmente interativo, homem-cultura (tradução nossa).

<sup>3</sup> Com efeito, não devemos duvidar que a comunicação documental é uma comunicação mediatizada e convencionalizada tanto em sentido tecnológico como ideológico (tradução nossa).

O ciclo documentário, na visão da Biblioteconomia, traduz o detalhamento técnico e operacional da informação e do documento, permitindo uma determinada leitura de seu conteúdo, em um contexto formalizado por regras pré-estabelecidas.

Lancaster (1993) chama as etapas de um ciclo documentário de funções técnicas, que podem ser expressas em cinco pontos fundamentais: seleção e aquisição; descrição e indexação; organização e manutenção; estratégia de busca e interação *versus* usuário.

O ciclo documentário representa a organização, em fluxo, do processo de disponibilização de informação, através da relação do documento no seu suporte e do conteúdo possível de interpretação.

As etapas do ciclo documentário, além de suas características técnicas gerais, estão em um processo de sujeição à uma política institucional, a ser definida posteriormente. Isto significa que as possibilidades de processamento de um documento podem ser múltiplas ou não, pois deverão estar ligadas somente a um domínio do conhecimento.

A indexação, que se situa no meio do ciclo documentário, define a partir de sua localização, o início de um novo processo intelectual, uma ruptura no ciclo documentário, como um ponto chave para a determinação dos possíveis produtos oferecidos no sistema. Produz, em sua prática, uma nova etapa de direcionamento implicado na própria relação usuário/pesquisador e se coloca, paradoxalmente, como o principal elo da cadeia e o ponto de vulnerabilidade sistêmica. Ela será responsável pelo perfeito funcionamento do sistema, em lugar de seus propósitos imediatos ou possibilitará uma busca para outros direcionamentos sem, no entanto, fugir do ciclo em que se encontra.

Esta relação de interpretação e de descrição do conteúdo está na função da tradução da linguagem natural para as linguagens de indexação. Quando tal relação estiver organizada em categorias, através da determinação do assunto principal e secundários, conduz no sistema de informação o armazenamento e pesquisa, traduzindo a expectativa da busca/resposta através

das palavras e/ou símbolos, às vezes indicando outras relações possíveis. As linguagens documentárias serão tratadas a seguir com maiores detalhes.

### **2.3 LINGUAGENS DOCUMENTÁRIAS**

A linguagem tornou-se, ao longo dos últimos anos, um tema de interesse central no âmbito da Ciência da Informação (CI). Essa intensificação do interesse passou a incorporar a necessidade da efetiva articulação de uma perspectiva mais ampla desse conceito, exigindo a aproximação da CI com as demais áreas do conhecimento que se dedicam ao estudo do fenômeno da linguagem na atualidade. Destaca-se, nesse contexto, a filosofia da linguagem, a análise do discurso, a teoria da terminologia, a teoria do conceito, a semântica, dentre outras.

A linguagem é compreendida, neste trabalho, como um conjunto organizado e compartilhado de signos que permitem a expressão do pensamento.

A explosão informacional, ampliada nos últimos anos devido às inovações tecnológicas, tornou evidente a utilidade das linguagens de indexação em contextos mais amplos, embora o controle de vocabulário seja uma preocupação e uma prática consolidada na área da Ciência da Informação há décadas.

Nesse sentido, a área lidera os estudos e os experimentos relativos à análise dos processos de produção da informação com vistas à criação de instrumentos de organização e recuperação cada vez mais adequados às necessidades dos grupos que deles fazem uso.

É consensual, no interior da CI, que um dos aspectos que mais carece de estudos específicos refere-se à elaboração de linguagens de indexação mais adequadas ao provimento e tradução das necessidades dos usuários em interação com sistemas de informações tradicionais ou eletrônicos.

As linguagens documentárias, também chamadas de linguagens de indexação, constituem-se em um sistema de termos e símbolos que se propõem a traduzir o conteúdo dos

documentos: combinam palavras preferenciais do vocabulário de determinado domínio do conhecimento, interligando a linguagem do autor do documento à do leitor/pesquisador.

O uso das linguagens documentárias, em bases terminológicas, como instrumento de representação/recuperação permite a comunicação entre o documento, a informação e o usuário, uma vez que ela se estabelece através desta linguagem. Entretanto, a informação registrada necessita de uma inter-relação dos conceitos, a conseqüente alteração no significado destes na área e de suas relações com as outras. Como conseqüência, percebe-se esforços científicos, cada vez maiores, no sentido de padronizar a terminologia usada pelas várias ciências.

Sendo assim, só o conjunto de termos estruturados em uma linguagem particular em sua forma e conteúdo, é que permitirá gerar novos termos e/ou alterar significados já existentes, que garanta ao pesquisador, no processo de busca e seleção da informação, encontrar a resposta de que necessita em uma linguagem padronizada e de qualidade. Isto revela que a identificação de conceitos relacionada à sistematização e representação de uma determinada área do conhecimento é de suma importância e, para que isso ocorra, à política de indexação deve ser amplamente tratada, no âmbito da instituição, ponto que se discute a seguir.

## **2.4 POLÍTICA DE INDEXAÇÃO**

A princípio, faz-se necessário conceituar política de indexação e compreender sua importância. De acordo com Carneiro (1985, p.221), uma política de indexação,

[...] deve servir como um guia para tomada de decisões, deve levar em conta os seguintes fatores: características e objetivos da organização, determinantes do tipo de serviço a ser oferecido: identificação dos usuários, para atendimento de suas necessidades de informação e recursos humanos, materiais e financeiros, que delimitam o funcionamento de um sistema de recuperação de informações.

Desse modo, a política de indexação de qualquer instituição deve ser planejada e permeada de cuidados para que o resultado das atividades desenvolvidas na indexação alcance seus objetivos de forma qualitativa.

Cesarino (1985, p.165) complementa a constatação de Carneiro explicando que uma política de indexação só pode ser estabelecida depois de observados os seguintes aspectos:

- identificação das características do usuário (áreas de interesse, nível, experiência, atividades que exercem);
- volume e características da literatura a ser integrada ao sistema;
- volume e características das questões propostas pelo usuário;
- número e qualidade dos recursos humanos envolvidos;
- determinação dos recursos financeiros disponíveis para criação e manutenção do sistema;
- determinação dos equipamentos disponíveis.

Assim, pode-se dizer que a política de indexação constitui-se em várias atividades complexas, que devem ser desenvolvidas em consonância com todo o contexto em que está inserido o sistema de informação, qual seja: a organização a que serve, a clientela a que se destina e os recursos financeiros, materiais e humanos disponíveis. Portanto, no momento do estabelecimento dessa política, todos estes condicionantes devem ser levados em consideração.

Os objetivos de uma política de indexação estão relacionados a definição das variáveis que afetam o desempenho deste serviço, o estabelecimento dos princípios e critérios que servirão de guia na tomada de decisões para sua otimização, a racionalização dos processos e a consistência das operações envolvidas.

Dentre os elementos que compõem uma política de indexação pode-se destacar:

1 - Cobertura de assuntos – identificação das áreas que necessitam de um tratamento aprofundado e das áreas a serem superficialmente tratadas;

2 - Processo de indexação - definição das variáveis que se referem ao:

a) Nível de exaustividade - medida de extensão em que todos os assuntos discutidos em certo documento são reconhecidos na operação de indexação e traduzidos na linguagem do sistema.

b) Nível de especificidade - extensão em que o sistema permite precisão ao especificar o assunto de um documento que esteja sendo processado.

c) Linguagem de indexação - afeta o desempenho de um sistema de recuperação de informação tanto na estratégia de busca (estabelece a precisão com que o técnico de busca pode descrever os interesses do usuário) quanto a indexação (estabelece a precisão com que o indexador pode descrever o assunto do documento).

d) Capacidade de revocação e precisão – exaustividade, revocação e precisão estão relacionadas. Quanto mais exaustivamente um sistema indexa seus documentos, maior será a revocação (número de documentos recuperados) na busca e, inversamente proporcional, a precisão será menor;

3 - Estratégia de busca – definição da responsabilidade para realização da busca de informações em um sistema.

4 - Tempo de resposta do sistema – definição do tempo permitido para ser consumido no momento da recuperação de informações úteis, sendo determinado pelas exigências de revocação e precisão de um sistema;

5 - Forma de saída – é o formato em que os resultados da busca são apresentados;

6 - Avaliação do sistema – identificação da forma como o sistema será avaliado visando a averiguar o nível de satisfação das necessidades de seus usuários, as falhas que estão ocorrendo e a forma como poderão ser corrigidas.

A partir das diretrizes estabelecidas pela política de indexação, os sistemas de informação constituem condições de desenvolver as atividades de representação temática dos documentos de forma racional e consistente.

A definição de uma política de indexação em plena consonância com a organização em que está inserido o SRI e com os usuários a que se destina, é indispensável para o alcance

dos objetivos de um sistema de informação: a recuperação das informações certas no momento adequado.

Dessa forma torna-se importante à delimitação e uso da linguagem artificial e natural para o processo, elemento destacado a seguir.

## **2.5 LINGUAGEM NATURAL**

Ao longo das três décadas de crescimento, expansão, utilização local e remota da informação armazenada em grandes sistemas de informação, inúmeros autores se dedicaram ao estudo das linguagens natural e controlada aplicadas à indexação e recuperação de informações.

A Linguagem Natural (LN) pode ser definida como àquela do discurso técnico-científico e, no contexto da recuperação da informação, Lancaster (1993, p.200) afirma que: “[...] a expressão normalmente se refere às palavras que ocorrem em textos impressos, considerando-se como seu sinônimo e expressão texto livre”. De fato, a linguagem natural é a expressão exata da língua utilizada corretamente em todas as atividades da vida cotidiana. De acordo com o contexto, o nível da língua será familiar, erudito, técnico, poético, literário, etc, e os termos específicos ou genéricos, por exemplo. Esta diversidade ilustra a riqueza e a sutileza da língua. Geralmente, os títulos das obras são redigidos em linguagem natural, o que facilita a sua compreensão pelos leitores.

Na linguagem natural, pode-se escolher um ou conjunto de termos para se fazer uma descrição, uma representação de um assunto, idéia, texto, situação, etc. O termo assim escolhido torna-se uma palavra-chave. Esta permite ao usuário identificar o que é importante ou o que é preciso reter para compreender o essencial de um determinado assunto. Consoante ao caso, a palavra-chave tem um sentido geral ou específico.

A utilização da palavra-chave é uma estratégia eficaz para delimitar um assunto na pesquisa. Também nesta perspectiva, quanto mais à palavra-chave for precisa e pertinente, mais fácil será a recuperação da informação.

Ademais, a utilização da palavra-chave torna-se essencial para fazer a ligação com as linguagens especializadas dos diversos domínios do conhecimento humano.

Bhattacharya (1974) apud Lopes (2002, p.3) analisou os experimentos realizados com a linguagem natural e o seu desempenho para recuperação nas áreas de aerodinâmica, ciência nuclear, física e biologia. Esses testes de avaliação do desempenho da LN, tanto no processo de indexação quanto no de busca, demonstraram que o uso da mesma como linguagem de indexação e/ou de recuperação é viável com um controle mínimo de terminologia.

Knapp (1982) apud Lopes (2002) resumiu brevemente os casos em que a busca em LN pode obter melhor desempenho quando estão relacionados a tópicos específicos; para temas atuais; para novas terminologias ainda não incluídas nas linguagens controladas (LC); para uma busca retrospectiva em que o conceito da LC é muito recente e não cobre os anos anteriores; quando o termo da LC é muito abrangente ou muito específico; para pesquisa em várias bases de dados; para identificação imediata de palavras de títulos dos documentos.

Nas bases de dados bibliográficos, os campos de busca em que se pode pesquisar usando apenas termos e/ou conceitos da LN, normalmente são os do título e resumo dos documentos. Nesses campos, cada palavra é automaticamente candidata a ser pesquisada, executando-se àquelas designadas pelos sistemas como não-significativas que compõem as listas de palavras proibidas.

Assim, as palavras remanescentes são usadas para criar índices que podem ser pesquisados na LN. Esse recurso pode ser utilizado para rastrear temas e seus respectivos conceitos terminológicos que ainda não foram incluídos na LC, ou seja, nos tesouros, nas

listas de cabeçalhos de assunto, nas tabelas de classificações especializadas, listas de descritores, códigos taxonômicos, nomenclaturas e outros.

Cabe ao indexador, que operacionaliza as estratégias de busca, a decisão tática de utilizar esses recursos no planejamento para, alternativamente, ampliar ou restringir os resultados que pretende, de acordo com o definido pelo usuário da informação.

As principais vantagens e desvantagens da LN, citadas na literatura especializada, estão expostas no quadro 1.

<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 1 - Linguagem natural: vantagens e desvantagens

FONTE: Lopes, 2002.

Como pode ser verificado no quadro acima, a linguagem natural possui vantagens e desvantagens. Diante destas questões cabe ao usuário a escolha da opção que melhor se ajuste às suas necessidades. No entanto, o diferencial está em saber aproveitar as oportunidades que aparecem com mais rapidez e eficácia, ou seja, estar atento aos resultados que trarão vantagens na recuperação da informação.

Em contra ponto com a linguagem natural, os SRI desenvolveram a linguagem controlada a ser exposta a seguir.

## **2.6 LINGUAGEM CONTROLADA**

A linguagem controlada é também conhecida como vocabulário controlado e pode ser definida 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 um vocabulário controlado possibilita, no planejamento da estratégia de busca, a recuperação, no campo específico de descritor, apenas daquelas palavras-chave listadas e admitidas antecipadamente pelo sistema. No amplo universo da linguagem, a controlada possui um *status* muito particular uma vez que, por meio dela pode-se representar, de maneira sintética, as informações materializadas nos textos.

Revedo o processo de busca de informação usando o vocabulário controlado, Brundage (1989) apud Lopes (2002) traçou um paralelo entre a linguagem científica e a linguagem controlada. Relembrou que os cientistas, de certa forma, estão familiarizados com o vocabulário controlado em suas áreas de especialização, portanto os paradigmas dos conhecimentos adquiridos nas disciplinas podem ser usados como modelo na prática, especialmente dos tesouros. Observou ainda que o indexador que operacionaliza as buscas dispõe de conhecimentos prévios sobre o uso da LC e da LN, porém o usuário/pesquisador

possui uma profunda intuição sobre a terminologia de seu campo de especialização e essas diferenças vão orientar os programas de treinamento para os diversificados grupos de participantes.

As principais vantagens e desvantagens do uso do vocabulário controlado na recuperação da informação estão sintetizadas no quadro 2.

<b>VANTAGENS</b>	<b>DESVANTAGENS</b>
Controle total do vocabulário de indexação, minimizando os problemas de comunicação entre indexadores e usuários.	Custos: a produção e manutenção da base de dados terá despesas maiores com a equipe de indexadores. Será necessário ainda manter pessoal especializado na atualização do vocabulário controlado.
Com o uso de um vocabulário controlado 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 usuários.
Redução no tempo de consulta à base, pois a estratégia de busca será mais bem elaborada com o uso do vocabulário controlado.	Desatualização do vocabulário controlado poderá conduzir a falsos resultados.

Quadro 2 – Vocabulário controlado: vantagens e desvantagens

FONTE: Lopes, 2002.

Com base nos dados do quadro anterior, verifica-se que o vocabulário controlado assim como a linguagem natural envolve diferentes questões em relação às vantagens e

desvantagens, que devem ser analisadas a luz do custo benefício para quem vai se utilizar dela, para que seja possível, a partir desse diagnóstico, decidir o que é mais viável adotar como política.

De maneira simplificada pode-se dizer que a linguagem documental tem dupla função:

- A primeira é servir como ferramenta de especialista (indexadores ou catalogadores) graças a qual, com a ajuda de termos normalizados (descritores, palavras-chave, cabeçalhos de assuntos) descreve o conteúdo do documento, em outras palavras, indexa as noções ou assuntos principais.
- A segunda é de servir como ferramenta de trabalho para o pesquisador que utiliza então a hierarquia, isto é, os termos normalizados e as suas relações para recuperar, sob um mesmo termo, a informação ou o documento que necessita ou ainda para traduzir as palavras-chave da linguagem natural em termos aceitos pelo sistema de informação.

As discursões até aqui efetuadas visaram constituir elementos que se relacionam diretamente no processo de avaliação estrutural dos tesauros propriamente dito, itens estes que nortearão, de forma segura, a compreensão mais ampla do foco deste estudo, qual seja, os tesauros e sua constituição, como poderá ser observado no próximo item.

## **2.7 TESAURO**

Tesauro é um vocabulário controlado e estruturado, formado por termos que tem entre si relações semânticas e genéricas, de equivalência, hierárquicas e associativas. Trata-se de um instrumento de controle terminológico que permite converter a linguagem natural dos documentos em linguagem controlada, já que representa, de maneira unívoca, seu conteúdo, com a finalidade de servir tanto para a indexação como para a recuperação da informação.

Os termos contidos em um tesauro correspondem a análise que é feita pelo indexador do texto ou material. De fato, ele é constituído de uma série de idéias ou conceitos relevantes que podem ser descritos por meio de termos ou descritores que serão convertidos em um ponto de acesso para a recuperação do documento. Tais considerações preliminares conduzem a necessidade de uma compreensão mais ampla do tema, a ser exposto a seguir.

### 2.7.1 Considerações Gerais

O termo tesauro, segundo Dodebei (2002), se origina do grego *thesaurós* e significa tesouro ou repertório. Este termo se popularizou com a publicação do dicionário analógico de Peter Mark Roget, em Londres, em 1852, intitulado *Thesaurus of english words and phrases*. Roget chamou de tesauro seu dicionário de palavras, uma vez que o termo também designa vocabulário, dicionário ou léxico. O dicionário do autor se diferenciava dos outros por ser um vocabulário organizado de acordo com o seu significado e não pela ordem alfabética e teve o mérito de estabelecer a denominação para vocabulários que relacionam seus termos por meio de algum tipo de relação de significado.

Nos anos 60, Vickery (1980 apud FOSKETT, 1985) apresentou quatro significados para o termo tesouro na literatura da Ciência da Informação, porém o mais importante dentre estes, é o de uma lista alfabética de palavras, onde cada uma delas é seguida de outras relacionadas a mesma.

Segundo Gomes (1990), o tesouro surgiu da necessidade de manipular grande quantidade de documentos especializados. Era preciso trabalhar um vocabulário mais específico e com uma estrutura mais depurada do que aquela presente nos cabeçalhos de assunto (remissivas e referências cruzadas tipo ver e ver também). Assim, além da especificidade, o surgimento desta ferramenta possibilitou uma melhor estrutura e as referências cruzadas deram lugar às relações hierárquicas, associativas e de equivalência, que

serão expostas posteriormente. Pelo fato desse novo instrumento possibilitar, através do agrupamento dos termos, o acesso a uma idéia, essas listas estruturadas passaram a ser chamadas de tesouros, por analogia com a obra de Roget, tendo como função a indexação/recuperação da informação.

Howerton (1965 apud CURRÁS, 1995, p.85) define tesouro como: “[...] uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou intuitivas”. Com efeito, o acesso constitui uma preocupação quando se elabora um tesouro, pois este deve facilitar ao usuário a busca e assegurar pelo mesmo a recuperação da informação desejada.

Outro trabalho relevante sobre o tema é o elaborado pela UNESCO (1973, p.13) que define o tesouro como sendo “[...] vocabulário controlado e dinâmico de termos relacionados semântica e genericamente, que cobre de forma extensiva um campo específico do conhecimento”. A UNESCO compartilha do princípio que o vocabulário seja atualizado constantemente e, dessa forma, suprima a necessidade do usuário.

Das primeiras iniciativas, como a de Mark Roget, até os dias atuais, os tesouros evoluíram em sua definição, construções teóricas e metodológicas, pela introdução de novos modelos cognitivos e abordagem centrada no usuário. Uma definição atual, resultante desta evolução, é a de Currás (1995, p.85), que afirma:

Tesouro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que a compõem – termos, simples ou compostos – encontram-se relacionados entre si sintática e semanticamente.

Em outras palavras, por linguagem especializada entende-se aquela que atua em um domínio restrito; por normalizada, compreende-se uma linguagem controlada; as unidades lingüísticas são termos e, finalmente, por pós-coordenada deseja-se indicar que os termos são combinados no momento de seu uso, em oposição às linguagens pré-coordenadas cujos

termos que designam assuntos complexos se coordenam previamente à sua utilização como as lista de cabeçalho de assunto, por exemplo.

Outra definição relevante é a do *National Information Standards Organization* (1986) citada por Moreira (2003, p.26), que propõe as linhas gerais para a construção, formatação e manutenção de tesouros monolíngues, sendo ele definido como um vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.

O emprego de tesouros nas tarefas de indexação e recuperação de informações, busca resolver o problema da alocação de documentos em classes de assuntos, não só por sua capacidade de controlar o vocabulário, mas porque é um instrumento que relaciona os descritores/termos de forma mais consistente, apresentando uma estrutura sintética simplificada e uma complexa rede de referências cruzadas. Isto permite ao especialista localizar, com mais facilidade, a palavra-chave requerida para uma busca. Apresenta, ainda, um relacionamento lógico e hierárquico dos descritores, o que contribui para a indexação dos documentos em nível específico e/ou genérico.

Segundo Foskett (1972), o principal objetivo do tesouro é o controle terminológico e, em concordância com Lancaster (1972), o autor lista como suas principais funções o controle de sinônimos e quase sinônimos; a distinção de homógrafos; a promoção de facilitadores para condução da busca por meio dos termos relacionados e das referências cruzadas, melhorando a consistência da indexação e transportando a linguagem de busca para a de indexação; e a redução do tempo e aumento da eficiência nas tarefas de indexação e recuperação de informações.

Deste modo observa-se que a função do tesouro é ampla, motivo pelo qual se discute tal aspecto a seguir com objetivo de permitir maior compreensão das exigências e mecanismos quando de sua elaboração.

### 2.7.2 Função do Tesouro

Segundo Currás (1995, p.34), o tesouro foi adotado "[...] na área de documentação, associado à forma de organização do vocabulário de indexação/recuperação". De fato, ele pode funcionar em um ambiente organizacional, na representação dos assuntos dos documentos e nas buscas informacionais. A representação dos assuntos dos documentos é realizada apenas pelo indexador que analisa-o e identifica seu conteúdo para depois traduzi-lo para os termos indicados em um tesouro.

O tesouro pode ser usado tanto para auxiliar na elaboração de consultas realizadas pelo usuário/pesquisador, em suas buscas informacionais, quanto pelo indexador durante o processo de classificação. Para os usuários, o tesouro, pela sua estrutura de termos e suas relações, auxilia a encontrar os melhores termo(s) que representem um assunto e descrevam uma obra.

Portanto, o tesouro é um componente de extrema relevância para um sistema de recuperação pois determinar quais termos podem ser empregados na busca para que esta tenha um resultado satisfatório além de permitir a introdução de novos descritores em sua estrutura e relações de modo a aproximar a linguagem do usuário à do sistema e realizar alterações dos já existentes.

São vários os tipos de tesouros que conduzem ao objetivo acima descrito, como se pode observar a seguir.

### 2.7.3 Tipos de Tesouros

Atendendo a diferentes aspectos, os tesouros podem variar conforme a língua, nível de especificidade e assunto que cobrem dos seguintes modos.

a) Em relação a língua – podem-se distinguir os tesouros pelos idiomas em que se pode escrever seus termos em:

- Monolíngues - o que contém descritores em uma única língua.
- Multilíngues - contém descritores em mais de uma língua.

b) Em relação ao nível de especificidade:

- Macrotresouros – os termos representam conceitos mais ou menos amplos, resultando em um menor número de descritores. Currás (1995) declara que um macrotresouro é composto de vários microtesouros, relacionados entre si por referências cruzadas, abrangendo uma ampla gama de assuntos.
- Microtesouros – os descritores denotam conceitos em um nível maior de especificidade e se referem a um domínio mais restrito.

As tipologias acima expostas demonstram que os tesouros são elementos complexos que configuram uma série de relações terminológicas como exposto no próximo item.

### 2.7.4 Tipos de Relações nos Tesouros

Como já foi mencionado o tesouro surgiu da necessidade de manipulação de grande quantidade de documentos especializados, onde é preciso trabalhar com vocabulário mais específico e uma estrutura mais articulada e integrada do que aquela presente nos cabeçalhos de assunto (remissivas e referências cruzadas tipo 'ver' e 'ver também'). Tudo isso em função da necessidade dos usuários em recuperarem documentos de uma área específica, na qual é necessário maior sistematização para a recuperação.

A estrutura de um tesauro é delineada pelos relacionamentos, às ligações e vinculações existentes entre os conceitos representados por termos. Nenhum termo existe em um tesauro sem ligação com outro, sempre determinada por seu significado (SVENONIUS, 2000).

São as relações existentes entre os termos que conferem a um tesauro uma multiplicidade de usos, desde o auxílio na própria função de indexação até o auxílio na efetiva recuperação dos documentos.

Currás (1995) apresenta as relações existentes entre os termos de um tesauro classificadas em relações de equivalência, hierárquicas e associativas, como:

a) Relações de equivalência - são aquelas que representam os sinônimos ou quase sinônimos de um termo. Indica que um conceito pode ser expresso por dois ou mais termos. Nesse caso, um deles deve ser determinado como preferencial e os demais como não preferenciais. Os indicadores de relacionamentos de equivalência são:

UP – Use para

USE – Use

**Exemplos:**

Linguagens de indexação

UP Linguagens de documentação

Linguagens de documentação

USE Linguagens de indexação

b) Relações hierárquicas – surgem da necessidade explicar a relação de superordenação ou subordinação, apontando os termos genéricos com conotação mais ampla e os termos específicos com conotação mais específica. Os indicadores de relacionamento hierárquico são:

TG – Termo Genérico

TE – Termo Específico

**Exemplos:**

Linhas aéreas

TG: Serviços de transporte

Serviços de Transporte

TE: Linhas aéreas

c) Relações Associativas – apontam as características comuns entre os conceitos que denotam um relacionamento simétrico, diferente das demais, de tal forma que um termo indica outro (veja também). São termos semântica ou conceitualmente associados, mas não são equivalentes. O indicador de relacionamento associativo é:

TR – Termo Relacionado

**Exemplos:**

Política

TR: Políticos

Assim expostas, os tipos de relação demonstram a cadeia de construção de um tesauro que influenciam nos aspectos inerentes à sua metodologia de elaboração, como será exposto a seguir.

### 2.7.5 Metodologia de elaboração dos Tesauros

É importante salientar alguns tópicos referentes a metodologia dos tesauros, visando evidenciar certos aspectos que serão importantes no processo de análise desta pesquisa e possibilitará um entendimento mais amplo sobre as questões que os envolvem.

De modo geral, pode-se organizar os trabalhos de elaboração de um tesouro subdividindo-o em três grandes partes, quais sejam: uma preliminar, uma de elaboração propriamente dita e uma de editoração.

Entende-se, como etapa preliminar, a constituição de uma equipe de trabalho, o planejamento geral do tesouro, a delimitação do tema central e o levantamento das principais fontes de coleta dos candidatos a descritor.

Pelo caráter interdisciplinar do trabalho de elaboração de um tesouro, a equipe deve ser constituída por profissionais bibliotecários, especialistas do domínio de conhecimento a ser coberto pelo tesouro e especialistas em informática. Cabe, a essa equipe, planejar toda a estrutura do tesouro. Nesse planejamento, deve ser explicitado o método de trabalho, o de coleta de termos, os critérios de determinação do descritor autorizado, a rede de relações, entre outros tópicos.

Ademais, em uma perspectiva de planejamento, salienta-se o que se denomina de delimitação temática, que refere-se fundamentalmente, à definição do tema central do tesouro, devidamente contextualizado, além da indicação dos temas periféricos importantes que devem dele constar bem como o tratamento a ser dado a esses tópicos. Nesse processo, entende-se que, além da delimitação do campo temático, ocorre a compreensão da área de especialidade como um todo e isso possibilita o estabelecimento das categorias nas quais o mesmo se organiza. Assim, é possível evitar-se configurações não representativas.

Finaliza-se a etapa preliminar quando ocorre o levantamento das fontes que serão utilizadas para a coleta dos descritores. As principais fontes são os próprios textos produzidos pelos especialistas do campo temático do tesouro. Dessa forma, os prováveis descritores serão coletados em uma situação discursiva real, contextualizados e, com a possibilidade de identificar outras denominações relativas a um mesmo conceito, ou seja, situações de sinonímia, de variações terminológicas ou ainda casos de homonímia. Também são fontes

importantes para posterior trabalho sobre as definições, os dicionários especializados e, principalmente, os bancos de dados terminológicos da área. Neste contexto se insere também o levantamento de outros tesouros que tratem da mesma temática.

A etapa de elaboração do tesouro, compreende a coleta dos candidatos a descritores, elaboração de um glossário, determinação dos descritores autorizados e os não-autorizados, validação dos descritores, estabelecimento das relações hierárquicas, associativas e de equivalência, além da elaboração da apresentação alfabética do tesouro.

A última etapa é a do arranjo alfabético dos descritores com as respectivas relações.

Os procedimentos seguintes referem-se à editoração, entrando nesta etapa a redação de uma apresentação criteriosa e elucidadora, entre outros tópicos. De modo geral, devem constar, nesta apresentação os objetivos, o âmbito ou a cobertura, a forma de apresentação, as abreviaturas usadas, a forma gráfica dos descritor e do não-descritor, as orientações de uso, o número total de termos, o critério para seleção dos mesmos, as fontes consultadas entre outras informações.

Tendo em vista que as proposições apresentadas neste estudo, têm valor de sugestões de diretrizes para elaboração do tesouro em biotecnologia, um instrumento essencial à apresentação/recuperação das informações que possam contribuir para qualificar um segmento relevante para a economia regional, a seguir enfatiza-se a área em tela visando contribuir para o amplo entendimento do corpus que posteriormente será analisado.

## **2.8 TESAUROS E A BIOTECNOLOGIA**

O tesouro sobre biotecnologia tem ampla relevância tendo em vista a inexistência de uma ferramenta dinâmica desenhada com o propósito de facilitar a descrição de conjuntos de dados e sua recuperação. Assim, buscar-se-á compor um conjunto de saberes relacionados

entre si a partir de aspectos hierárquicos, por associação e por equivalência, estabelecendo relações de acordo com o conceito ou conceitos que cada termo representa.

A palavra biotecnologia é formada por três termos de origem grega:

- *bio*, que quer dizer vida;
- *logos*, que significa conhecimento, e
- *tecno*, que designa a utilização prática da ciência.

De acordo com o Glossário de biotecnologia (2005, p.10), a palavra tem como significado: “[...] aplicar processos biológicos à produção de materiais e substâncias para uso industrial, medicinal, farmacêutico, etc”.

A biotecnologia agrega conhecimento nas áreas de microbiologia, bioquímica, genética, engenharia química e informática. Envolve a manipulação de agentes biológicos como microrganismos, células e moléculas (enzimas, anticorpos, DNA, etc.), resultando em bens como alimentos, bebidas, produtos químicos, energia, produtos farmacêuticos, pesticidas, etc. Contribui com serviços como a purificação da água, tratamentos de resíduos, controle de poluição, etc, para elevar a qualidade de vida da população.

O termo biotecnologia foi empregado pela primeira vez em 1919, por um engenheiro agrícola da Hungria. Entretanto, em um sentido amplo, a humanidade vem utilizando a biotecnologia desde a Antigüidade. Em 1800 a.C. a população utilizava os processos de fermentações para a fabricação de vinho, pão, queijo, etc.

No início do século XX desenvolveram-se as técnicas de cultura de tecidos e, a partir de meados do mesmo, surgiram novos horizontes com a Biologia Molecular e com a Informática que permite a automatização e o controle das plantas em níveis industriais.

Subseqüentemente, o homem tem ampliado as técnicas de manipulação dos seres vivos, promovendo um desenvolvimento significativo nas mais variadas áreas, como agricultura, medicina, indústria química, etc.

Foi depois dos anos 70, com cientistas americanos, que a biotecnologia concentrou seus esforços nas pesquisas com o DNA (material genético) e, com isso, foi possível criar os organismos geneticamente modificados (OGMs), também conhecidos como transgênicos.

Deste modo, a biotecnologia é o ramo da ciência que pesquisa e promove a transferência de genes de uma espécie para outra, a fim de atribuir a esta última, características naturais da primeira. A utilização da biotecnologia tem possibilitado o surgimento de produtos em todas as áreas: plantas geneticamente modificadas, técnicas para a criação de medicamentos, hormônios, anticorpos, enzimas, entre outros.

Diante o exposto, faz-se necessário abordar em maiores detalhes, os assuntos relacionados a aplicação da biotecnologia nas diversas áreas do saber, com a função de possibilitar o entendimento da abordagem do campo de ação que ela envolve.

## **2.9 APLICAÇÕES DA BIOTECNOLOGIA**

A biotecnologia é uma ciência em constante evolução, que está sempre oferecendo produtos novos e mais eficientes, em diversas áreas. Atualmente, os setores que mais se beneficiam do desenvolvimento desta tecnologia são a agricultura e a indústria farmacêutica mas, as suas aplicações não se restringem a esses dois setores. Combustíveis, plásticos e tecidos são outros exemplos para os quais a biotecnologia poderá trazer grandes benefícios, como pode ser observado a partir das áreas específicas:

- **Agricultura:** plantas geneticamente modificadas, que são tolerantes a herbicidas e/ou resistentes a insetos e pragas, já são uma realidade e diversos países do mundo. No ano de 2001 foram cultivados 52,6 milhões de hectares de plantas geneticamente modificadas nos 13 países que já aprovaram o cultivo comercial deste tipo de planta.

Essas plantas, de propriedades econômicas, simplificam o controle das daninhas, gerando economia para os agricultores e benefícios para os consumidores, que têm à sua disposição alimentos originários de plantas que foram menos afetadas por doenças. Já está em fase de pesquisa e/ou teste plantas enriquecidas com vitaminas e outras substâncias, frutas que permanecem maduras e adequadas para consumo por mais tempo e, no futuro próximo, plantas que poderão funcionar como vacinas, combatendo doenças que afetam milhões de pessoas em todo o mundo, como a diarreia, cegueira noturna, sarampo, leishmaniose e outras.

- **Medicina:** a primeira aplicação comercial ocorreu em 1982, com a produção da insulina para o tratamento de diabetes. O gene que produz a insulina humana foi isolado e transferido para uma bactéria. Enquanto elas crescem e vivem em tanques de fermentação, produzem a proteína insulina que é então isolada e purificada para o trato do diabetes. Esse processo é fundamental, pois antes dele, era necessário utilizar insulina suína ou bovina, menos eficientes do que a humana, para o tratamento da doença.

Além da insulina, hormônios de crescimento humano, *kits* para diagnóstico de doenças, vacinas, anticorpos e o interferon humano (substância natural sintetizada no organismo humano para a defesa contra vírus) também são originários da biotecnologia.

- **Indústria plástica:** a partir da modificação genética de bactérias, a indústria já começa a desenvolver plásticos que se decompõem mais rapidamente no meio ambiente e, por isso, poluem muito menos o solo, as águas e o ar.

A indústria têxtil também poderia se beneficiar da biotecnologia, ao utilizar, por exemplo, algodão geneticamente modificado que produz fibras mais resistentes ou, ainda, fibras coloridas naturalmente evitando o uso de pigmentações artificiais, o

que diminuiria os gastos com o tingimento, que também causa impactos ambientais.

- **Outras áreas:** a biotecnologia não trabalha apenas com alimentos, indústria farmacêutica, de tecidos e plásticos. Existem também pesquisas para buscar outras formas de gerar energia, utilizando matéria como a madeira, girassol, milho, soja e cana-de-açúcar, por exemplo, para produzir biocombustíveis, que não prejudicam o meio ambiente e que são mais baratos.

Deste modo, uma vez constituído o arcabouço teórico desta pesquisa e exposto os princípios que permitem compreender o campo do conhecimento eleito para o estudo, faz-se necessário apontar, no próximo Capítulo, o percurso metodológico adotado para atingir os objetivos propostos.

## CAPÍTULO III

Neste capítulo apresentam-se os procedimentos metodológicos utilizados para o desenvolvimento do trabalho, destacando os métodos utilizados, o universo e a composição da amostra, além das variáveis de análise da pesquisa, de modo a favorecer o entendimento das análises efetuadas.

### PERCURSO METODOLÓGICO

#### MÉTODO

A pesquisa, de caráter qualitativo, está baseada no método bibliográfico documental e exploratório-descritivo, sendo que buscou-se explorar as seguintes temáticas: informação, linguagem documentária, tesouros, biotecnologia.

#### 3.1 UNIVERSO E AMOSTRA

Através de levantamento preliminar na Internet, identificou-se dez tesouros contendo assuntos como alimentos, bebidas, agricultura, fármacos e outros tantos relacionados a biotecnologia. Dos tesouros identificados, arrolou-se oito por serem os mais significativos no que tange a proeminência para a área, sendo desses eleitos três que compuseram a amostra desta pesquisa.

- a) Tesouro Cadeia Alimentar – TCA - organizar informação sobre alimentação e nutrição. Envolve vários tipos de agentes sociais, tais como: produtor, distribuidor, especialista no tema, elaborador de políticas públicas, consumidor e/ou apreciador da boa mesa.
- b) Macrothesaurus OECD *Chapter Heading* – relaciona áreas da agricultura, indústria, biologia, bioquímica, alimentos, nutrição, fármacos, biotecnologia, etc.
- c) Tesouro de Alimentos e Bebidas HONselect – relaciona exhaustivamente a área de bebidas e alimentos, focos de atuação da biotecnologia.

- d) Thesaurus Canadian de Ciência e Tecnologia – cobre exaustivamente a área de engenharia, construção civil, ciência e tecnologia, dentro dessa última comporta um item que trabalha com a biotecnologia.
- e) Thesaurus EnVoc - vocabulário controlado do sistema INFOTERRA, que trabalha com informação sobre meio ambiente.
- f) Tesouro Alfabético Conceitual – aborda às áreas de agricultura e alimentos.
- g) Tesouro do BCDAM – enfoca a área de meio ambiente, em especial a recuperação de petróleo, tratamento do lixo, purificação da água. Assuntos tratado dentro da biotecnologia.
- h) Glossário de Biotecnologia para a Agricultura e Alimentação – enfocando a agricultura e os alimentos com especial atenção à biotecnologia.

Nesta pesquisa, selecionou-se os tesouros de Alimentos e Bebidas HONselect; o Tesouro Cadeia Alimentícia - TCA e o Glossário de Biotecnologia para a Agricultura e Alimentação, pela abrangência ampla que eles apontam, por estarem voltados diretamente para bens e produtos da área de biotecnologia e por apresentarem os itens selecionados para análise.

### **3.2 VARIÁVEIS DE ANÁLISE**

Considerando os parâmetros de avaliação dos tesouros expostos na literatura por Lancaster (1993) e Urdiciain (1998), foram adotadas as seguintes variáveis de análise para atingir aos objetivos propostos por esta pesquisa:

- a) Forma de apresentação – alfabética, sistemática ou gráfica.
  - Alfabética – os descritores e os não descritores se agrupam em uma única seqüência alfabética junto com suas relações.

- Sistemática – estruturado em duas partes. A primeira é a principal, contém as categorias ou hierarquias e a segunda é a auxiliar, consta de um índice alfabético que conduz os usuários a seção semântica correspondente a que pertence o termo.
  - Gráfica – os termos se apresentam em forma de gráfico, onde se associam os termos relacionados.
- b) Índice e/ou introdução – identificação de instrumento que permita ao usuário ser auxiliado em relação ao manuseio dos tesouros explicitando suas características no âmbito de aplicação.
- c) Tamanho – levantamento do número de descritores que compõe o tesouro.
- d) Tipos de Relacionamento – estabelecimento de comparações por semelhança e diferenças entre os conceitos, quer seja para defini-los, nomeá-los ou ordená-los.
- e) Quantidade de notas esclarecedoras registradas por tesouros – chamadas de Nota de Escopo, que são utilizadas para indicar o conceito do termo. É indicada na estrutura por:
- NE – Nota de Escopo
- f) Morfologia das palavras – uso que se faz nos tesouros das formas das palavras, singular ou plural, e se os descritores se expressam por meio de entradas diretas, sem inversão dos termos, respeitando a ordem natural das expressões.
- g) Idiomas – facilidade para os usuários, pois na utilização dos tesouros pode selecionar a língua que mais tem afinidade.

### **3.3 PROCEDIMENTOS DE COLETA DE DADOS**

Após seleção pela Internet, realizou-se a análise individualizada de cada um considerando as variáveis eleitas. Posteriormente, a partir da análise individual que priorizou

o exame das características consideradas essenciais para um tesouro com bom nível, efetuou-se a análise comparativa. O resultado de ambas análises estão expostos no Capítulo seguinte.

## CAPÍTULO IV

Neste capítulo será abordada a análise individualizada dos tesouros selecionados e, posteriormente, apresenta-se o exame global visando compor a dimensão estabelecida pelos objetivos da pesquisa em curso.

### ANÁLISE E DISCUSSÃO DOS RESULTADOS

#### 4.1 ANÁLISE INDIVIDUAL DOS TESAuros

##### 1) Alimentos e Bebidas HONselect

###### a) Caracterização

O primeiro tesouro em análise aborda assuntos relacionados a saúde, alimentos e bebidas. Foi criado pela Fundação HON - *Health On the Net*, organismo não governamental com *status* consultivo especial no Conselho Econômico e Social das Nações Unidas, que tem como missão orientar a crescente comunidade de consumidores a buscar informações médicas confiáveis e precisas. A figura 9 destaca a página inicial do tesouro, mostrando seus detalhes estruturais.

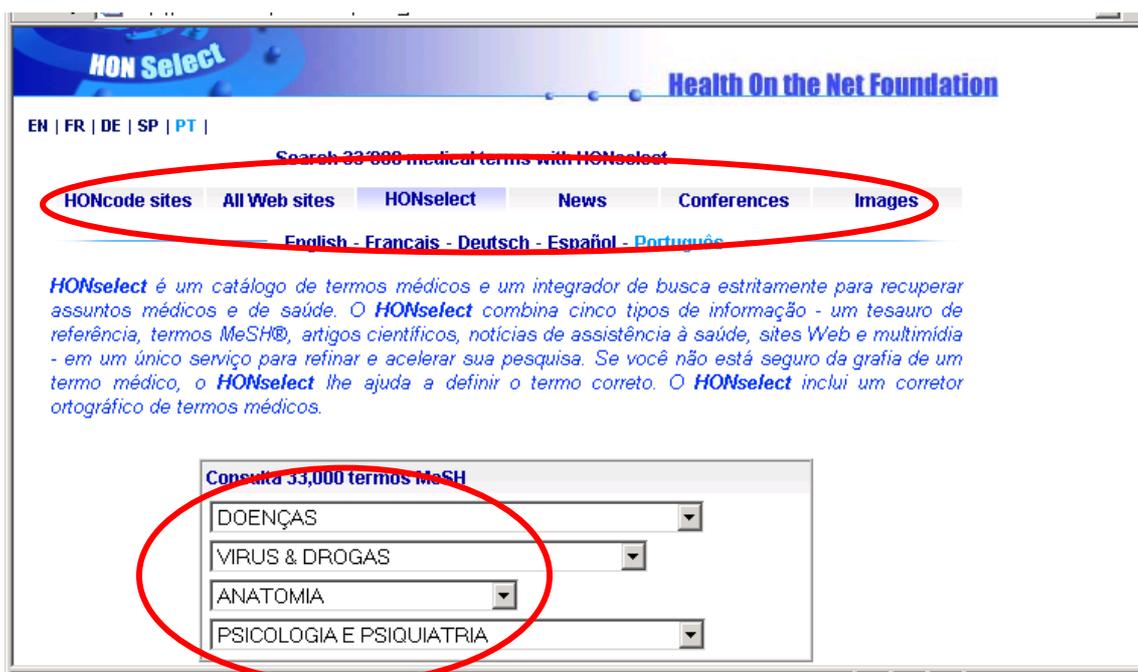


Figura 9 – Página principal do tesouros HONselect  
FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

O *HONselect* disponibiliza aos usuários, uma imensa variedade de informações, todas disponíveis na Internet facilitando dessa forma o acesso as buscas *on line*.

Trabalha com vários assuntos, visando sempre o desenvolvimento mundial. Também proporciona através de sua página, como pode ser visualizado na figura 9, informações completas de buscas a vários sítios médicos através da *web*, onde o usuário pode consultar a base de dados em nível mundial sobre conferências médicas além de permitir uma busca avançada de 6.800 imagens e vídeos médicos.

A consulta aos 33.000 termos disponíveis no tesouro em estudo pode ser feita escolhendo-se uma das quatro categorias – Doenças; Vírus e drogas; Anatomia; Psicologia e psiquiatria existentes, conforme pode ser verificado no destaque da figura acima.

Na figura 10 pode ser visualizado a área principal sobre tecnologia, alimentos e bebidas, que remeterá aos tópicos sobre biotecnologia, foco desse trabalho. Cabe ainda explicitar que este tesaurus disponibiliza para seus usuários uma lista de termos favoritos, de doenças raras e um campo onde pode ser enviado sugestões aos colaboradores.

Hierarquia - Inglês - Francês - Alemão - Espanhol - Português

- ANATOMIA
- ORGANISMOS
- DOENÇAS
- COMPOSTOS QUÍMICOS E DROGAS
- TÉCNICAS E EQUIPAMENTOS
- PSICOLOGIA E PSIQUIATRIA
- CIÊNCIAS BIOLÓGICAS
- CIÊNCIAS FÍSICAS
- ANTROPOLOGIA, EDUCAÇÃO, SOCIOLOGIA E FENÔMENOS SOCIAIS
- TECNOLOGIA E ALIMENTOS E BEBIDAS**
  - Alimentos e Bebidas
  - Tecnologia, Indústria e Agricultura
- HUMANIDADES
- CIÊNCIA DA INFORMAÇÃO
- PESSOAS
- ASSISTÊNCIA À SAÚDE
- LOCALIZAÇÕES GEOGRÁFICAS

MeSH 2005 © U.S. National Library of Medicine (NLM)®  
MeSH 2005 © DeCS Health Science Descriptor - Bireme (Português)

Consulta - Nova busca

Favoritos Lista de Doenças Raras Add HONselect

Figura 10– Página mostrando a área de Tecnologia, alimentos e bebidas.

FONTE: [http://www.hon.ch/HONselect/index\\_pt](http://www.hon.ch/HONselect/index_pt)

No campo dos Favoritos encontra-se, como visualizado no destaque da figura 11, vários assuntos elencados em ordem alfabética, facilitando dessa forma a busca por parte do usuário.



Lista de Favoritos: [english](#) - [français](#) - [deutsch](#) - [español](#) - [português](#)

- **Abdome**
- **Agentes Antiinflamatórios, Agentes Anti-Reumáticos e Mediadores da Inflamação**
- **Agentes do Sistema Nervoso Central**
- **Alimentos e Bebidas**
- **Anormalidades Cardiovasculares**
- **Antraz**
- **Artrite**
- **Artrite Infecciosa**
- **Artrite Reumatóide**
- **Artropatias**
- **Asma**
- **Asma Induzida por Exercício**
- **Ações Químicas e Utilizações**
  
- **Bactérias**
- **Broncopatias**
  
- **Cancer**
- **Características da População**
- **Cardiopatias**
- **Ciência da Informação**
- **Ciências Naturais**

Figura 11 – Lista de favoritos

FONTE: [http://www.hon.ch/HONselect/index\\_pt](http://www.hon.ch/HONselect/index_pt)

Na figura 11, em relação ao item Favoritos que aparece em destaque na figura 10 no final da página, pode-se verificar que ao clicá-lo tem-se acesso a lista em ordem alfabética de assuntos relacionados as partes do corpo humano, tipos de medicamentos, alimentos e bebidas, anormalidades do coração, doenças, ciências naturais e da informação, comportamentos, bactérias, entre outros assuntos.

A figura a seguir, demonstra a relação, em ordem alfabética, de nome completo de doenças raras, por exemplo.



Figura 12 – Lista de doenças raras  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt](http://www.hon.ch/HONselect/index_pt)

Na figura 12, que também é um desdobramento da figura 10, quanto ao item lista de doenças raras, cada um desses termos apresentados em ordem alfabética, tem desdobramentos que estão relacionados ao tipo de doenças que se busca informações. Dentro do *link* selecionado aparecerão outras informações mais detalhadas a respeito da mesma.

### ***b) Forma de apresentação***

Como pode ser verificado na figura 13, o tesouro *HONselect* indexa saberes em áreas como medicina, mais também com outras que inferem na ciências da vida. Dentre elas está a de tecnologia, alimentos e bebidas, que aponta um desdobramento para alimentos e bebidas, o foco do estudo da área de biotecnologia.



Figura 13– Página mostrando as áreas trabalhadas pelo HONselect  
FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Dentro do campo de alimentos e bebidas, o tesouro *HONselect* facilita a busca, uma vez que relaciona a definição, sinônimos e os termos específicos pelos quais os assuntos podem ser recuperados, conforme pode ser visualizado na figura 14.



The screenshot displays the HONselect search interface. At the top, there is a logo for 'HON Select' and the text 'Health On the Net Foundation'. Below this, a search bar contains the text 'Alimentos e Bebidas' and a dropdown menu set to 'a palavra'. To the right of the search bar are buttons for 'Busca, Busque' and 'Limpar'. Below the search bar, there is a dropdown menu for 'Conteúdo sobre "Alimentos e Bebidas":'. The main content area shows the search results for 'Alimentos e Bebidas', which are circled in red. The results include the definition 'Edible or potable substances.', synonyms 'Bebidas e Alimentos /', and specific terms 'Bebidas' and 'Alimentos'.

HONcode sites All Web sites HONselect News Conferences Images  
Busca em HONselect - Inglês - Francês - Alemão - Espanhol - Português  
Alimentos e Bebidas Busca, Busque Limpar  
a palavra em termo MeSH  
Conteúdo sobre "Alimentos e Bebidas":  
Hierarquia - Inglês - Francês - Alemão - Espanhol - Português  
**Alimentos e Bebidas**  
Definição: Edible or potable substances.  
Sinônimos: Bebidas e Alimentos /  
Termo(s) específico(s):  
Bebidas Alimentos

Figura 14– Página mostrando a área de alimentos e bebidas.

FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Na figura 15, o destaque ressalta que o tesouro *HONselect* dissecar os assuntos relacionados a bebidas, trazendo também a definição, os sinônimos e os termos mais específicos relacionados ao tema.



The screenshot displays the HONselect search interface. At the top, it says "Health On the Net Foundation" and "Search 33,000 medical terms with HONselect". Below this, there are navigation links for "HONcode sites", "All Web sites", "HONselect", "News", "Conferences", and "Images". A search bar contains the word "Bebidas" and a dropdown menu is set to "a palavra". To the right of the search bar are buttons for "Busca, Busque" and "Limpar". Below the search bar, there is a dropdown menu for "Conteúdo sobre 'Bebidas:'". Further down, there are language options: "Hierarquia" and "- Inglês - Francês - Alemão - Espanhol - Português". The main content area is titled "Bebidas" and contains the following information:

- Definição:** Liquids that are suitable for drinking. (From Merriam Webster Collegiate Dictionary, 10th ed)
- Sinônimos:** Chás de Ervas / Chás Medicinais /
- Termo(s) específico(s):**
  - Bebidas Alcoólicas
  - Leite
  - Substitutos do Leite
  - Bebidas Gasosas
  - Café
  - Águas Minerais
  - Chá

Figura 15– Página mostrando o assunto de bebidas.  
FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Na figura 16, verifica-se que o item bebidas pode apresentar outros desdobramentos mais específicos como bebidas alcoólicas, trazendo também, como nos itens anteriores, a definição e os termos mais particulares. Esses detalhamentos encontrados na estrutura deste tesaurus, favorece e assegura ao usuário uma pesquisa aprofundada pois, a partir do momento que ele seleciona o termo procurado, saberá se é ou não relevante para o seu trabalho.

The screenshot shows the Health On the Net Foundation search interface. At the top, it says "Search 33,000 medical terms with HONselect". Below this, there are navigation links for "HONcode sites", "All Web sites", "HONselect", "News", "Conferences", and "Images". A search bar contains the text "Bebidas Alcoólicas" and "Busca, Busque" with a "Limpar" button. Below the search bar, there are dropdown menus for "a palavra" and "em termo MeSH". A section titled "Hierarquia" shows "Bebidas Alcoólicas" selected. Below this, the definition is given as "Drinkable liquids containing ETHANOL.". Under "Termo(s) específico(s):", there are links for "Absinto (extrato)", "Cerveja", and "Vinho". Under "Veja também:", there is a link for "Etanol". A red oval highlights the definition, specific terms, and related terms section.

Figura 16– Página mostrando o assunto de bebidas.  
FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

O tesaurus *HONselect* apresenta ainda uma lista, em ordem alfabética, de termos favoritos, ou seja, relacionados as grandes áreas no qual atua, conforme pode ser observado na figura 17.



The screenshot shows the HONselect website interface. At the top left is the HON Select logo, and at the top right is the Health On the Net Foundation logo. Below the logos is a navigation bar with the text "Lista de Favoritos: english - français - deutsch - español - português". The main content area displays a list of favorite terms in alphabetical order, each preceded by a blue circular icon. The terms listed are: Abdome, Agentes Antiinflamatórios, Agentes Anti-Reumáticos e Mediadores da Inflamação, Agentes do Sistema Nervoso Central, Alimentos e Bebidas, Anormalidades Cardiovasculares, Antraz, Artrite, Artrite Infecciosa, Artrite Reumatóide, Artropatias, Asma, Asma Induzida por Exercício, Ações Químicas e Utilizações, Bactérias, Broncopatias, and Cancer.

**HON Select**

**Health On the Net Foundation**

Lista de Favoritos: english - français - deutsch - español - português

- Abdome
- Agentes Antiinflamatórios, Agentes Anti-Reumáticos e Mediadores da Inflamação
- Agentes do Sistema Nervoso Central
- Alimentos e Bebidas
- Anormalidades Cardiovasculares
- Antraz
- Artrite
- Artrite Infecciosa
- Artrite Reumatóide
- Artropatias
- Asma
- Asma Induzida por Exercício
- Ações Químicas e Utilizações
- Bactérias
- Broncopatias
- Cancer

Figura 17– Página mostrando lista de favoritos em ordem alfabética.

FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

### c) Índice e/ou introdução

Como pode ser verificado na figura 18, o tesouro em análise comporta uma página caracterizando o *HONselect* e de que forma esta estruturado, mostrando ao usuário que o mesmo pode ser acessado por uma palavra ou parte da mesma para, com isso, facilitar a recuperação da informação. Disponibiliza ainda ao pesquisador, um corretor ortográfico para ajudar, no caso de dúvidas, a definir o termo pretendido.

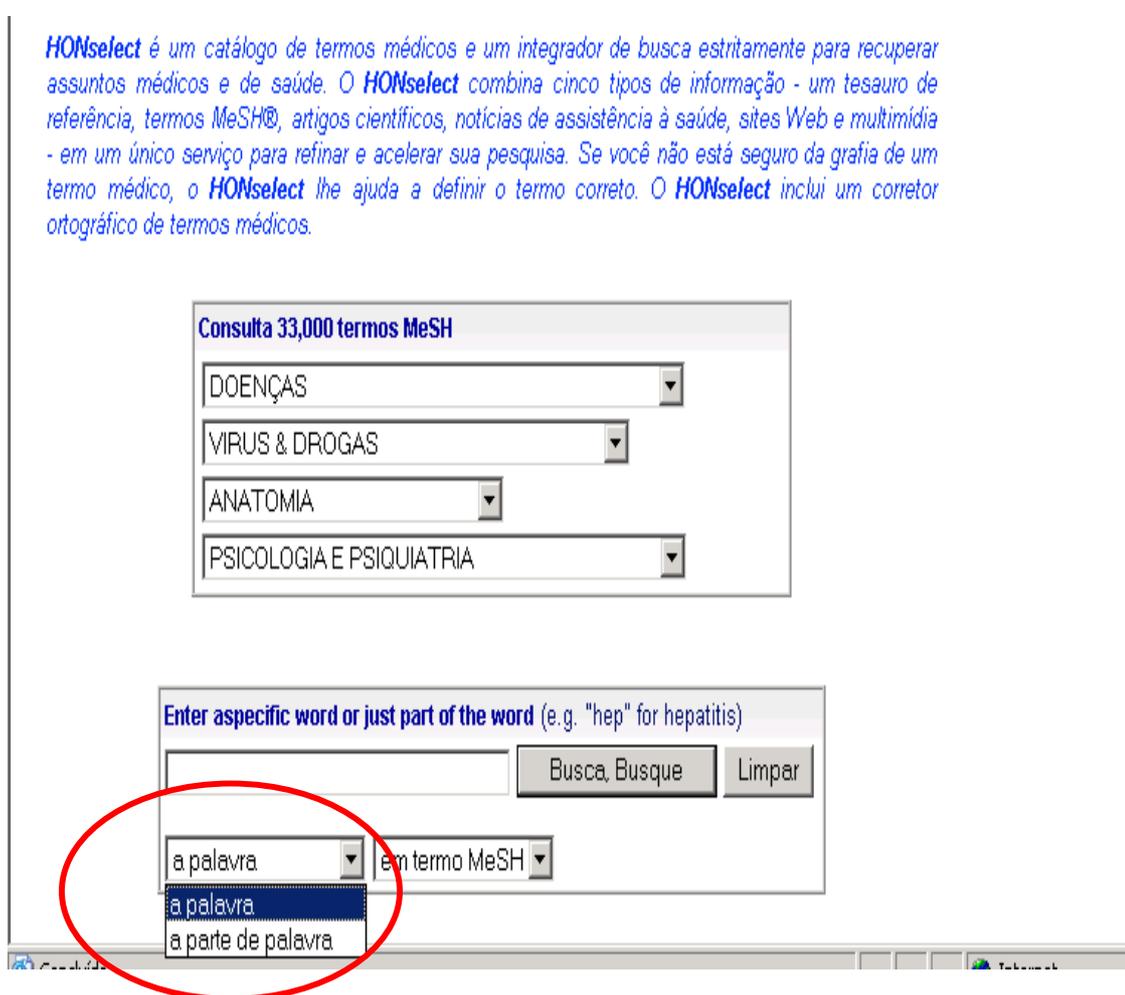


Figura 18– Página contendo introdução.  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Tais mecanismos, presentes no tesouro conforme destacado na figura acima, é denotativo da relevância que a organização deste instrumento de linguagem controlada oferece para a interatividade quando do seu manuseio.

#### d) Tamanho

Conforme pode ser verificado na figura 19, o tesouro *HONselect* possui 33.000 (trinta e três) mil termos, o que demonstra um elevado grau de relevância de sua abrangência, pois tal quantidade, facilita a recuperação da informação para o usuário em menor tempo e com maior nível de cobertura.

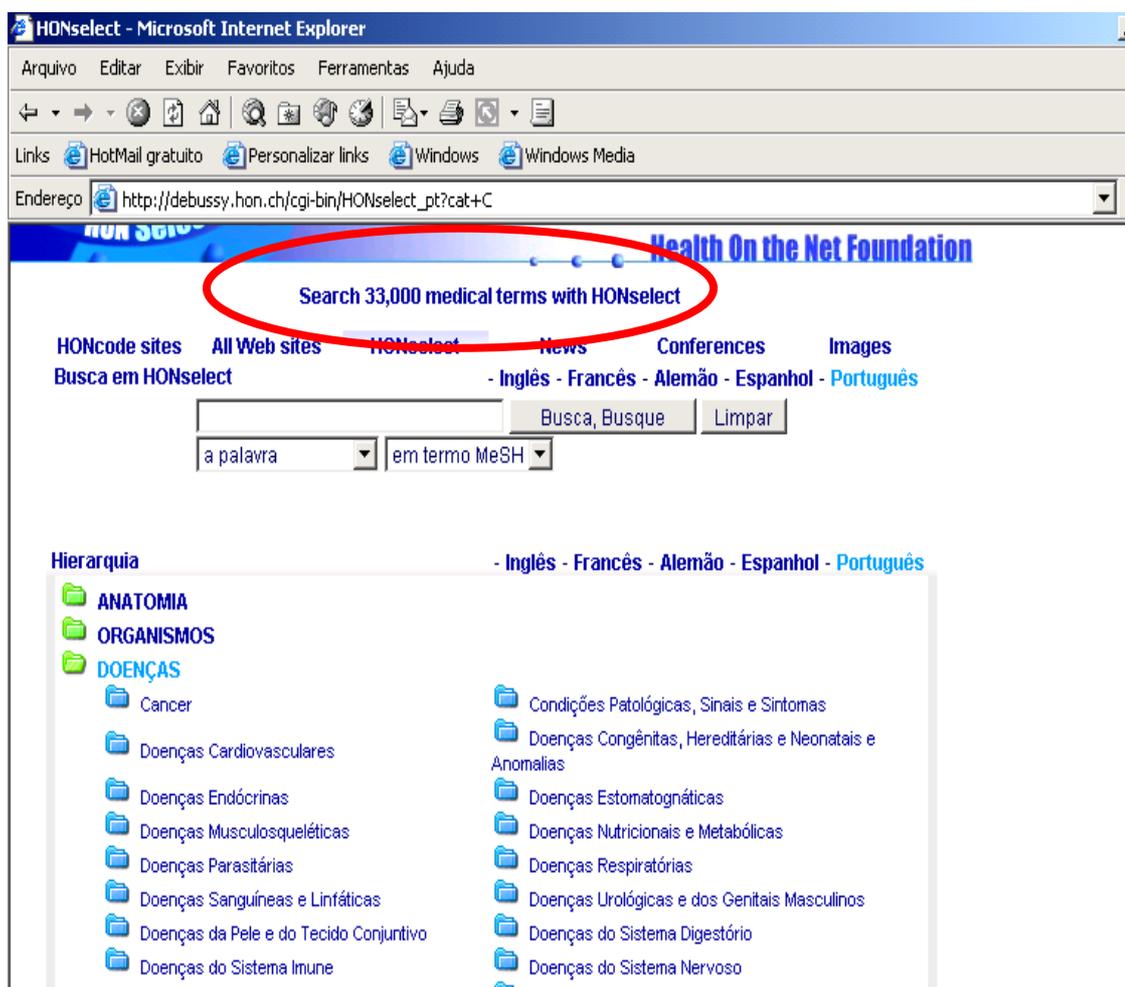


Figura 19– Página mostrando a quantidade de termos.  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Apoiado nos aspectos teóricos elencados anteriormente, é possível apontar que o *HONselect* é um macrotesauros, pois relaciona assuntos de várias áreas, constituindo objeto de microtesauros especializados.

**e) Tipo de relacionamento**

Como pode ser verificado nas figuras 20, 21 e 22 o tesouro HONselect trabalha com os três tipos de relacionamentos, pois apresenta comparações por semelhanças e diferenças entre os conceitos, quer seja para defini-los, nomeá-los ou ordená-los favorecendo, dessa forma, a recuperação da informação por parte do usuário que, mesmo não possuindo prática no manuseio para busca, o faz de forma clara e facilitada pelos elementos expostos.

The screenshot displays the HONselect search interface. At the top, it says "Search 33,000 medical terms with HONselect". Below this, there are navigation links: "HONcode sites", "All Web sites", "HONselect", "News", "Conferences", and "Images". A search bar contains the word "Vinho" (circled in red). Below the search bar, there are dropdown menus for "a palavra" and "em termo MeSH". A "Busca, Busque" button and a "Limpar" button are also visible. Below the search bar, there is a dropdown menu for "Conteúdo sobre 'Vinho':". Underneath, there are language options: "- Inglês - Francês - Alemão - Espanhol - Português". The word "Vinho" is displayed in red. Below it, the definition is shown: "Definição: Fermented juice of fresh grapes or of other fruit or plant products used as a beverage." Below the definition, there is a section titled "Veja também:" (circled in red) with a link to "Vitis".

Figura 20– Página mostrando a relação associativa.  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

A figura acima mostra uma busca sendo realizada a partir da palavra vinho. O sistema apresenta a definição do mesmo e remete, através da relação associativa, a um termo

alternativo. Esse tipo de estruturação facilita a busca para o usuário, pois o sistema remete automaticamente a outras expressões existentes relacionados ao mesmo assunto.

A relação de equivalência é, conforme exposto anteriormente, a que trabalha com o termo preferido e o não preferido, onde duas ou mais palavras são consideradas, para fins de indexação, como referentes ao mesmo conceito.

The screenshot shows the HONselect search interface. At the top, it says "Search 33,000 medical terms with HONselect". Below this are navigation links: "HONcode sites", "All Web sites", "HONselect", "News", "Conferences", and "Images". There is a search bar containing "Vitis" and a dropdown menu set to "a palavra". To the right of the search bar are buttons for "Busca, Busque" and "Limpar". Below the search bar is a dropdown menu for "Conteúdo sobre 'Vitis:'".

Below the search bar, there are language options: "Hierarquia" followed by "- Inglês - Francês - Alemão - Espanhol - Português".

The main content area displays the term "Vitis" in red. Below it is a definition: "Definição: A plant genus in the family VITACEAE, order Rhamnales, subclass Rosidae. It is a woody vine cultivated worldwide. It is best known for grapes, the edible fruit and used to make WINE and raisins." Below the definition are synonyms: "Sinónimos: Uvas / Uva-Passa / Videira / Uvas-Passa /".

Below the synonyms is the section "Veja também:" with a link to "Vinho".

At the bottom, there is a section "Veja relacionado:" with a link to "Vitaceae", which is circled in red in the original image.

Figura 21– Página mostrando a relação equivalência.  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

A figura 21 destaca um tipo de relação de equivalência viabilizada pelo tesouro em estudo. Utiliza o termo **veja relacionado**, para direcionar o usuário a busca do assunto procurado, ou seja, o instrumento permite que o usuário observe também outras opções para localizar a informação que busca.

The screenshot shows the HONselect website interface. At the top, it says "HON Select" and "Health On the Net Foundation". Below that, it prompts to "Search 33,000 medical terms with HONselect". There are navigation links for "HONcode sites", "All Web sites", "HONselect", "News", "Conferences", and "Images". A language selection menu is visible with options: "- Inglês - Francês - Alemão - Espanhol - Português". The search input field contains "Vitaceae" and has buttons for "Busca, Busque" and "Limpar". Below the search field, there are dropdown menus for "a palavra" and "em termo MeSH". A section titled "Hierarquia" also has language options. The main content area displays "Vitaceae" in red, followed by a definition: "Definição: A plant family of the order Rhamnales, subclass Rosidae, class Magnoliopsida, best known for the VITIS genus, the source of grapes." Below the definition, it lists "Sinónimos: Vitáceas /". A section titled "Termo(s) específico(s):" is circled in red and lists "Ampelopsis", "Cissus", and "Vitis" with small icons next to each name.

Figura 22– Página mostrando a relação hierárquica.  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Na figura 22 pode ser visualizado a relação hierárquica, que se baseia em graus ou níveis de superordenação ou subordinação, destacando membros ou partes do termo. Tal elemento é indicativo de novas buscas que poderão ser mais efetiva e apresentar maior relevância àquilo que é investigado.

*f) Notas esclarecedoras*

A nota de escopo ou esclarecedora é utilizada para indicar o conceito de um termo. Em alguns casos, antes da expressão, é apresentada a sigla NE, o que significa *não é obrigatório*.

The screenshot shows the HONselect website interface. At the top, there is a logo for 'HON Select' and 'Health On the Net Foundation'. Below this, a search bar contains the text 'Vinho'. To the right of the search bar are buttons for 'Busca, Busque' and 'Limpar'. Below the search bar, there are two dropdown menus: 'a palavra' and 'em termo MeSH'. A dropdown menu below that shows 'Conteúdo sobre "Vinho:"'. Below this, there is a 'Hierarquia' section with language options: '- Inglês - Francês - Alemão - Espanhol - Português'. The main content area displays the word 'Vinho' in red, followed by a definition: 'Definição: Fermented juice of fresh grapes or of other fruit or plant products used as a beverage.' This definition is circled in red. Below the definition, there is a 'Veja também:' section with a link to 'Vitis'.

Figura 23– Página mostrando nota esclarecedora.

FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Como é possível verificar na figura 23, no lugar da sigla NE encontra-se a palavra **DEFINIÇÃO**, neste caso tal função é igual a da Nota de Escopo.

### g) Morfologia das palavras

O que se verifica na figura 24, em relação a forma de entrada das palavras para busca e localização do termo é que o tesaurus *HONselect* aceita que seja utilizado termos no plural, porém os remetem para o singular no momento da busca, direcionando para algumas opções da solicitação desejada, ou seja, o termo no singular. Isto demonstra a preocupação, por parte dos especialistas responsáveis, em facilitar o manuseio em vista da recuperação desejada.

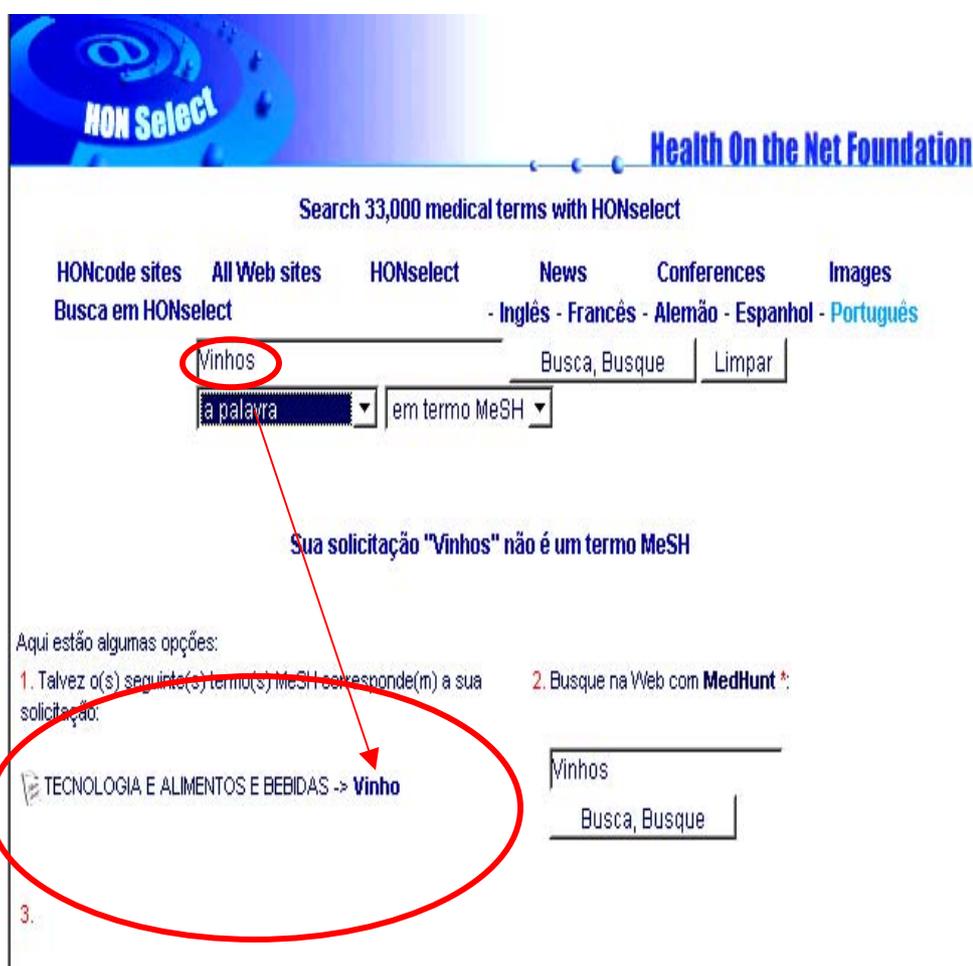


Figura 24– Página mostrando a morfologia das palavras.  
 FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Tal elemento também é indicativo da ordenação da terminologia de modo a favorecer uma ampla recuperação, assegurando o grau de precisão como exposto anteriormente.

### ***h) Idiomas***

A diversidade de idiomas existentes neste tesaurus representa uma ajuda fundamental para viabilizar o processo de comunicação. Desta forma pode-se verificar na figura 25, que o tesaurus *HONselect* adota o inglês, francês, alemão, espanhol e português, demonstrando o alto grau de preocupação com o usuário que pode acessar no idioma que tem mais afinidade ou mesmo verificar a variância da terminologia, elemento próprio da linguagem documentária, como já exposto.



Figura 25 - Página mostrando os idiomas.  
FONTE: [http://www.hon.ch/HONselect/index\\_pt.html](http://www.hon.ch/HONselect/index_pt.html)

Ainda no que tange a grande oferta de idiomas, é possível também inferir que a preocupação com o usuário se faz presente quando este tesaurus oferece seus produtos através de cinco línguas denotando a intencionalidade de se constituir em um instrumento universalmente utilizado.

## 2) Tesouro Cadeia Alimentícia - TCA

### a) Caracterização

O Tesouro Cadeia Alimentícia – TCA foi elaborado pela Informação e Tesouro - INFOTHES, visando disseminar e transferir conhecimento e informação sobre alimentação e nutrição. A temática do TCA foi escolhida pela atualidade do assunto e por envolver vários tipos de agentes sociais, seja como: produtores, distribuidores, especialistas, elaboradores de políticas públicas, consumidores.

A elaboração do TCA tem como origem a construção do Microtesauros Alimentos para o Serviço Nacional de Aprendizagem Industrial – SENAI, em 1999, e a atuação no Centro de Documentação e Informação da ONG Instituto Polis que atua em uma área voltada para política de segurança alimentar e nutricional.

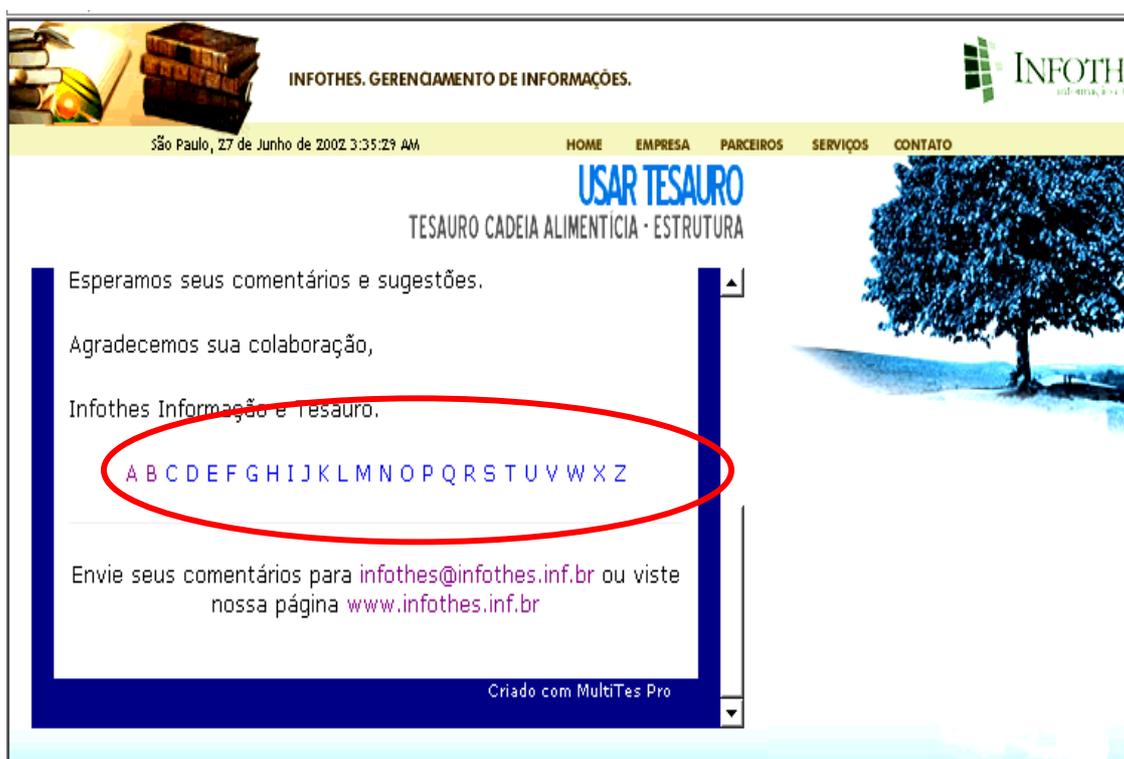


Figura 26– Forma de apresentação do TCA.  
FONTE: <http://www.thesaurus.eti.br/>

A figura 26 expressa a página inicial do tesouro, mostrando seus detalhes estruturais e focando a letras do alfabeto, onde o usuário deverá selecionar a letra inicial da palavra ou termo que será de seu interesse na busca. Na página principal também podem ser encontrados *links* onde o usuário poderá obter informações a respeito da empresa, parceiros, serviços disponíveis e um endereço eletrônico para eventuais comentários e/ou sugestões a respeito do mesmo.

### b) Forma de apresentação

Conforme pode ser visualizado na figura 27, a forma de entrada e a grafia dos termos (palavras ou expressões) utilizados seguem o padrão dos dicionários da língua portuguesa e aparecem em ordem alfabética, facilitando a busca das palavras, uma vez que o usuário já está familiarizado com as entradas encontradas nos dicionários.

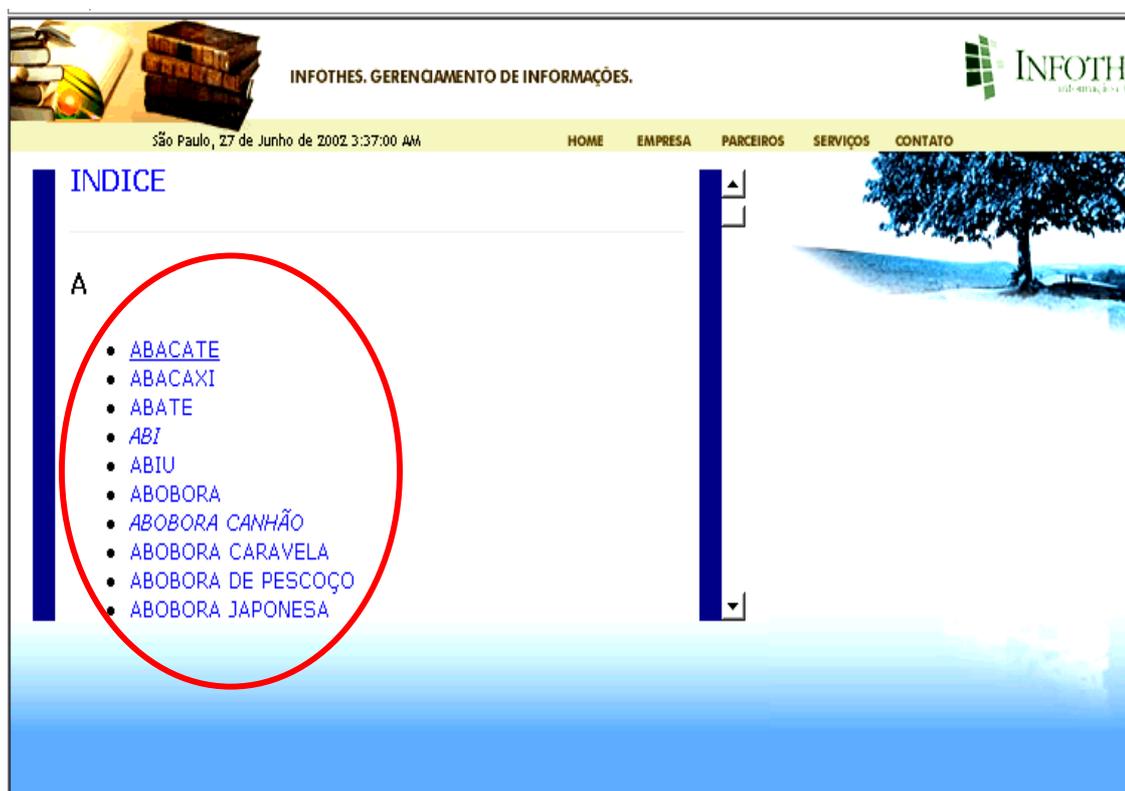


Figura 27 - Forma de apresentação do TCA.  
FONTE: <http://www.thesaurus.eti.br/>

c) *Índice e/ou introdução*

O índice é uma ferramenta que permite o usuário localizar uma informação sem a necessidade de acessar diretamente um documento/informação. São listas alfabéticas de nomes de pessoas, materiais, geográficos, etc.

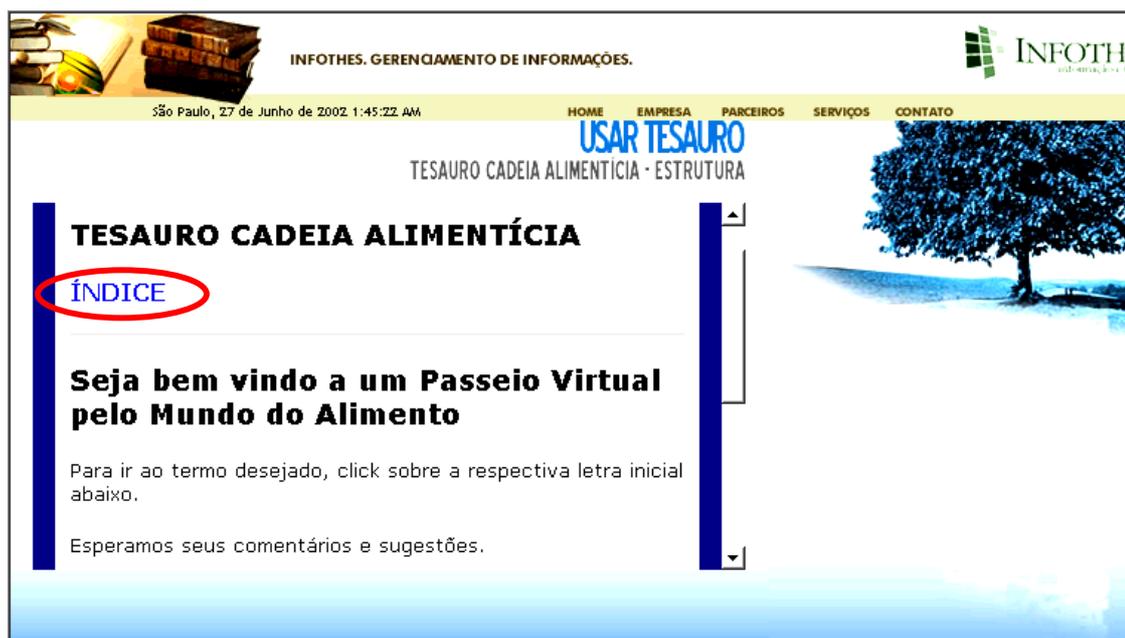


Figura 28 - Índice do Tesouro Cadeia Alimentícia  
FONTE: <http://www.thesaurus.eti.br/>

Como pode ser verificado na figura 28, no TCA consta um índice que orienta o usuário a manuseá-lo com facilidade. A existência de tal mecanismo, seguido da explicação sobre a forma como o termo desejado pode ser localizado bem como a existência expressa de um canal de diálogo com o usuário (Esperamos seus comentários e sugestões) são indicativos da busca da integração almejada com os utilizadores de modo a favorecer uma boa recuperação.

Verifica-se que além de orientar o usuário, o índice elaborado em ordem alfabética no TCA remete o usuário ao termo desejado, bastando para isso clicar sobre a respectiva letra inicial da palavra, conforme pode ser visualizado na figura 29.



Figura 29 – Índice alfabético do Tesauro Cadeia Alimentícia.

FONTE: <http://www.thesaurus.eti.br/>

Na figura acima também é possível observar que no momento em que o usuário seleciona a letra e o termo, este é sublinhado para indicar o que está sendo pesquisado. A barra de rolagem, a esquerda, viabiliza o conhecimento dos termos elencados pelo tesauros.

*d) Tamanho*

Em relação à quantidade de termos disponibilizados pelo TCA, não foi encontrada nenhuma fonte que informasse precisamente tal variável de análise, impossibilitando um exame sobre a dimensão do tesouro.

*e) Tipo de relacionamento*

Conforme pode ser verificado na figura 30, o TCA está estruturado para interagir com todos os relacionamentos no campo das idéias, onde conceitos são comparados e representados por relações hierárquicas e associativas e no campo semântico (comparação de palavras) e são representados pela relação de equivalência.

The screenshot shows the website 'USAR TESAURO' with the following content:

INFOTHES. GERENCIAMENTO DE INFORMAÇÕES.

INFOTHE  
informações em

São Paulo, 27 de Junho de 2002 2:58:16 AM

HOME EMPRESA PARCEIROS SERVIÇOS CONTATO

USAR TESAURO

TESAURO CADEIA ALIMENTICIA - ESTRUTURA

NC: A utilização da biotecnologia tem possibilitado o surgimento de produtos de ponta em todas as áreas: plantas geneticamente modificada, vacinas, medicamentos, anticorpos, enzimas e hormônios, entre outros.

TG: PRODUÇÃO DO ALIMENTO

TE: ENGENHARIA GENÉTICA  
MELHORAMENTO GENÉTICO

TR: ALIMENTO TRANSGÊNICO  
DNA  
ENZIMA  
LONGA VIDA

Figura 30 - Página mostrando os tipos de relacionamentos.  
FONTE: <http://www.thesaurus.eti.br/>

A figura acima também permite inferir que o tesauro adota uma terminologia técnica, da área da Ciência da Informação, para expor os tipos de relacionamentos que oferecem, o que pode não favorecer, em um primeiro momento um fácil entendimento e assimilação da estrutura pelo usuário.

*f) Notas esclarecedoras*

Como pode ser verificado na figura 31, o TCA apresenta um elevado número de exposições de notas esclarecedoras, que aparecem seguida aos termos constantes do tesauro. Em alguns casos, foram acrescentadas notas complementares com outras informações sobre o conceito.

The screenshot shows a website interface for 'INFOTHES. GERENCIAMENTO DE INFORMAÇÕES.' with a navigation menu (HOME, EMPRESA, PARCEIROS, SERVIÇOS, CONTATO) and a date/time stamp (São Paulo, 27 de Junho de 2002 2:22:41 AM). The main content area is titled 'USAR TESAURO' and 'TESAURO CADEIA ALIMENTICIA - ESTRUTURA'. A section titled 'BIOTECNOLOGIA' is highlighted with a red oval. The text within the oval reads: 'NE: Ramo da ciência que pesquisa e promove a transferência de genes de uma espécie para outra, a fim de atribuir a esta última características naturais da primeira.' and 'NC: A utilização da biotecnologia tem possibilitado o surgimento de produtos de ponta em todas as áreas: plantas geneticamente modificada, vacinas, medicamentos, anticorpos, enzimas e hormônios, entre outros.' Below the oval, classification codes are listed: 'TG: PRODUÇÃO DO ALIMENTO' and 'TE: ENGENHARIA GENÉTICA'. A tree image is visible on the right side of the page.

Figura 31 – Notas esclarecedoras.  
FONTE: <http://www.thesaurus.eti.br/>

Cabe destacar também, ainda observando a figura 31, que as definições expostas são destacadas pelas barras azuis e a seta que permite a rolagem facilitar o contato com os demais elementos da estrutura existente.

*g) Morfologia das palavras*

Como pode ser verificado na figura 32, todas as palavras constantes do TCA tem como forma de entrada o singular.



Figura 32 – Morfologia das palavras.  
FONTE: <http://www.thesaurus.eti.br/>

### h) Idiomas

Em relação ao idioma, o TCA, conforme pode ser visualizado na figura 33, somente emprega o português, restringindo seu acesso àqueles usuários que dominam a língua portuguesa.

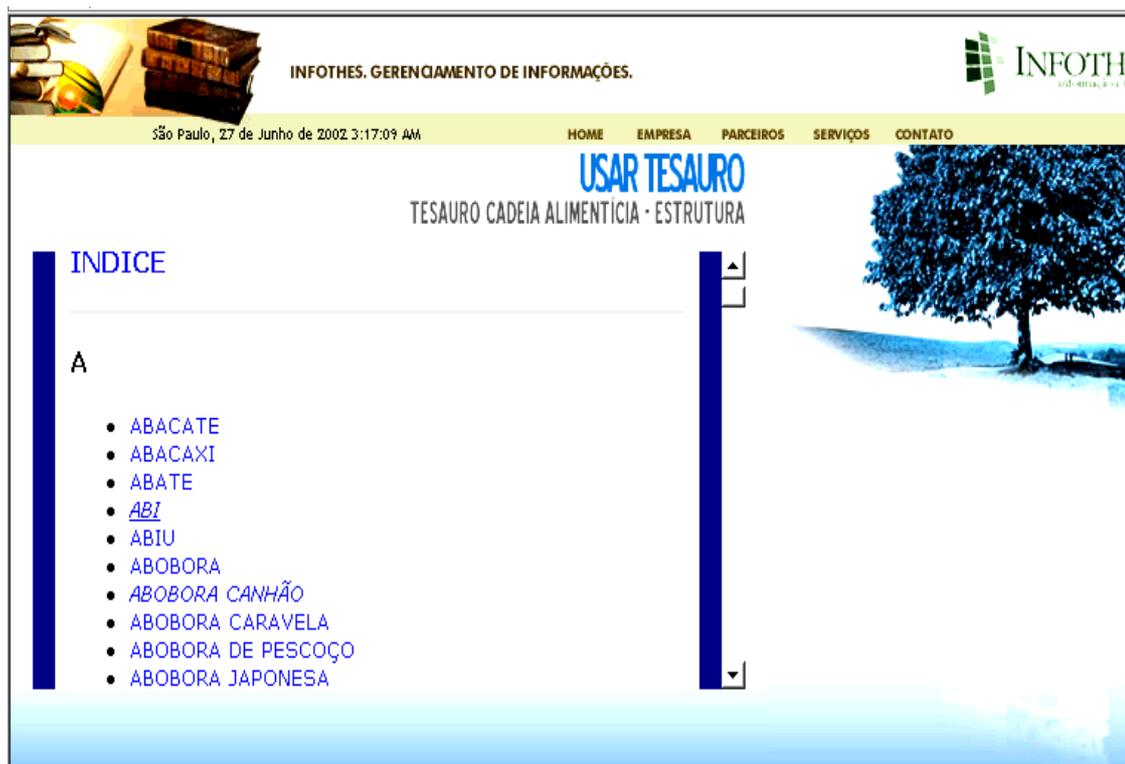


Figura 33 – Página mostrando o idioma do TCA.

FONTE: <http://www.thesaurus.eti.br/>

No que tange o aspecto do idioma, é importante destacar que a oferta é direcionada para o público que pretende atingir tendo em vista o objetivo de sua construção como exposto anteriormente. Tal situação demonstra que a elaboração busca atender as expectativas do público para o qual se dirige.

### 3) Glossário de Biotecnologia em Alimentação e Agricultura

#### a) Caracterização

O Glossário de Biotecnologia em Alimentação e Agricultura constitui uma versão espanhola do documento *Glossary of Biotechnology for Food and Agriculture* publicado pela Organização das Nações Unidas para Agricultura e Alimentação - FAO, em 2001. Esta publicação suscitou grande interesse mundial por tratar de áreas importantes ao desenvolvimento humano, como a biotecnologia, tão em evidência nos dias atuais. Levando em conta esta relação tão positiva e, com a finalidade de disponibilizar para um público mais amplo, a FAO, procedeu e disponibilizou a tradução do mesmo em outros idiomas, com o objetivo de facilitar a tradução de textos relativos a biotecnologia.



ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA AGRICULTURA Y LA ALIMENTACIÓN

**LA BIOTECNOLOGÍA en LA ALIMENTACIÓN Y LA AGRICULTURA**

- [Declaración de la FAO sobre biotecnología](#)
- [Documentos FAO](#)
- [Actividades de la FAO](#)
- [Foro Electrónico](#)
- [Revisiones sectoriales](#)
- [Noticias](#)
- [Eventos](#)
- **[Glosario](#)**
- [Contactos](#)
- [Documentos sobre las políticas nacionales](#)
- [FAO-BioDeC](#)
- [Enlaces](#)

**Glosario de biotecnología para la agricultura y la alimentación**

(actualmente disponible en [árabe](#), [español](#), [francés](#) e [inglés](#))

**PÁGINA INICIAL**

[Árabe](#)

[Inglés](#)

[Francés](#)

[Español](#)

**AYUDA**

**Acerca del glosario**

**Listado alfabético de términos**

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#)  
[N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

**Buscar un término**

[\[Búsqueda\]](#) [\[Borrar\]](#)

Figura 34 – Página principal do glossário de biotecnologia em alimentação e agricultura.  
 FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

Na figura 34, pode ser visualizado a página principal, bem como sua estrutura. O glossário apresenta uma lista atualizada de termos relacionados a biotecnologia, incluindo a engenharia genética. Tem como objetivo oferecer fontes de referências seguras e adequadas aos pesquisadores da área. Como pode ser verificado, o tesouro em foco disponibiliza ainda através de *links*, uma grande oferta de informações como: documentos relacionados a instituição, atividades da FAO, notícias, eventos, etc.

### b) Forma de apresentação

Em um tesouro, os termos e suas inter-relações podem ser apresentadas de várias maneiras, porém o mais comum são serem expostas em ordem alfabética.

Y LA ALIMENTACIÓN

## LA BIOTECNOLOGÍA *en* LA ALIMENTACIÓN Y LA AGRICULTURA

- [Declaración de la FAO sobre biotecnología](#)
- [Documentos FAO](#)
- [Actividades de la FAO](#)
- [Foro Electrónico](#)
- [Revisiones sectoriales](#)
- [Noticias](#)
- [Eventos](#)
- [Glosario](#)
- [Contactos](#)
- [Documentos sobre las políticas nacionales](#)
- [FAO-BioDeC](#)
- [Enlaces](#)

**Glosario de biotecnología para la agricultura y la alimentación**  
(actualmente disponible en [árabe](#), [español](#), [francés](#) e [inglés](#))

**PÁGINA INICIAL**

- [Árabe](#)
- [Inglés](#)
- [Francés](#)
- [Español](#)

**AYUDA**

**Acerca del glosario**

**Listado alfabético de términos**

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#)  
[N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

**Buscar un término**

[Búsqueda] [Borrar]

Figura 35 – Página mostrando a lista de entrada em ordem alfabética.  
 FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

No caso do tesouro em análise, como pode ser verificado na figura 35, as entradas dos termos são feitas por lista em ordem alfabética, o que facilita a busca por parte do usuário em duas situações a saber: a primeira que esta acostumado a lidar com esse tipo de entrada; e a segunda, o tempo de busca é mínimo, pois basta clicar na letra inicial do termo que procura.

Na figura 36 verifica-se uma listagem de termos iniciados pela letra B, todos em ordem alfabética. Acima das palavras consta como resultado da busca, 105 termos iniciando por esta letra. Isto implica analisar que no tesouro em questão existe um alto grau de exaustividade.

**LA BIOTECNOLOGÍA** en LA ALIMENTACIÓN Y LA AGRICULTURA

[Declaración de la FAO sobre biotecnología](#)  
[Documentos FAO](#)  
[Actividades de la FAO](#)  
[Foro Electrónico](#)  
[Revisiones sectoriales](#)  
[Noticias](#)  
[Eventos](#)  
[Glosario](#)  
[Contactos](#)  
[Documentos sobre las políticas nacionales](#)  
[FAO-BioDeC](#)  
[Enlaces](#)

**Glosario de biotecnología para la agricultura y la alimentación**  
 (actualmente disponible en [árabe](#), [español](#), [francés](#) e [inglés](#))

[PÁGINA INICI](#)

[AYUDA](#)

105 Resultados de la búsqueda

[BABS](#)  
[BAC](#)  
[Bacillus thuringiensis](#)  
[bacilo](#) [Bacteria](#)  
[bacteria](#)  
[bacteria lisogénica](#)  
[bactericida](#)  
[bacteriocina](#)  
[bacteriostático](#)  
[bacteriófago](#)  
[bacteriófago atemperado](#)  
[baculovirus](#)  
[banco](#)  
[banco de clones](#)  
[banco de genes](#)  
[banco de genes en campo](#)  
[bando cromosómico](#)  
[barnasa](#)  
[basal](#)

Figura 36 – Página mostrando a lista de entrada em ordem alfabética – letra B.  
 FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

***c) Índice e/ou introdução***

O item índice e/ou introdução não foi localizado. Porém, no caso específico do tesauro analisado, a existência de tal elemento pode ser desnecessária, visto que a estrutura da página principal do tesauros como foi visualizado anteriormente, é bastante completa, podendo ser caracterizado como um portal e, dessa forma, facilitar a recuperação da informação pelo usuário.

***d) Tamanho***

No tesauro de Biotecnologia foram encontrados mais de 3.000 (três) mil termos com suas respectivas definições. Mediante essa análise, verifica-se que sua capacidade é extensa e com isso aumenta seu nível de exaustividade.

**e) Tipo de relacionamento**

Na figura 37, pode ser verificado a forma como o tesauros permite o cruzamento dos termos pesquisados, com outros do mesmo assunto. Assim, o usuário/pesquisador tem várias oportunidades de busca, que mesmo não sendo a principal, chegará ao resultado com alto grau de pertinência.

**biotecnología**

**DEFINICIÓN**

- 1. "Cualquier aplicación tecnológica que utiliza sistemas biológicos, organismos vivos, o algunos de sus derivados para crear o modificar productos o procesos para usos específicos" (Convenio de Diversidad Biológica). 2. "Interpretado en sentido más estricto, [...] el conjunto de diferentes tecnologías moleculares tales como la manipulación y transferencia de genes, el tipado de ADN y la donación de plantas y animales" (Declaración de la FAO sobre biotecnología)

**TÉRMINO INGLÉS CORRESPONDIENTE**

- [biotechnology](#)

**TÉRMINO FRANCÉS CORRESPONDIENTE**

- [biotechnologie](#)

**TÉRMINOS RELACIONADOS**

- [ADN](#)

[Comentarios sobre este término](#) [Sugerencia de un nuevo término para el Glosario](#)

Figura 37 – Página mostrando tipos de relacionamentos.

FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

Cabe ainda destacar, como observado na figura 37, que o usuário tem acesso a um comentário sobre o termo e a seu correspondente em outro idioma, denotando a preocupação pela ampla compreensão da terminologia pesquisada.

Pode ser visualizado na figura 38 que, além da definição do termo selecionado o tesouro disponibiliza ao usuário o termo correspondente na língua francesa e espanhola. Abaixo como sinalizado, o instrumento de busca do tesouro em análise, remete a outros termos relacionados a palavra em questão. Tal elemento facilita a busca para o usuário, que demandará de menor tempo para recuperar a informação solicitada.

**ADN**

**DEFINITION**

- Abr. de ácido desoxirribonucleico. Largo polímero de desoxirribonucleótidos. El ADN constituye el material genético de la mayoría de los organismos y orgánulos que se conocen; normalmente se encuentra formando una doble hélice, aunque algunos genomas virales contienen ADN de una sola cadena y otros, ARN de una o de doble cadena. Véase: par de bases, código genético

**CORRESPONDING FRENCH TERM**

- [ADN](#)

**CORRESPONDING SPANISH TERM**

- [ADN](#)

**RELATED TERMS**

- [ARN](#), [código genético](#), [doble hélice](#), [ADN bicatenario](#), [desoxirribonucleótido](#), [proteína de unión al ADN monocatenario](#), [ADN recombinante](#), [Mitochondrial DNA](#), [replicación del ADN](#)

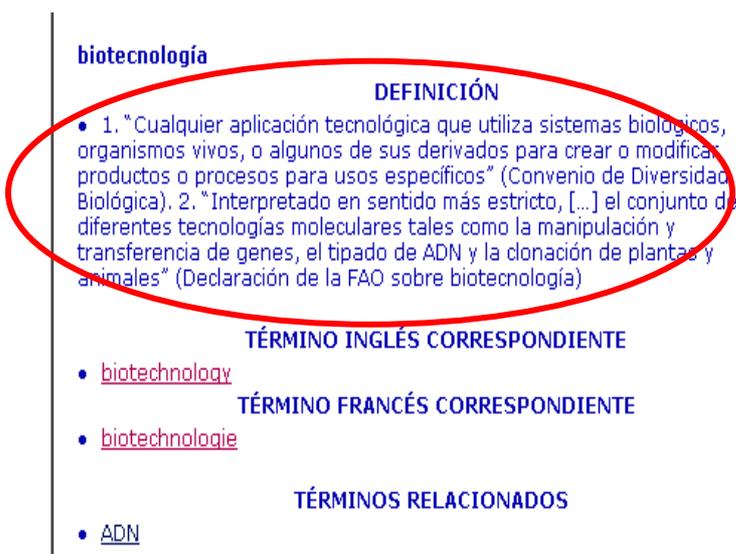
[comments on this term](#) [suggest new terms](#)

Figura 38 – Página mostrando tipos de relacionamentos.

FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

*f) Notas esclarecedoras*

Na figura 39 pode se verificar que o tesaurus apresenta a palavra pesquisada e em seguida expõe sua definição. Remete também a outros termos que tenham relação em outros idiomas, favorecendo dessa forma a busca para o usuário.



**biotecnología**

**DEFINICIÓN**

- 1. "Cualquier aplicación tecnológica que utiliza sistemas biológicos, organismos vivos, o algunos de sus derivados para crear o modificar productos o procesos para usos específicos" (Convenio de Diversidad Biológica). 2. "Interpretado en sentido más estricto, [...] el conjunto de diferentes tecnologías moleculares tales como la manipulación y transferencia de genes, el tipado de ADN y la clonación de plantas y animales" (Declaración de la FAO sobre biotecnología)

**TÉRMINO INGLÉS CORRESPONDIENTE**

- [biotechnology](#)

**TÉRMINO FRANCÉS CORRESPONDIENTE**

- [biotechnologie](#)

**TÉRMINOS RELACIONADOS**

- [ADN](#)

Figura 39 – Página mostrando notas esclarecedoras.

FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

O elemento destacado expressa a existência de notas esclarecedoras como itens que buscam favorecer a compreensão do termo que está sendo pesquisado.

### ***g) Morfologia das palavras***

Em relação a morfologia dos termos verifica-se, na figura 40, que todos os descritores aparecem na forma singular. Verifica-se que este tipo de entrada facilita a busca, porém o usuário terá que, antes de iniciá-la saber como utilizar o instrumento de busca.

[banco](#)  
[banco de clones](#)  
[banco de genes](#)  
[banco de genes en campo](#)  
[bandeo cromosómico](#)  
[barnasa](#)  
[basal](#)  
[base](#)  
[base nitrogenada](#)  
[basípeto](#)  
[basófilo](#)  
[beta-ADN](#)  
[beta-galactosidasa](#)  
[beta-glucuronidasa](#)  
[beta-lactamasa](#)  
[beta-sitosterol](#)  
[BEV](#)  
[BFGF](#)  
[biental](#)  
[bio-](#)  
[bioacumulación](#)  
[bioanálisis](#)  
[bioaumentación](#)  
[biocatálisis](#)  
[biochip](#)  
[biocontrol](#)  
[bioconversión](#)  
[hindenradahle](#)

Figura 40– Página mostrando morfologia das palavras.

FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

Ainda em relação à forma, conforme pode se visualizado na figura 41, quando solicitado através do campo de busca um termo no plural, o instrumento de controle informa que não há nenhum registro. Dessa forma o usuário terá que efetivar a busca, no caso do tesouro em questão, utilizando a palavra no singular.

**LA BIOTECNOLOGÍA** *en* LA ALIMENTACIÓN Y LA AGRICULTURA

- Declaración de la FAO sobre biotecnología
- Documentos FAO
- Actividades de la FAO (actualmente disponible en [árabe](#), [español](#), [francés](#) e [inglés](#))
- Foro Electrónico
- Revisiones sectoriales
- Noticias
- Eventos
- Glosario**
- Contactos
- Documentos sobre las políticas nacionales
- FAO-BioDeC
- Enlaces

**Glosario de biotecnología para la agricultura y la alimentación**

[PÁGINA INICIAL](#)

[AYUDA](#)

**No se ha encontrado ningún registro**

Listado alfabético de términos

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#)  
[N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

**Buscar un término**

[\[Busqueda\]](#) [\[Borrar\]](#)

Figura 41– Página mostrando morfologia das palavras.  
 FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

A inexistência de mecanismos de recuperação da informação que favoreçam uma dinâmica forma de busca, pode constituir uma barreira para uso deste instrumento de linguagem controlada.

Na figura 42, pode-se observar que a busca pela palavra bactéria, quando solicitada na forma singular, aponta para dois resultados (bactéria e bactéria lisogénica), também na forma singular. Dessa forma, confirma-se que todos os termos deverão ser recuperados entrando por essa categoria.

The image shows a screenshot of the FAO glossary search interface. On the left is a navigation menu with links such as 'Declaración de la FAO sobre biotecnología', 'Documentos FAO', 'Actividades de la FAO', 'Foro Electrónico', 'Revisiones sectoriales', 'Noticias', 'Eventos', 'Glosario', 'Contactos', 'Documentos sobre las políticas nacionales', 'FAO-BioDeC', and 'Enlaces'. The main content area is titled 'Glosario de biotecnología para la agricultura y la alimentación' and includes a language note: '(actualmente disponible en árabe, español, francés e inglés)'. There are links for 'PÁGINA INICIAL' and 'AYUDA'. The search results section, titled 'Resultados de la búsqueda', shows two results: 'bacteria' and 'bacteria lisogénica'. Below the results is a section 'Listado alfabético de términos' with alphabetical navigation links from A to Z. At the bottom, there is a search box containing the text 'bacteria' and buttons for '[Búsqueda]' and '[Borrar]'. Red circles highlight the search results and the search input field.

Figura 42– Página mostrando forma de busca.

FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

A forma ampla de recuperação por mais de um termo pode, ao mesmo tempo que oferece outras opções, dificultar a precisão uma vez que a consulta a nota esclarecedora só ocorrerá quando outro *link* for acionado.

## h) Idiomas

Conforme pode ser visualizado na figura 43, o Glossário de Biotecnologia encontra-se disponível em árabe, espanhol, francês e inglês, representando elemento fundamental para a ampla comunicação no âmbito da ciência e tecnologia.

**LA BIOTECNOLOGÍA en LA ALIMENTACIÓN Y LA AGRICULTURA**

- [Declaración de la FAO sobre biotecnología](#)
- [Documentos FAO](#)
- [Actividades de la FAO](#)
- [Foro Electrónico](#)
- [Revisiones sectoriales](#)
- [Noticias](#)
- [Eventos](#)
- **Glosario**
- [Contactos](#)
- [Documentos sobre las políticas nacionales](#)
- [FAO-BioDeC](#)
- [Enlaces](#)

**Glosario de biotecnología para la agricultura y la alimentación**  
(actualmente disponible en [árabe](#), [español](#), [francés](#) e [inglés](#))

[PÁGINA INICIAL](#)

[Árabe](#)

[Inglés](#)

[Francés](#)

[Español](#)

[AYUDA](#)

[Acerca del glosario](#)

[Listado alfabético de términos](#)

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#)  
[N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

[Buscar un término](#)

[\[Búsqueda\]](#) [\[Borrar\]](#)

Figura 43– Página mostrando os idiomas.

FONTE: [http://www.fao.org/biotech/index\\_glossary.asp?lang=es](http://www.fao.org/biotech/index_glossary.asp?lang=es)

Ainda em relação a figura 43, cabe destacar que o Glossário esta disponibilizado não somente no espanhol, francês e inglês como encontrados em outros tesouros, mas o idioma árabe, uma novidade nesse instrumento de controle terminológico. Verifica-se, dessa forma, a preocupação por parte da FAO em disseminar a informação de forma exaustiva e em nível mundial.

As análises até aqui efetuadas visaram nortear o exame quanto à estrutura individualizada de cada tesouro relacionado à área da biotecnologia, itens relevantes que direcionarão e contribuirão de forma segura, a compreensão da análise global, que será exposta a seguir.

## 4.2 ANÁLISE GLOBAL

A biotecnologia promove impactos em diversos setores industriais por meio de técnicas inovadoras e pela promoção de verdadeira revolução no tratamento de doenças, no uso de novos medicamentos para aplicação humana e animal, na multiplicação e reprodução de espécies vegetais e animais, no desenvolvimento e melhoria de alimentos, na utilização sustentável da biodiversidade, na recuperação e tratamento de resíduos, dentre outras áreas com potencial crescente de aplicação.

Dessa forma, o que se verifica é que muitas instituições de pesquisa estão produzindo e disponibilizando informações relevantes para contribuir com o desenvolvimento da área tendo em vista que a pesquisa é ampla e demanda de largo espaço.

Assim, a grande oferta de informações em biotecnologia associada a transdisciplinaridade da área, exige que os saberes sejam disponibilizados de forma ordenada visando favorecer sua localização e recuperação. Isto é possível de ser observado pela oferta de instrumentos de linguagem controlada que favoreçam a recuperação da informação.

O quadro 3, a seguir, sintetiza as análises efetuadas no corpus selecionado.

VARIÁVEIS SELECIONADAS	TESAUROS ANALISADOS		
	Alimentos e Bebidas HONselect	Tesouro Cadeia Alimentícia - TCA	Glossário de Biotecnologia para a Agricultura e Alimentação
Forma de apresentação (alfabética, sistemática ou gráfica)	Alfabética	Alfabética	Alfabética
Índice e/ou introdução	Possui	Possui	Não encontrado
Tamanho (nº de descritores)	33.000	Não encontrado	+ de 3.000
Tipos de Relacionamento (equivalência, hierárquicas, associativas).	Equivalência, hierárquicas, associativas.	Equivalência, hierárquicas, associativas.	Equivalência, hierárquicas, associativas.
Notas esclarecedoras	Possui	Possui	Possui
Morfologia das palavras (singular, plural)	Singular	Singular	Singular
Idiomas	Espanhol, alemão, inglês, francês e espanhol.	Português	Árabe, espanhol, francês e inglês.

Quadro 3 – Tesauros avaliados com suas variáveis

A partir da análise e interpretação dos dados coletados, pode-se inferir, a partir das variáveis eleitas, as seguintes considerações.

1) Forma de apresentação (alfabética, sistemática ou gráfica)

Com relação a forma de apresentação dos tesauros analisados, todos os termos, preferidos ou não-preferidos, estão organizados em uma só seqüência alfabética. A

modalidade alfabética é a forma tradicional de apresentação dos tesouros. Esta ordem é prática e permite que se localize rapidamente o termo desejado.

## 2) Índice e/ou introdução

Dois deles, o Alimentos e Bebidas *HONselect* e o Tesouro Cadeia Alimentícia – TCA, possuem índice que permite ao usuário ser auxiliado em relação a utilização, características e aplicação dos mesmos. Contudo, no Glossário de Biotecnologia para a Agricultura e Alimentação, tal instrumento não foi encontrado. É oportuno destacar que a existência do índice é de fundamental importância no tesouro pois, é ele quem vai auxiliar o usuário em relação ao seu manuseio, explicitando suas características no campo de aplicação.

A prevalência pelo emprego do elemento analisado demonstra que sua apresentação não deve ser desconsiderada tendo em vista o favorecimento do uso.

## 3) Tamanho (nº de descritores)

Dos tesouros analisados, o de Alimentos e Bebidas *HONselect*, é o maior pois comporta trinta e três mil termos relacionados a área de alimentos, bebidas e saúde. O segundo é o Glossário de Biotecnologia para a Agricultura e Alimentação com mais de três mil descritores. No Tesouro Cadeia Alimentícia - TCA não foi encontrado um número especificado de termos.

Todos os tesouros analisados podem ser enquadrados em relação ao nível de especificidade como macrotesouros, pois, abrangem uma ampla gama de assuntos relacionados entre si, de diversas áreas do conhecimento humano.

#### 4) Tipos de Relacionamento (equivalência, hierárquicas, associativas)

Todos os tesouros analisados adotam os três tipos de relacionamentos, o de equivalência, hierárquico e associativos. Tal condição demonstra que esses elementos são fundamentais para o bom desenvolvimento e recuperação da informação por parte dos usuários. São as relações entre os termos que conferem a um tesouro uma multiplicidade de usos, que vai da função até o auxílio na efetiva recuperação da informação.

#### 5) Quantidade de notas esclarecedoras registradas por tesouros

Todos os tesouros analisados possuem notas esclarecedoras. As notas esclarecedoras são instrumentos importantes dentro da estrutura de um tesouros, pois, permite ao usuário ter maior clareza de significado em relação ao termo pesquisado e possibilita ainda relacionar com outros termos.

#### 6) Morfologia das palavras (singular, plural)

Quanto a morfologia das palavras, todos trabalham utilizando termos no singular. O objetivo principal do tesouro é dar assistência ao usuário (pesquisador ou indexador) de maneira que ele consiga encontrar o termo que represente um determinado significado para o que se procura, ou seja, com a ajuda do tesouro, o usuário no momento da busca poderá identificar termos alternativos, que permitirá descrever a informação contida no documento de forma mais adequada.

Destaca a recuperação no singular para o termo buscado no plural, o que representa uma preocupação com a eficiência da recuperação.

## 7) Idiomas

As línguas representam um subsídio fundamental para a comunicação. Dessa forma pode-se verificar que o tesaurus de Alimentos e Bebidas *HONselect* é o que mais se destaca pois trabalha com uma diversificação bem ampla, pois se encontra traduzido para o espanhol, alemão, inglês, francês e espanhol.

Em segundo aparece o Glossário de Biotecnologia para a Agricultura e Alimentação se encontra em árabe, espanhol, francês e inglês. O idioma árabe infere surpresa, pois é uma novidade em relação aos tesaurus analisados. Já o Tesouro Cadeia Alimentícia - TCA se apresenta somente na língua portuguesa, fazendo com que seu acesso seja restrito aqueles usuários que dominem o idioma em questão. Ainda em relação a língua, os três tesaurus analisados podem ser enquadrados como multilíngues, pois, comportam descritores em mais de um idioma.

A diversidade de idioma facilita o acesso ao qual o tesouro se propõe uma vez que está disponível para acesso em rede.

### **Implicações**

Para construção de um tesaurus em biotecnologia para a Amazônia, em função das análises realizadas é fundamental observar:

#### a) Forma de apresentação

A biotecnologia é uma área multidisciplinar, como destacado e a construção de uma linguagem controlada deve levar em consideração os aspectos inerentes a sua cadeia produtiva. Por ser multidisciplinar, o usuário de um tesaurus em biotecnologia também tem uma formação variada.

Deste modo, o instrumento deverá constituir várias formas de apresentação de modo a atender as expectativas do usuário final, podendo adotar uma ordem sistemática.

b) Índice/Introdução

Ter índice indicando o modo de entrada para favorecer uma precisa recuperação. Em vista da heterogeneidade de possíveis usuários, convém que se adote uma linguagem clara e precisa.

c) Tamanho

Por ser uma área complexa onde vários assuntos estarão relacionados entre si, o tesauro em biotecnologia será considerado um macrotesauro. Nesse caso não será possível definir a quantidade máxima de termos que serão empregados no mesmo.

d) Tipo de Relacionamento

Como já exposto, os tesauros analisados apresentam os três tipos de relacionamento visando favorecer uma boa qualidade de uso.

Desse modo, o tesauro em biotecnologia comportará os relacionamentos de equivalência, hierárquico e associativo.

e) Notas Esclarecedoras

Como já exposto, o tesauro de biotecnologia envolve diversas áreas do conhecimento como a química, biologia, física, etc., e por ter esse alto grau de complexidade, se faz necessário utilizar as notas esclarecedoras ou de escopo, para que o usuário se situe e tenha clareza dos termos utilizados.

f) Morfologia das palavras

Quanto a morfologia das palavras, todos os tesouros analisados trabalham utilizando termos no singular. Dessa forma o tesouros em biotecnologia trabalhará utilizado o termo no singular, porém como verificado no tesouros *HONselect*, o usuário poderá entrar na forma plural que o mesmo através do instrumento de busca o remeterá para o termo no singular.

g) Idioma

A Amazônia é fruto de estudo de diversos grupos de pesquisa no mundo.

Biotecnologia é uma área de ponta com pesquisas fortalecidas através da integração de pesquisadores em âmbito mundial.

Desse modo o tesouros em construção será disponibilizado em três idiomas, o português, inglês e espanhol.

## CONCLUSÃO

Neste espaço expressa-se, em sua primeira parte, as conclusões obtidas a partir da coleta de dados realizada em atendimento aos objetivos a que se propôs a pesquisar. Na segunda parte, registra-se as recomendações para trabalhos futuros em relação ao tema tesouros e biotecnologia.

O objetivo geral deste estudo se constituiu no exame da composição das linguagens documentárias presentes nos tesouros em biotecnologia. Em termos específicos, objetivou-se investigar as peculiaridades estruturais de três tesouros em biotecnologia, tendo como subárea produtos alimentícios e fármacos; comparar a composição estrutural dos tesouros analisados; apontar elementos que favoreçam a geração de um tesouros em biotecnologia para a Amazônia.

Um trabalho desta natureza não apresenta conclusões fechadas, apenas evidencia alguns pontos problemáticos, sinaliza possibilidades, prepara para futuras análises. Com este posicionamento, os dados obtidos, ao longo deste estudo, permite tecer algumas considerações.

Os sistemas de informações devem ser gerenciados de forma que o seu ciclo de vida inclua análise, projeto, implementação, evolução operacional, deterioração e substituição, possibilitando assim a criação de um sistema novo. Entretanto, as metodologias de sistemas de informação foram desenvolvidas para favorecer uma abordagem metódica de análise, no qual, o planejamento estratégico de sistemas de informações seja um componente essencial da organização, e que, uma vez implantado exigirá manutenção constante.

A análise conceitual vista aqui, diz respeito à interpretação do que o usuário está a procura, enquanto a tradução é disponibilização para atender as suas necessidades; pois os termos selecionados do vocabulário do sistema, está na representação dessa necessidade de informação, que poderá ser considerada uma estratégia de busca.

Dentre as linguagens controladas, os tesouros constituem um dos meios mais utilizados para a indexação e recuperação de documentos e/ou informações. Sendo assim, deve-se verificar se atendem às necessidades atuais de seus clientes consumidores de informação.

O objetivo principal do tesouro é oferecer assistência ao usuário (pesquisador ou indexador) de maneira que ele consiga encontrar o termo que represente um determinado significado para o que se procura, ou seja, com sua ajuda o usuário, no momento da busca, poderá identificar termos alternativos, que permitirá localizar a informação contida no documento de forma adequada.

Após o estabelecimento de um sistema de informação e da definição do tipo de linguagem de recuperação, pode-se considerar que para um bom desempenho, o tesouro precisa de uma estrutura complexa e estar adaptado ao campo de interesse e das necessidades dos usuários. Deve permitir alterações necessárias, sem a qual certas áreas do conhecimento não poderiam ser desenvolvidas com maiores detalhes. No entanto, caso não exista uma linguagem de recuperação adequada, será necessário rever o que foi desenvolvido, sem contudo, desprezar o já existente, que poderá fornecer subsídios para reordenação.

Assim, o volume de informação disponível atualmente, em especial nas áreas científicas e tecnológicas, fez despertar para a necessidade de normalização e controle dos vocabulários, principalmente devido ao aparecimento de ambigüidades terminológicas, o que vem dificultando o rápido acesso à informação armazenada.

Enfim, a importância do tesouro como recurso auxiliar em um sistema de recuperação da informação, e o grande volume de dados disponibilizados através das tecnologias da informação faz verificar que esses dados precisam ser tratados de forma padronizada e com metodologia capaz de facilitar a geração de informações pertinentes.

Por fim, pretende-se que os resultados deste estudo possam subsidiar medidas que contribuam para a elaboração e manutenção do tesouro em biotecnologia da Amazônia, e dessa forma favorecer o desenvolvimento da região através de suas riquezas naturais.

## REFERÊNCIAS

- AZEVEDO NETTO, Carlos Xavier de. **Signo, sinal, informação: as relações de construção e transferência de significados**. Disponível em: [www.documents/biblioteca.com.br](http://www.documents/biblioteca.com.br). Acesso em: 18/04/2005.
- BAEZA-YATES; RIBEIRO NETO, B. *Modern information retrieval*. [S.l.]: Addison-Wesley, 1999.
- CABRÉ, M. T. **La terminologia: teoria, metodologia, aplicaciones**. Barcelona: Antártida/Empúries, 1993.
- \_\_\_\_\_. Terminologie et dictionnaires. **Meta**, v.39, n.4, p.589-597, 1994.
- \_\_\_\_\_. Terminologie ou terminologies? Spécialité linguistique ou domaine interdisciplinaire? **Meta**, v.36, n.1, p.56-63, 1991.
- CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.14, n.2, p.221-241, set. 1985.
- CESARINO, Maria Augusta da Nóbrega. Sistemas de recuperação da informação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.14, n.2, p.157-68, set. 1985.
- CINTRA, A. M. M. et al. Do termo ao descritor. **Rev. Comun. & Artes**, São Paulo, v.17, n. 28, p.75-82, jan./abr. 1994.
- CHARTIER, Roger. **A ordem dos livros: leitores, autores e biblioteca na Europa entre os séculos XIV e XVIII**. Brasília: UNB, 1998.
- COELHO NETTO, José Teixeira. **A biblioteca como modelo de sistema de comunicação**. R. Bras. Bibliotecon. e Doc, São Paulo, n. 11, v. 1/2, p. 29-32, jan./jun.1978.
- COHEN, D. M. **O consumidor da informação documentária: o usuário de sistemas documentários visto sob a lente da análise documentária**. São Paulo: ECA/SP, 1995. (Dissertação de Mestrado).
- CONSELHO DE INFORMAÇÃO SOBRE BIOTECNOLOGIA. **Glossário de biotecnologia**. 2005. Disponível em: [www.cib.org.br](http://www.cib.org.br). Acesso em: 30.06.2006.
- CURRÁS, Emília. **Tesauros: linguagens terminológicas**. Brasília: IBICT, 1995.
- DODEBEI, Vera Lúcia Doyle. **Tesauro: linguagem de representação da memória documentária**. Niterói: Intertexto; Rio de Janeiro: Interciência, 2002.
- ECO, Umberto. **Conceito de texto**. São Paulo: T. A. Queiroz, 1984.
- \_\_\_\_\_. Marco Pólo e o unicórnio. In: \_\_\_\_\_. **Kant e o ornitorrinco**. Rio de Janeiro: Record, 1988.

\_\_\_\_\_. **Tratado geral de semiótica**. São Paulo: Perspectiva, 1980.

FERNEDA, Edberto. **Recuperação de informação**: análise sobre a contribuição da Ciência da computação para a Ciência da Informação, São Paulo: [s.n.], 2003.

FOSKETT, A. C. **A abordagem temática da informação**. São Paulo: Polígono, 1973.

\_\_\_\_\_. **A study of the role of categories in a thesaurus for education documentation**. Strasbourg: Council of Europe, 1972.

FOSKETT, D. **Subject and information analysis**. New York: Marcel Dekker, 1985.

GOMES, Hagar Espanha (Org.). **Manual de elaboração de tesouros monolíngues**. Brasília: PNB, 1990.

GONZALES DE GOMEZ, M. N. A representação do conhecimento e o conhecimento da representação: algumas questões epistemológicas. **Ciência da Informação**, Brasília. V.22, n.3, p.217-222, set./dez., 1993.

GUINCHAT, Claire; MENO, Michel. **Introdução geral às ciências e técnicas da informação e documentação**. Trad. Míriam V. da Cunha. 2.ed. aum. Brasília, IBICT, 1994.

GUTIÉRREZ, Antonio García. **Teoría da la indización**: nuevos parámetros de investigación. Transinformação, Campinas, v.1, n.2, p.147-159, maio/ago., 1989.

[HTTP://WWW.FAO.ORG/BIOTECH/INDEX\\_GLOSSARY.ASP?LANG=ES](http://www.fao.org/biotech/index_glossary.asp?lang=es). Acesso em: 04/6/2007.

[HTTP://WWW.HON.CH/HONSELECT/INDEX\\_PT.HTML](http://www.hon.ch/honselect/index_pt.html). Acesso em: 04/6/2007.

[HTTP://WWW.THESAURUS.ETI.BR/](http://www.thesaurus.eti.br/). Acesso em: 04/6/2007.

LANCASTER, F. W. **Construção e uso de tesouro**: curso condensado. Brasília: IBICT, 1987.

\_\_\_\_\_. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 1993.

\_\_\_\_\_. **Vocabulary control for information retrieval**. Washington: Information Resources Press, 1972.

LARA, M. L. G. (a). Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. **Ciência da Informação**, Brasília, v.22, n.3, p.223-226, set./dez. 1993.

\_\_\_\_\_. (b). Linguagens documentárias, instrumentos de mediação e comunicação. **Rev. Bras. Biblioteconomia e Documentação**, São Paulo, v.26, n.1/2, p.72-80, jan. /jun. 1993.

\_\_\_\_\_. (c). **A representação documentária**; em jogo a significação. São Paulo: [s.n.], 1993. (Dissertação de Mestrado ECA/USP).

\_\_\_\_\_. O unicórnio (o rinoceronte, o ornitorrinco...), a análise documentária e a linguagem documentária. **Data Gama Zero – Rev. de Ciência da Informação**, Brasília, v.2, n.6, p.1-12, dez. 2001.

LE COADIC, Yves-François. **A ciência da informação**. Brasília, Briquet de Lemos, 1996.

LESKA, B. M. Theoretical and methodological problems og therminology. In: INTERNATIONAL SYMPOSIUM, 1979, Warsaw. **Proceedings...** [S.l.: s.n.], 1979.

LIMA, Vânia Mara Alves. **Terminologia, comunicação e representação documentária**. São Paulo: [s.n.], 1998. (Dissertação de Mestrado).

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.1-18, jan. 2002.

LOPES, L. C. **A informação e os arquivos: teorias e prática**. Niteroi: EDUFSCAR; EDUFF, 1996.

MIRANDA, A. L. C. Globalizacion y sistemas de informacion: nuevos paradigmas y nuevos desafios. **Ciência da Informação**, Brasília, v.3, n.3, p.308-313. 1996.

MOREIRA, Alexandra. **Tesauros e ontologias: estudo de definições presentes na literatura das áreas das Ciências da Computação e da Informação, utilizando-se o método analítico-sintético**. 2003. 150f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.

NASCIMENTO, Luiz Antonio do. Proposta de um sistema de recuperação de informação para extranet de projeto. São Paulo: Escola Politécnica da Universidade de São Paulo, 2004.

PIGNATARI, Décio. **Informação, linguagem, comunicação**. São Paulo: Ateliê Editorial, 2002.

PINTO, Virgínia Bentes. Indexação documentária: uma forma de representação do conhecimento registrado. **Perspectiva em Ciência da Informação**, v.6, n.2, p. 223-234, jul./dez. 2001.

ROBREDO, Jaime. **Da ciência da informação revisitada aos sistemas humanos de informação**. Brasília: Thesaurus; SSRR Informações, 2003.

SCHAFF, A. **Introdução à semântica**. Rio de Janeiro: Civilização Brasileira, 1968.

SHENK, David. **Data smog: surviving the information glut**. New York: Harper Edge, 1997.

SVENONIUS, Elaine. The intellectual foundation of information organization. Cambridge, MIT Press, 2000.

TÁLAMO, M. F. G. M. **Linguagem documentária**. São Paulo: APB, 1997.

UNESCO. **Guia para la creación y desenvolvimiento de thesaurus multilingues, técnicos y científicos, destinados a la recuperación de dados.** Montevideo: CINTERFOR Documentation, 1973.

URDICIAIN, Blanca Gil. Evaluación semántica y estructural de tesauros. **Rev. General de Información y Documentación**, v.8, n.2, p.193-199, 1998.

VÉRON, E. **Ideologia, estrutura e comunicação.** São Paulo: Cultrix, 1970.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)