

Programa Geral de Informação (PGI) e UNISIST
UNESCO

Construção e uso de tesauros:

curso condensado

MCT
IBICT

5032-6 546/87

INSTITUTO
DE DESENVOLVIMENTO
EM CIÊNCIAS E TECNOLOGIA

Elaboração de tesouros

Programa Geral de Informação (PGI) e UNISIST
UNESCO

50326

BIBLIOTECA
DO
L. B. I. C. T

Construção e uso de tesouros: curso condensado

F. W. LANCASTER

025.4
L245
et. 3
CDU

MCT

Ministério da Ciência e Tecnologia

 **CNPq**

CONSELHO NACIONAL DE DESENVOLVIMENTO
CIENTIFICO E TECNOLÓGICO

IBICT



INSTITUTO BRASILEIRO
DE INFORMAÇÃO
EM CIÊNCIA E TECNOLOGIA

© UNESCO 1985

Direitos de tradução cedidos ao Instituto Brasileiro de
Informação em Ciência e Tecnologia (IBICT)

Tradução, com adaptações, do original inglês *Thesaurus
Construction and Use. A Condensed Course*, elaborado por
F.W. Lancaster (PGI-85/WS/11).

Traduzido por Cesar Almeida de Menezes Silva
Revisto por Odilon Pereira da Silva

MEN: 4915

Instituto Brasileiro de Informação em Ciência e Tecnologia COBIBLIOTECA
Proc.
Livr.
Preço Doação - Ce\$ 1,00
Nº registro 546/87 - 07.10.87

PHL 003955

Lancaster, F. W.

Construção e uso de tesouro : curso condensa-
do / F. W. Lancaster; trad. de César Almeida de
Menezes Silva; rev. de Odilon Pereira da Silva.
— Brasília : IBICT, 1987.

106 p. (UNESCO; PGI-85/WS/11)

1. Tesouro-construção. 2. Tesouro-uso. I. Sil-
va, César Almeida de Menezes, trad. II. Silva,
Odilon Pereira da, rev. III. Título.

CDU 025.45

Instituto Brasileiro de Informação em Ciência e Tecnologia
(IBICT)

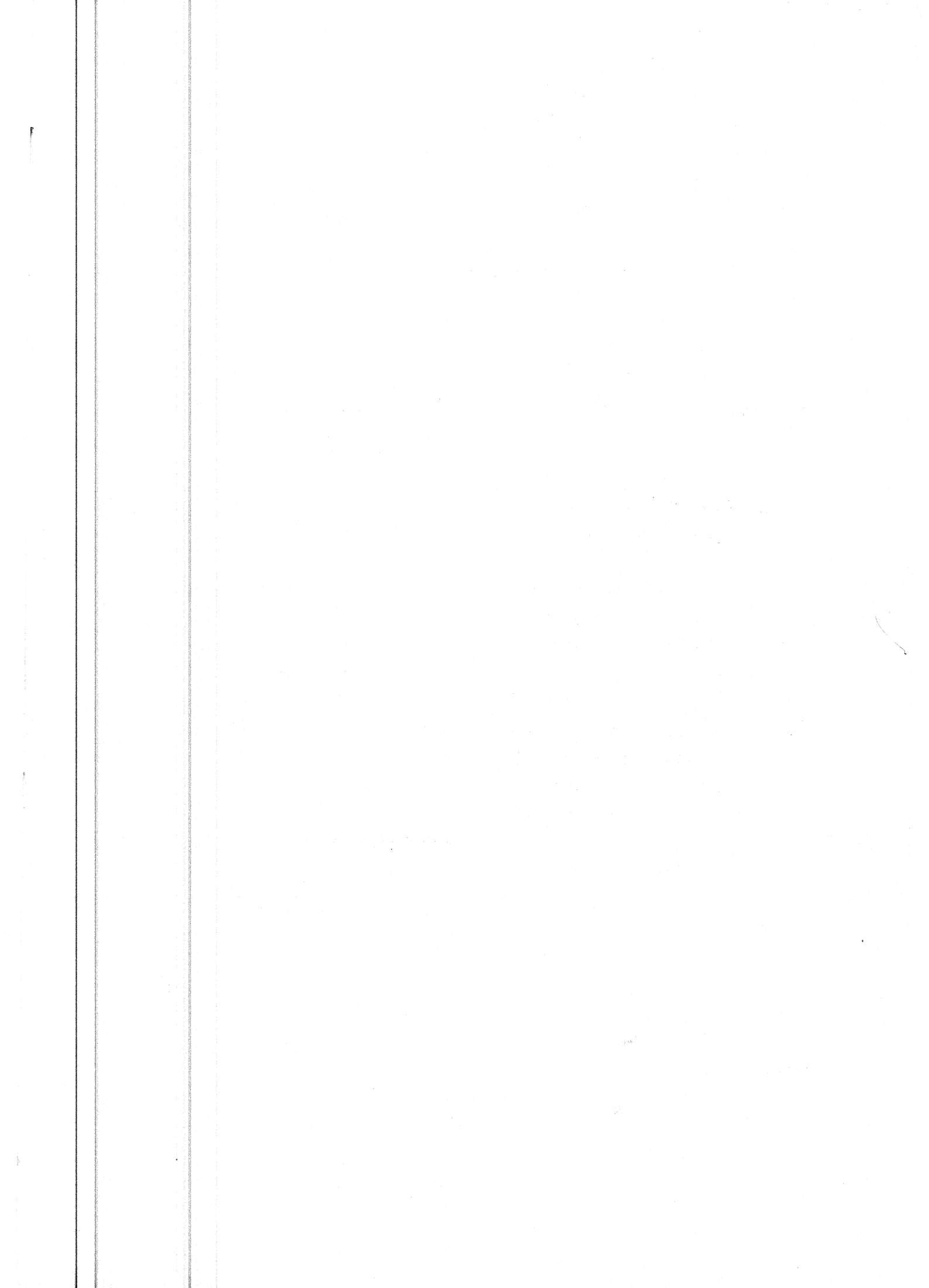
SCN - Quadra 2 - Bloco K
70710 Brasília, DF

Tel. (061) 225 96 77 Telex 612481 CICT BR

BIBLIOTECA
DO
L. B. I. C. T.

SUMÁRIO

INTRODUÇÃO	5
UNIDADE	
1. Finalidade do controle de vocabulário	9
2. Componentes básicos de um vocabulário controlado....	14
3. Coleta dos termos	18
4. Organização dos termos	22
5. A relação hierárquica	26
6. A relação associativa	28
7. Características dos descritores.....	29
8. O vocabulário de entrada.....	33
9. Notas explicativas e identificadores	35
10. Formato e apresentação do tesauro	37
11. Crescimento e atualização.....	43
12. Utilização do computador.....	45
13. Fatores do vocabulário que afetam o desempenho dos sistemas de informação	51
14. Sistemas de linguagem natural	59
ILUSTRAÇÕES.....	63
BIBLIOGRAFIA	113



INTRODUÇÃO

A origem deste curso foi o Seminário Regional sobre Linguagens de Indexação, organizado sob o patrocínio da UNESCO pelo Centro Argentino de Información Científica y Tecnológica (CAICYT), e realizado em Buenos Aires, em agosto de 1978. O autor do presente volume foi o principal instrutor naquele seminário.

O CAICYT preparou, posteriormente, um curso, derivado do próprio seminário, sob a forma de ilustrações (transparências, na verdade) com o texto correspondente. Este curso foi publicado em 1981 (*Curso sobre Lenguajes de Indización y Construcción de Tesoros*. Publicação da UNESCO. PGI.81/WS/31).

Em 1984, a UNESCO firmou um contrato com o autor para preparar uma versão inglesa do referido curso, o que resultou neste volume.

O volume é constituído de duas partes principais: um conjunto de 84 "ilustrações", que poderão ser transformadas em diapositivos, ou transparências, e um texto para explicar e desenvolver as ilustrações. Difere consideravelmente da versão em espanhol. De um modo particular, concentra-se quase exclusivamente na construção e nas características dos tesoros, enquanto que o Seminário na Argentina se ocupou também de outras áreas.

O texto está dividido em 14 unidades temáticas, mas sem nenhuma preocupação em torná-las "simétricas": algumas são bastante curtas, outras bem mais desenvolvidas.

O curso não pretende ser um programa para autodidatas, conquanto possa ser usado com essa finalidade. Foi, de fato, preparado para servir a um instrutor — as ilustrações (transparências) sendo, então, usadas como recursos visuais, e o texto mais como um guia para o que poderia ser acrescentado nas aulas.

Não deverá ser considerado completo em si mesmo. Há inúmeros lugares em que o texto sugere tópicos que devem ser desenvolvidos pelo

instrutor (por exemplo: diferenças entre tesouros e listas de cabeçalhos de assuntos). É de se esperar que os instrutores acrescentem outros exemplos de sua própria lavra aos aqui oferecidos, sobretudo exemplos que ilustrem tesouros de especial interesse para os estudantes que frequentam determinados cursos. Além do mais, é importante que muitos exemplos de diferentes tipos de tesouros sejam postos à disposição dos estudantes, para manuseio e exame, e que cada estudante possua um exemplar da publicação da UNESCO, *Guidelines for the Establishment and Development of Monolingual Thesauri* (2.ed 1981).

Finalmente, um curso sobre elaboração de tesouros **deve** revestir-se de um cunho prático. De cada estudante, ou pequeno grupo de estudantes, deve-se exigir a compilação de um pequeno tesouro — de aproximadamente 200 termos — em alguma área de assunto.

O tesouro deve ser estruturalmente completo, no sentido de que deve incluir todos os tipos de componentes — TG, TE, TR, notas explicativas etc., e todas as relações devem indicar suas recíprocas. Além disso, deve o tesouro conformar-se com as exigências das Diretrizes da UNESCO, sob todos os aspectos. Com o propósito de ajudar os estudantes na tarefa de construir esses tesouros, será necessário supri-los com uma variedade de fontes — livros de referência, resumos etc. — de onde possam extrair os termos.

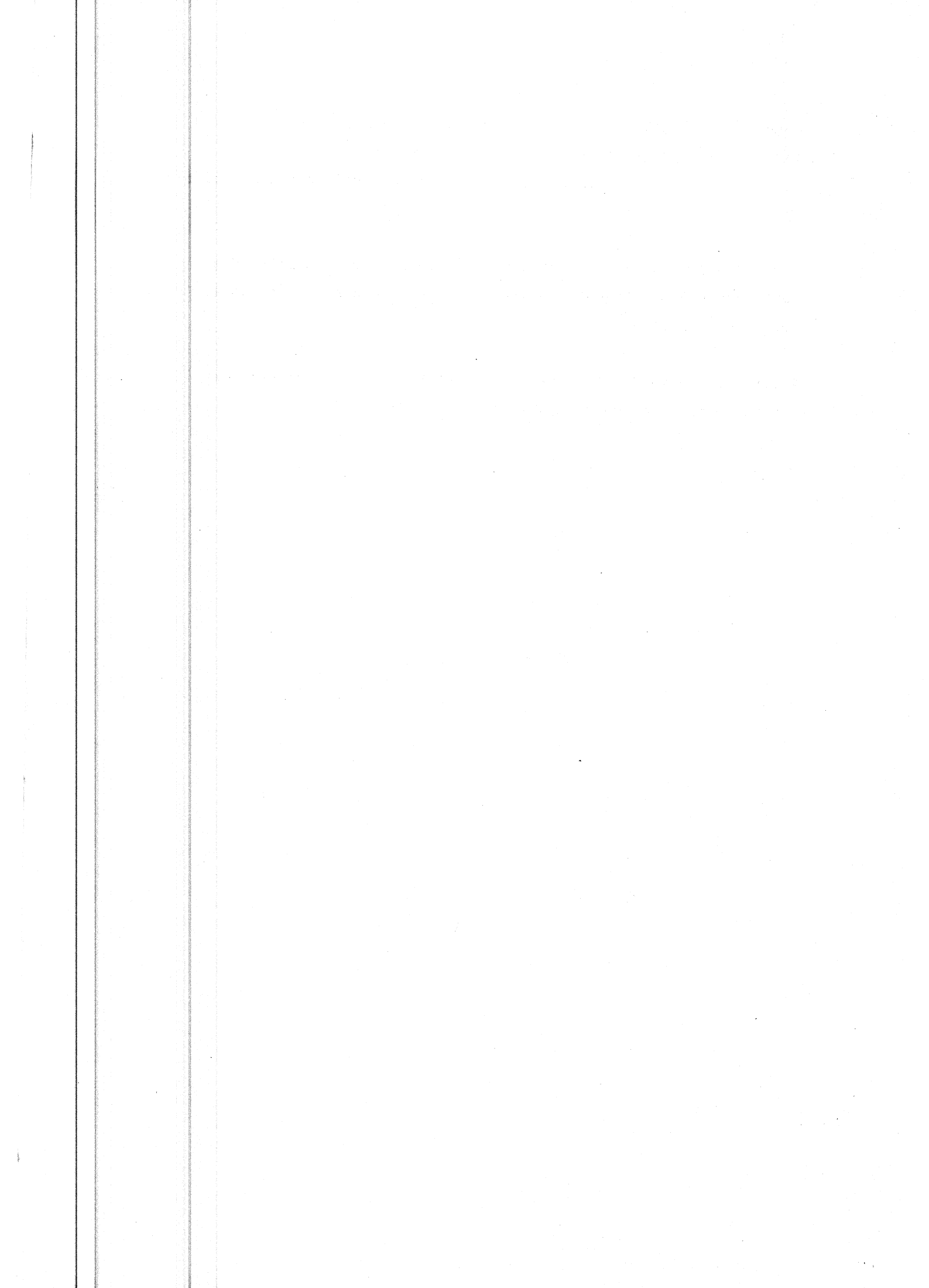
Este volume não é um livro-texto. Para mantê-lo relativamente pequeno, alguns tópicos tiveram tratamento um tanto superficial. Maior detalhamento em todos os tópicos que cobre, bem como em outros com os mesmos relacionados, pode ser obtido na obra de Lancaster. *Vocabulary Control for Information Retrieval* (2.ed., Arlington, Va., Information Resources Press, 1985), apresentada numa seqüência semelhante à adotada no presente curso.

Na maioria dos casos, as instruções e exemplos oferecidos neste curso seguem as Diretrizes da UNESCO. Todavia, o curso não se destina apenas a ilustrar o uso daquelas diretrizes, mas a cobrir a gama maior de assuntos relacionados com a elaboração e o uso dos tesouros.

Como decorrência, outras terminologias, processos e práticas, amplamente empregados, embora não inteiramente conformes às Diretrizes, são com frequência mencionados e apresentados a título de ilustração.

Esses desvios devem provocar discussão em classe a respeito das vantagens e desvantagens das diferentes abordagens e práticas em contextos diferentes.

O autor gostaria de agradecer à British Standards Institution a permissão para reproduzir páginas do *ROOT Thesaurus*.



UNIDADE 1. FINALIDADE DO CONTROLE DE VOCABULÁRIO

Objetivos

1. Ilustrar o papel que o vocabulário desempenha nos sistemas de recuperação da informação;
2. apresentar alguns dos problemas que soem ocorrer quando esse vocabulário não é controlado;
3. mostrar as funções básicas de um vocabulário controlado.

Ilustração 1

A expressão **recuperação da informação** é geralmente empregada para designar as atividades relacionadas com a localização de documentos (livros, periódicos, artigos, relatórios e outras formas) que tratam de um assunto particular, e um **sistema de recuperação da informação** consiste em um grupo de atividades e componentes destinados a facilitar o acesso ao conteúdo dos documentos.

Os componentes de maior importância são discutidos aqui. Do universo de documentos disponíveis, os considerados relevantes, tendo em vista os interesses de uma determinada comunidade, são selecionados e adquiridos por algum centro de informação (que pode ser uma biblioteca convencional).

A catalogação descritiva e a **catalogação por assunto**, ou **indexação por assunto**, são aplicadas aos documentos selecionados. Teoricamente, **catalogação por assunto** e **indexação por assunto** são a mesma coisa, mas a primeira em geral se refere à descrição do assunto de uma publicação em sua inteireza (por exemplo: livros, periódicos), enquanto a outra muito provavelmente se aplica a partes de publicações (por exemplo: artigos de periódicos). Neste curso, o termo **indexação** se refere à **indexação por assunto**, muito embora os princípios discutidos se apliquem também à **catalogação por assunto**.

O primeiro passo em indexação, a **análise conceitual**, envolve a decisão sobre de que trata o documento. O segundo passo, a **tradução**, envolve a

seleção de termos para representar o conteúdo do documento. Pode-se referir-se a eles como **termos de indexação** ou, visto que eles descrevem o assunto dos documentos, **descritores**.*

O conjunto completo de termos de indexação utilizados em um sistema de recuperação pode ser denominado **vocabulário** ou **linguagem de indexação** daquele sistema.

No centro do sistema de recuperação encontram-se duas bases de dados: uma consistindo dos próprios documentos, e a outra, das entradas que representam esses documentos e proporcionam o acesso aos mesmos sob a forma de termos de indexação selecionados para lhes representar o assunto. Essa segunda base de dados pode ser considerada como um índice do conteúdo da primeira.

Na maioria das bibliotecas, a base de dados de documentos consiste basicamente de livros arrumados em estantes, enquanto que o índice é provavelmente constituído por entradas em um fichário. Em outras situações, o índice estará automatizado (em fita magnética ou disco), ou sob a forma de microfilme, ou mesmo impresso como um livro convencional, embora possam existir outras alternativas.

Na ilustração, o termo **consultas** se refere a necessidades de informação oriundas do universo de usuários do sistema. Os consulentes estarão, via de regra, à procura de documentos, ou parte de documentos que tratem de determinado assunto, embora possam, ao invés, estar à procura de respostas para um problema factual.

No contexto de um centro de informação, o que ocorre com as consultas que chegam é semelhante ao que ocorre com os documentos recebidos.

* Diretrizes para indexação por assunto podem ser encontradas na norma internacional ISO DIS 5963: *Methods for examining documents, determining their subjects and selecting indexing terms.*

A **análise conceitual**, aqui, diz respeito à interpretação daquilo que o consulente na verdade está à procura, enquanto que **tradução** é a representação das necessidades do consulente, através de termos selecionados do vocabulário do sistema. A representação da necessidade de informação daí decorrente pode ser considerada uma **estratégia de busca**.

Quando se aplica essa estratégia de busca ao índice, aquelas entradas que preenchem os requisitos da estratégia são "recuperadas". Isso pode conduzir posteriormente à recuperação dos documentos correspondentes no acervo.

Deveria ficar evidente na Ilustração 1 que o vocabulário empregado em um sistema de recuperação é um componente de importância capital, uma vez que o assunto, tanto das consultas quanto dos documentos, deve ser representado por termos tomados a essa linguagem.

Na maioria dos sistemas de recuperação, o vocabulário empregado será um **vocabulário controlado**. Tal vocabulário nada mais é do que um conjunto limitado de termos que devem ser utilizados por indexadores e usuários. Em outras palavras, o vocabulário determinará os termos que podem ser empregados e os que não o podem. Além disso, os vocabulários controlados, em geral, apresentam alguma forma de "estrutura" (isto é, os termos encontram-se organizados de maneira a poderem evidenciar relações importantes), embora isso não seja elemento essencial à sua definição.

No contexto da recuperação da informação, o antônimo de **vocabulário controlado** é **linguagem natural**. Esta última significa que o assunto pode ser descrito por meio de quaisquer palavras ou frases tais como ocorrem nos próprios documentos, sem limitações. É possível operar um sistema de recuperação sem qualquer controle de vocabulário, o que será discutido adiante neste curso. Certos problemas, entretanto, ocorrem, quando não se impõe controle algum sobre o vocabulário, como se pode ver na Ilustração 2.

Ilustração 2

Todos os termos aqui podem ser vistos como de alguma forma relacionados com falhas nos metais. Alguns dizem respeito a tipos de falhas (por exemplo: fratura, ruptura), alguns a possíveis causas da falha (por exemplo: corrosão, fadiga), alguns ainda a propriedades dos metais que podem afetar sua tendência a apresentar falhas (por exemplo: ductibilidade, resistência), e assim por diante.

Os termos aparecem em ordem alfabética, da mesma forma como podem ser encontrados em um catálogo de assunto, ou num índice, onde eles estariam misturados com muitos outros a respeito de tópicos inteiramente diversos.

Para a pessoa em busca de informação relacionada com falhas nos metais, vários problemas vêm ilustrados nesse exemplo:

- a. alguns termos listados são sinônimos, ou quase-sinônimos: "falha" e "defeito", por exemplo;
- b. algumas palavras são ambíguas quando retiradas do contexto. Por exemplo: "ruptura" pode referir-se ao processo de fratura ou ao ato de violação de um contrato, no campo do direito, enquanto termos como "fadiga" e "resistência" podem ser ditos tanto dos seres humanos quanto dos materiais;
- c. termos relacionados quanto ao significado ficam muito distanciados uns dos outros na ordem alfabética.

Embora pareça impossível que alguém se disponha a fazer uma pesquisa sobre todos os aspectos da falha, é possível que alguém queira recuperar toda informação que trate de uma forma de falha, como, por exemplo, a fratura. A sequência alfabética, infelizmente, separa muitos termos atinentes ao assunto: defeito, ductilidade, falha, fragilidade, fratura, resistência, ruptura etc. A menos que esses termos sejam agrupados ou relacionados de alguma forma, fica muito difícil para um usuário tê-los a todos em mente.

Ilustração 3

É exatamente esse tipo de problema que o vocabulário controlado tenta solucionar:

- a. palavras ou frases sinônimas, ou quase-sinônimas, são fundidas: uma é adotada, e da outra é feito algum tipo de remissiva no vocabulário;
- b. palavras que têm muitos significados ou contextos podem aparecer mais de uma vez; podem-se usar explicações entre parênteses ou quaisquer outros recursos que evitem ambigüidade;
- c. termos com significados que guardem relação entre si são associados de alguma forma. São representados aqui elos entre alguns dos termos da Ilustração 2. Descrevem-se dois tipos de relações. O primeiro, a relação entre uma coisa e suas espécies (isto é, gênero e espécie): fratura e ruptura são sempre tipos de falhas. O segundo, uma relação mais transitória; por exemplo: o **processo** de falha pode ser induzido por outros processos, como a corrosão ou a fadiga, e pode ser influenciado por propriedades dos materiais, como a resistência e a fragilidade. A relação genérica, permanente, é às vezes denominada **paradigmática** ou **a priori**, enquanto a de natureza mais transitória (uma falha pode ser devida à corrosão, mas nem sempre, e da corrosão nem sempre resulta uma falha) é, às vezes, denominada **sintagmática** ou **a posteriori**.

Ilustração 4

Os objetivos do controle de vocabulário em um sistema de recuperação da informação podem ser sumariados assim:

- a. Promover representação consistente do assunto, pelos indexadores e consulentes (evitando por conseguinte a dispersão de materiais correlacionados) através do controle de sinônimos e quase-sinônimos, bem como pelo estabelecimento das distinções entre homógrafos;
- b. facilitar o desenvolvimento de uma busca exaustiva a respeito de algum tópico, através do agrupamento de termos cujos significados

estão paradigmática ou sintagmaticamente relacionados.

A indexação tende a ser mais consistente quando o vocabulário adotado é controlado. Há mais probabilidade de os indexadores coincidirem na escolha dos termos necessários à indexação de um tópico, quando os mesmos são relacionados em uma lista preestabelecida. O vocabulário controlado tende, pois, a fazer coincidir a linguagem dos indexadores com a dos pesquisadores.

Um vocabulário controlado não precisa ser mais do que um conjunto limitado de termos que devem ser usados por indexadores e pesquisadores. Na maioria dos casos, entretanto, é introduzida uma estrutura, de tal maneira que os termos, cujos significados estão inter-relacionados, são agrupados ou acasalados de alguma forma. Isso ajuda tanto o indexador quanto o usuário a escolher os termos mais apropriados para representar um tópico particular, e identificar todos os termos necessários ao desenvolvimento de uma busca relativamente exaustiva sobre o assunto.

O vocabulário controlado de que o tesouro é uma variedade - é, com muita frequência, adotado para padronizar descritores ou cabeçalhos de assunto que representam o conteúdo das publicações ou interesses individuais (isto é, estratégias de busca e perfis de interesse adotados em Disseminação Seletiva de Informação - DSI). Em princípio, entretanto, tal recurso pode ser adotado em qualquer situação em que a padronização da terminologia seja requerida.

UNIDADE 2. PRINCIPAIS COMPONENTES DE UM VOCABULÁRIO CONTROLADO

Objetivos

1. Mostrar que um vocabulário controlado deve conter elementos classificados a par de elementos alfabéticos;
2. mostrar diferentes maneiras de como esses elementos podem ser apresentados;

3. ilustrar a estrutura de um tesouro convencional;
4. mostrar que um tesouro encerra uma classificação hierárquica completa.

Ilustração 5

Um vocabulário controlado deve conter dois elementos intercomplementares: um arranjo sistemático (classificado) de termos e um alfabético.

Ilustração 6

Aqui é mostrada uma classificação parcial de termos que descrevem tipos de bibliotecas. Tem a forma de uma árvore genealógica, como se usa em Genealogia, de tal forma que se torna fácil perceber que **bibliotecas departamentais** são uma subdivisão de **bibliotecas de faculdade**; estas, uma subdivisão de **bibliotecas universitárias**, e assim por diante. A subdivisão se baseia rigorosamente na relação gênero/espécie.

Bibliotecas departamentais são **tipos** de bibliotecas de faculdade, que são, por sua vez, **tipos** de bibliotecas universitárias.

Ilustração 7

O esquema de árvore genealógica funciona bem com hierarquias relativamente pequenas de termos, mas hierarquias muito complexas são de difícil esquematização por essa forma, porque pode resultar impossível apresentar a hierarquia toda em apenas uma página. Além do mais, a disposição em forma de árvore genealógica desperdiça espaço.

Neste exemplo é apresentada uma classificação hierárquica idêntica. Aqui, entretanto, os termos aparecem com um arranjo que se assemelha aos esquemas da classificação bibliográfica convencional adotados em bibliotecas (como a Classificação Decimal Universal ou a da Biblioteca do Congresso). Vários níveis de recolhido são empregados para indicar os níveis de hierarquia.

Ilustração 8

Essa se assemelha mais ainda a uma classificação bibliográfica, porque cada termo recebeu agora um número de classificação. Esses números são denominados notação da classificação. Alguns estudantes de Biblioteconomia são levados a crer que a notação é essencial a um esquema de classificação. Não o é. A ilustração 8 não é um esquema de classificação mais do que o é a Ilustração 7 ou mesmo a 6. As ilustrações 6 e 7, como apresentadas, são verdadeiros esquemas de classificação cobrindo uma parte da terminologia de "bibliotecas". A notação serve apenas para manter o esquema numa seqüência e para funcionar como um muito conveniente símbolo abreviativo de cada termo útil, por exemplo, na arrumação de livros em estantes de bibliotecas.

Ilustração 9

As ilustrações 6-8 apresentam termos de uma forma sistemática (classificada), mas não oferecem a abordagem alfabética. Devem oferecer, como suplemento, algum tipo de índice alfabético, para que fique evidente para o usuário onde, no esquema de classificação, um determinado termo aparece. Este é um exemplo simples de entradas de índice alfabético baseado na classificação da ilustração 8. Se a classificação é apresentada sem notação (ilustrações 6 e 7), o máximo que o índice pode fazer é remeter à página onde uma determinada hierarquia aparece — o que ilustra uma das desvantagens de não ter uma notação num esquema de classificação.

Ilustração 10

Nos exemplos 6-8, o arranjo principal dos termos é o classificado. E ele é explicitamente classificado. O índice alfabético é, na verdade, um suplemento, embora seja um suplemento indispensável ao uso eficiente de um esquema de classificação de qualquer porte. Um esquema de classificação se compõe de dois elementos constitutivos, distintos mas inter-relacionados: um classificado e outro alfabético.

A Ilustração 10 é uma apresentação alternativa do mesmo conjunto de termos biblioteconômicos. É explicitamente alfabética. Entretanto, a classificação hierárquica, que é explícita nas ilustrações 6-8, foi aqui embutida na ordenação alfabética de uma maneira implícita, através do emprego de referências cruzadas. Aqui um gênero é chamado de **termo genérico** (TG), e uma espécie de **termo específico** (TE). Assim, podemos notar que BIBLIOTECAS UNIVERSITÁRIAS possui duas espécies e um gênero, enquanto que BIBLIOTECAS possui quatro espécies e nenhum gênero (isto é, ele é o termo mais alto na árvore genealógica).

A Ilustração 10 é uma ilustração bastante simplificada de como os termos são apresentados em um tesouro de recuperação da informação. O tesouro difere do esquema de classificação descrito anteriormente porque:

- (1) o arranjo é explicitamente alfabético, mas implicitamente classificado;
- (2) os elementos alfabético e classificado são apresentados numa única seqüência;
- (3) em qualquer ponto de entrada só é oferecido um único nível hierárquico para cima e outro para baixo. Assim, é possível ver-se, a partir do termo BIBLIOTECAS UNIVERSITÁRIAS, que BIBLIOTECAS DE FACULDADE está um nível abaixo na hierarquia, mas deve-se ir ao termo BIBLIOTECAS DE FACULDADE para perceber que BIBLIOTECAS DEPARTAMENTAIS encontra-se um nível mais abaixo na mesma hierarquia. Essa estrutura de "um degrau acima, um degrau abaixo" não é essencial aos tesouros, mas economiza espaço quando o instrumento utiliza o papel como suporte físico;
- (4) a relação TG/TE apresenta reciprocidade simétrica. Se ao termo B corresponde como TE o termo Y, então ao termo Y deve corresponder como TG o termo B.

Embora não ostensiva, a classificação implícita na Ilustração 10 é tão completa e correta quanto a apresentada de forma mais explícita nas ilustrações 6-8. Na verdade, a Ilustração 10 procede diretamente da 7.

Além do mais, uma vez que as ilustrações 6, 7, 8 e 10 apresentam um esquema de classificação idêntico, um arranjo poderia ser derivado do

outro automaticamente. Quer dizer, poder-se-ia criar um programa de computador para derivar 10 e 6, 7 ou (mais facilmente devido à notação) de 8, enquanto outro programa poderia ser criado para derivar 6 ou 7 de 10. A Ilustração 8, também, poderia ser automaticamente derivada da 10. Neste caso, entretanto, o programa precisaria ser expandido, para permitir chegar-se à notação algorítmicamente.

Convém notar que o tesouro adotado em recuperação da informação é diferente de um tesouro literário, como o de Roget. Difere também das listas de cabeçalhos de assuntos adotadas tradicionalmente em bibliotecas. A lista de cabeçalhos de assuntos não encerra em si um esquema completo de classificação. Além do mais, cabeçalhos de assuntos são descritores mais "complexos" do que os descritores de um tesouro, no sentido de que eles tendem a ser mais "pré-coordenados" (ver o texto das Ilustrações 73-75).

Ilustração 11

Em todos os exemplos **biblioteconômicos** apresentados até agora, apenas um tipo de relação foi mostrado: o dos termos que aparecem numa mesma hierarquia. Mas termos de **hierarquias diferentes** também podem ser relacionados, como no caso dos pares ligados por setas na Ilustração 11(a).

Um vocabulário controlado unirá também termos que são intimamente relacionados, mesmo que apareçam em hierarquias diferentes. No tesouro consegue-se isso com o uso da remissiva **termo relacionado (TR)**, conforme visto em 11(b). As remissivas TG/TE associam termos da mesma hierarquia, enquanto a remissiva TR associa termos de hierarquias diferentes. Esse assunto é discutido com maiores detalhes mais adiante.

UNIDADE 3. COLETA DOS TERMOS

Objetivos

1. Distinguir as abordagens "de cima para baixo" e "de baixo para

- cima" na construção de tesouros;
2. distinguir "garantia literária" e "garantia de uso" e mostrar a importância desses dois princípios;
 3. mostrar como esses dois princípios podem ser adotados na coleta de termos para a construção do tesouro.

Ilustração 12

Antes de construir um tesouro inteiramente novo, um centro de informação fará muito bem em examinar um tesouro já existente, para verificar se por acaso não satisfaz às necessidades locais. Em alguns casos um tesouro preexistente pode ser modificado para se tornar de utilidade. Isso poderá significar expandi-lo em diversas áreas de assunto. Uma forma especial de expandir é construir um **microtesouro** adequado à estrutura de um instrumento mais geral (por exemplo: um tesouro sobre diabetes dentro da estrutura hierárquica de um tesouro médico preexistente).

Em outras ocasiões pode existir um tesouro adequado noutra idioma, podendo ser traduzido para o idioma do centro de informação.

Criar um tesouro inteiramente do nada pode ser um empreendimento muito dispendioso, que deveria ser levado a cabo somente quando se tem certeza de que nenhum dos instrumentos já existentes pode ser adotado ou adaptado.

Ilustração 13

Na construção do tesouro, podem ser identificadas quatro etapas fundamentais. A primeira delas consiste na coleta da matéria-prima — a terminologia do assunto a cobrir.

Ilustração 14

São apresentadas duas abordagens à construção de um tesouro. Na abordagem de cima para baixo, ou "dedutiva", que frequentemente exige o trabalho em equipe, as categorias principais de termos são identifi-

casas e subdivididas de cima para baixo. Dessa forma, um grupo de bibliotecários, construindo um tesouro sobre Ciência da Informação, reconhece a necessidade de uma categoria "tipos de bibliotecas". Adotando "bibliotecas" como o cabeçalho mais amplo, começam a subdividir a categoria, conforme mostrado na ilustração 6.

Problemas principais a destacar nessa abordagem:

1. É difícil prever todas as categorias ou hierarquias de que se vai necessitar;
2. a característica empregada para subdividir um gênero pode não conduzir à classificação mais adequada às necessidades dos usuários. Por exemplo: a classe "brinquedos" poderia ser subdividida por material, cor, idade da clientela a que se destina, método de locomoção etc. Para um fabricante de brinquedos, uma dessas características (talvez a idade) pode ser muito mais importante do que qualquer uma das outras.

A abordagem empírica (isto é, de baixo para cima, ou "indutiva") à construção de tesouros é exatamente o oposto, e tende a ser muito mais confiável. Os termos que ocorrem na área são coletados de fontes diferentes, e só é criada uma categoria de termos se ela parecer importante ou útil. Por exemplo: uma subdivisão de brinquedos por cor (azuis, vermelhos), embora teoricamente possível, pode não significar nada para o fabricante para quem o tesouro está sendo criado.

Na abordagem empírica, podem ser identificados dois princípios importantes: garantia literária e garantia de uso.

Ilustração 15

O princípio da garantia literária (também conhecida por garantia bibliográfica) foi enunciado por Wyndham Hulme há mais de 70 anos. O contexto foi o da classificação bibliográfica: um termo (número de classificação) se justifica apenas quando se sabe existir literatura sobre o assunto.

Extrapolado para a construção de tesouros, o princípio pode ficar assim: um termo se justifica apenas se ocorre dentro da literatura de um determinado assunto com algum grau de frequência.

A maneira mais eficiente de coletar a terminologia de um assunto é buscá-la naquelas fontes de referência com probabilidade de conter uma alta concentração de termos. As fontes têm que ser razoavelmente atualizadas.

A melhor de todas as fontes, talvez, seja uma publicação de resumos dedicada ao assunto, ou a seção adequada de uma publicação de âmbito mais amplo (por exemplo: a seção sobre Psicologia Educacional do *Psychological Abstracts*).

Trabalhando de forma absolutamente manual, o compilador do tesouro terá de examinar os resumos, registrando palavras e frases que pareçam relevantes para a descrição do assunto. Essas palavras e frases podem ser sublinhadas nos resumos ou transcritas para fichas. Eventualmente, podem ser transcritas para formas legíveis por máquina, para manipulação pelo computador.

É uma medida muito sensata caminhar para trás no tempo, começando pelo último fascículo dos resumos. O compilador pára a coleta quando atinge um ponto em que a ocorrência dos termos parece ter diminuído, isto é, quando começam a ser encontrados muito pouco termos novos.

A coleta de termos por essa forma manual é uma tarefa tediosa e que consome muito tempo. O processo pode ser acelerado consideravelmente quando o texto dos resumos pode ser obtido na forma de uma base de dados computadorizada. Neste caso, podem ser criados programas para imprimir em ordem descendente de frequência uma lista de palavras que ocorrem nos títulos e resumos. Pode-se fazer o mesmo com pares de palavras ou grupos de três, quando ocorrem com significativa frequência na base de dados. Além do mais, seria possível criar listagens impressas de co-ocorrência de palavras nos resumos, de grande valia na seleção dos TRs a serem usados no tesouro.

Ilustração 16

A seleção de termos a partir da literatura (garantia literária) é importante. Mas é igualmente importante que os termos coletados representem as necessidades dos usuários do serviço de informação que vai adotar o tesouro. Alguns termos podem ocorrer na literatura, mas serem de pouco interesse para um determinado grupo de usuários. Por exemplo: os usuários podem necessitar de menos especificidade do que o que a literatura sugere para certos tipos de termos.

O compilador do tesouro faria muito bem em identificar os termos que traduzem os interesses temáticos dos usuários, ao mesmo tempo em que os coleta da literatura. Estas abordagens são muito mais complementares do que alternativas.

Se o serviço de informação existe há algum tempo e colecionou registros de buscas de literatura realizadas para os usuários no passado, tais registros são, provavelmente, uma rica fonte de terminologia. Questionários também podem ser úteis (enviados a todos os usuários em potencial ou a uma amostragem deles, dependendo do número envolvido) na coleta de descrições que indiquem os interesses atuais da população a ser servida. No caso de uma comunidade de pesquisa, títulos/resumos de publicações desses usuários constituem uma valiosa fonte de termos.

Em certas situações é possível combinar numa única operação a garantia literária com a garantia de uso. Usuários representativos recebem cópias das publicações que mais de perto refletem seus atuais interesses profissionais. Os usuários, então, sublinham os termos que tornam essas publicações interessantes para eles.

UNIDADE 4. ORGANIZAÇÃO DOS TERMOS

Objetivos

1. Apresentar os padrões para a construção de tesouros;
2. mostrar como os termos podem ser organizados em categorias e em hierarquias dentro das categorias;

3. mostrar como um arranjo alfabético do tesouro pode ser sistematicamente derivado do arranjo hierárquico;
4. mostrar possibilidades alternativas de apresentação do componente alfabético de um tesouro.

Ilustração 17

O primeiro tesouro para recuperação de informação data de 1959. Muita experiência na construção de tesouros foi adquirida no início dos anos 60, especialmente nos Estados Unidos. Essa experiência foi sendo registrada paulatinamente, levando à publicação de diretrizes e, posteriormente, ao estabelecimento de normas para construção de tesouros (nos Estados Unidos, Reino Unido, França e Alemanha, por exemplo).

A UNESCO publicou diretrizes para a construção de tesouros monolíngües em 1970, e estas acabaram por levar ao estabelecimento da norma internacional ISO 2788, publicada em 1974. Um edição revista das diretrizes da UNESCO, publicada em 1981, tornou-se a base de um projeto de edição revista da ISO 2788.

Diretrizes da UNESCO para construção de tesouros multilíngües apareceram em 1976 e foram revistas em 1979, resultando na ISO DIS 5964 (1983). As formas de tesouro discutidas neste curso são compatíveis com as *Guidelines for the Establishment and Development of Monolingual Thesauri*, da UNESCO (2.ed. 1981), a menos que se diga algo em contrário.

Ilustração 18

Uma vez coletados os termos da literatura, dos usuários, ou ambos, eles precisam ser organizados em grandes categorias e em hierarquias dentro dessas categorias. Devem ser estabelecidas as relações inter-hierárquicas. Finalmente, o arranjo sistemático deve ser "invertido" para formar um arranjo alfabético.

Ilustração 19

Suponhamos que alguém estivesse construindo um tesouro sobre "irrigação" e que tivesse coletado em fichas uma grande quantidade de termos. As fichas podem agora ser arrumadas em pilhas, de tal forma que todas as fichas de uma determinada pilha pertençam ao mesmo "tipo" (isto é, estejam dentro da mesma categoria). Presumivelmente, uma categoria será reservada a tipos de irrigação, outra a tipos de plantações a irrigar, outra a tipos de solos, outra ainda a zonas climáticas, e assim por diante.

Ilustração 20

Uma vez identificadas as categorias, o passo seguinte consiste em organizar cada uma delas em hierarquias. Na maioria dos casos, a categoria inteira pode ser organizada com resultados práticos em uma grande e única hierarquia, como seria o caso de produtos agrícolas, que se dividem em grandes tipos, grandes subtipos etc. Em alguns poucos casos, entretanto, poderá fazer mais sentido não tentar adotar o esquema de hierarquia única. Por exemplo: uma das categorias identificadas por primeiro pode ser "propriedades". Todavia, posteriormente, o compilador decide que este termo geral é muito amplo e não muito útil como "termo superior" da hierarquia. Ele constrói, ao invés, várias hierarquias separadas: propriedades físicas, propriedades mecânicas, forma, cor etc.

Este exemplo mostra um desenvolvimento parcial da hierarquia "plantações". Da maneira como apresentada aqui, é uma verdadeira classificação hierárquica baseada na relação gênero/espécie: laranjas são um tipo de frutas cítricas; frutas cítricas são um tipo de fruta; e frutas são um tipo de produto agrícola.

Ilustração 21

O conjunto completo de hierarquias desenvolvido a partir dos termos sobre irrigação constituirá uma classificação da terminologia sobre irrigação. Uma vez satisfeito com as hierarquias estabelecidas, o

compilador inverte a classificação para criar o tesouro alfabético. Cada termo torna-se uma entrada e são mantidas suas relações hierárquicas, expressas em TGs e TEs.

Este exemplo mostra um grupo de entradas de tesouros, derivadas da Ilustração 20. Todas as relações gênero/espécie foram mantidas, e **são perfeitamente recíprocas**: visto que "limões" mostra "frutas cítricas" como seu TG, "frutas cítricas" deve mostrar "limões" como um de seus TEs. As ilustrações 20 e 21 são apenas duas maneiras diferentes de mostrar o mesmo esquema de classificação: 21 derivou-se de 20, e 20 poderia ter sido derivada de 21.

Ilustração 22

O esquema "de um degrau acima, um degrau abaixo" (descrito anteriormente) é usado na Ilustração 21. Ele é recomendado, quando não por outras razões, pela de economizar espaço no tesouro impresso.

Às vezes, adotam-se outros recursos para evidenciar a hierarquia, como na Ilustração 22. À esquerda são mostrados todos os níveis de termos específicos em uma única sequência alfabética. Não há nada que recomende isso: desperdiça espaço e não distingue os níveis hierárquicos. À direita é apresentado um arranjo alfabético-classificado, que tem a vantagem de mostrar todos os níveis hierárquicos num relancear de olhos. A única desvantagem é que, quando impresso, resultará em um instrumento muito volumoso. Para um tesouro relativamente pequeno e especializado, entretanto, pode ser a melhor forma de apresentação.

Ilustração 23

Quando adotado, o tipo alfabético-classificado deverá ser explícito em identificar os níveis hierárquicos: TE 1 é um termo específico em primeiro nível, TE 2 em segundo, e assim por diante.

UNIDADE 5. A RELAÇÃO HIERÁRQUICA

Objetivos

1. Distinguir as relações "genérica", "partitiva" e "exemplar";
2. mostrar como deve ser usada a relação TG/TE;
3. mostrar como a relação TG/TE pode ser indicada em um tesauro impresso.

Ilustração 24

A relação TG/TE acabará sendo do tipo gênero/espécie (isto é genérica) em quase todas as circunstâncias. Para ser um TE legítimo, um termo deve representar uma espécie da coisa identificada por seu TG.

Ilustração 25

A relação todo/parte (partitiva) via de regra não será considerada uma TG/TE. Por exemplo: em um tesauro sobre construção de móveis, o termo "pernas de mesa" deverá aparecer como TE subordinado a "pernas" e não a "mesas".

Por convenção, entretanto, certas relações partitivas são frequentemente tratadas como se fossem genéricas. Partes do corpo e nomes geográficos são os melhores exemplos. Faz sentido fazer de "aorta" um TE subordinado a "coração", embora ela seja uma parte e não um tipo de coração, o mesmo valendo para a localização geográfica. As Diretrizes da UNESCO reconhecem explicitamente duas outras situações em que a relação partitiva deveria ser tratada como genérica: campos de estudo (Física pode ser um TE de Ciência) e estruturas sociais (Divisões podem ser um TE de Exércitos).

Ilustração 26

As Diretrizes da UNESCO permitem a um tesauro reconhecer outros dois tipos de relações TG/TE, um genérico e outro partitivo, de tal forma que os termos "partes" podem ser agrupados com o termo "todo" ao qual se referem. Neste caso, faz-se uma distinção entre TEG (termo espe-

cífico genérico) e TEP (termo específico partitivo), e os dois grupos de termos são listados separadamente. A relação é recíproca: neste exemplo, "navios de guerra" terá como TEG "navios", e "cascos" terá como TEP "navios". Esse procedimento é útil no sentido de que evita a introdução de termos artificiais para funcionar como TG dos termos partes: por exemplo: "partes do navio" ou "componentes do navio".

Ilustração 27

As Diretrizes também permitem ser a relação "exemplar" tratada como TG/TE, quando isso parecer útil. Um termo "exemplar" será algum tipo de nome próprio - de pessoa, grupo, prédio etc. A "Catedral de Brasília" não é rigorosamente um tipo de catedral (um **tipo** seria, por exemplo, Moderna), mas um exemplo (exemplar).

Ilustração 28

Em um esquema de classificação facetada construído de forma correta, o princípio de divisão (ou "característica") empregado para derivar as facetas deve ser expresso clara e explicitamente.

Ilustração 29

Em um tesouro, entretanto, tais distinções não são, em geral, necessárias. Esta entrada é perfeitamente aceitável, embora os TEs representem três tipos de divisão diferentes (isto é, facetas): aplicação, país e estilo.

Ilustração 30

Não obstante, a identificação explícita de facetas em um tesouro é de se recomendar, particularmente, quando muitos termos estão envolvidos. O tipo de entrada aqui ilustrado evitaria uma longa lista de TEs heterogêneos sob o termo APLICAÇÕES EM ARQUITETURA (TE Arquitetura residencial etc.).

Ilustração 31

Dentro de um tesouro um termo terá, normalmente, apenas um TG (isto é, ele aparecerá em apenas uma hierarquia). Em alguns casos, entretanto, ele poderá receber legitimamente dois ou mais TGs. Quanto mais geral o tesouro, maior a probabilidade de que isso ocorra. Havendo a possibilidade, é melhor pôr um termo na hierarquia à qual ele sempre pertence: um diamante **pode ser** um instrumento para cortar, mas é **sempre** uma pedra preciosa.

Todavia, a relação hierárquica deve refletir os interesses da comunidade a quem serve. Por exemplo: em um tesouro para a indústria metalúrgica haverá pouca utilidade de uma hierarquia de pedras preciosas. Aqui, presumivelmente, "diamantes" apareceria apenas como "instrumentos para cortar".

Ilustração 32

Alguns tesouros poderão indicar, para cada termo, não apenas seu TG, mas também seu TS (isto é, o "termo superior" na hierarquia onde ele aparece). Nesta ilustração (ver Ilustração 20) o TG de "laranjas" é "frutas cítricas", enquanto seu TS é "produtos agrícolas".

UNIDADE 6. A RELAÇÃO ASSOCIATIVA

Objetivos

1. Distinguir a relação hierárquica (TG/TE) da associativa (TR);
2. dar exemplo de tipos de termos que talvez devessem ser associados por TRs.

Ilustração 33

Podem ser formuladas regras precisas para reger a relação TG/TE.

A relação TR, ou associativa, é muito menos susceptível de regulamentações.

A última coisa que se pode dizer com algum caráter de definido é que qualquer par de termos considerados inter-relacionados quanto ao sentido no contexto de um determinado tesouro pode ser associado através da remissiva TR, **contanto que eles não apareçam na mesma hierarquia**. Isso equivale a dizer que um TR é um conector inter-hierárquico, e não intra-hierárquico. É inteiramente desnecessário usar um TR para associar **arquitetura residencial** com **arquitetura religiosa** ou com **arquitetura naval**. Eles são irmãos, parceiros co-iguais na mesma hierarquia e a relação se torna evidente, indo para cima na hierarquia em direção ao termo genérico, **arquitetura**, no caso.

Por outro lado, as setas, neste exemplo, mostram termos que poderão ser legitimamente associados com **arquitetura religiosa**, desde que eles procedem de hierarquias inteiramente outras.

Ao contrário da relação TG/TE, não é imprescindível que a relação TR seja recíproca. Não se está obrigado a tornar **arquitetura religiosa** em um TR de **iconografia** apenas porque a recíproca é verdadeira. Entretanto, em princípio, é aconselhável adotar a reciprocidade.

Ilustração 34

Exemplificam-se aqui algumas relações frequentemente expressas por meio de um TR. A lista não se esgota, embora ela pareça incluir as relações mais importantes e frequentes. Dever-se-á aplicar a reciprocidade em todos os casos aqui apresentados (por exemplo: causa/efeito, a par de efeito/causa).

UNIDADE 7. CARACTERÍSTICAS DOS DESCRITORES

objetivos

1. Definir um descritor legítimo em termos de:

- (a) forma vocabular,
 - (b) singular/plural,
 - (c) seqüência vocabular;
2. orientar quanto à problemática dos "compostos", isto é, quanto à complexidade dos descritores.

Ilustração 35

Quase invariavelmente os descritores acabarão sendo substantivos ou algum tipo de locução. A maioria dessas locuções será composta com a participação de adjetivos, mas podem ocorrer também algumas locuções preposicionadas. Verbos na forma infinitiva e participial não devem ser adotados como descritores. Adjetivos que se encontram dissociados de seu substantivo devem ser evitados. Os descritores devem estar na forma direta, de preferência. Isso leva à vezes os alunos ao engano de querer agrupar os termos relacionados com LIVRO. É absolutamente desnecessário fazer isso - consegue-se o agrupamento por outras formas (por exemplo: indo hierarquia acima até o gênero "livro").

Ilustração 36

A maioria dos substantivos se refere à quantidade. São "quantitativos". Representam eles entidades mensuráveis através de números ("quantos?") e deveriam se apresentar sempre na forma do plural. Menos frequentes são os nomes que se referem a volume, e representam itens mensuráveis em termos de quantidade ("quanto?"). Devem ser representados na forma do singular.

Pode-se abrir exceção para substantivos normalmente considerados do singular, ou imaginados como sendo do singular em determinado contexto. Faz sentido falar-se de "olhos" e "ouvidos" e no entanto dizer-se "nariz". Em um tesouro sobre navios pode-se usar "ponte de comando" em vez de "pontes de comando", partindo do princípio de que por norma há apenas uma ponte de comando em cada navio.

Ilustração 37

Evite-se redundância desnecessária em um tesauro especializado. Em um tesauro sobre Biblioteconomia, o termo "cooperação" pode ser facilmente interpretado como "cooperação entre bibliotecas"; "prédios" como "prédios de bibliotecas", e assim por diante. Não há utilidade alguma em encontrar metade dos descritores do tesauro começando pela palavra "biblioteca".

Ilustração 38

Uma das maiores dificuldades que encontra um compilador de tesauro é descobrir até que ponto decompor uma frase em seus elementos constitutivos. Em um extremo encontra-se a decisão de manter a frase "ruído dos motores de automóveis" como um único termo (descriptor); no outro, a de desmembrá-la em diversos termos. Também é possível adotar um meio-termo. A decisão vai depender de vários fatores, incluindo-se entre eles o âmbito do tesauro e as convenções do idioma em questão (por exemplo: a praxe alemã *versus* a praxe inglesa).

Ilustração 39

As **Diretrizes** da UNESCO estabelecem algumas regras para ajudar na tomada de decisão, baseadas no princípio da "fatoração sintática". Elas são muito complexas e devem ser estudadas no contexto das próprias **Diretrizes**. Este exemplo tenta apresentar uma abordagem mais simples. Não está preocupado em apresentar todos os casos, mas apenas os considerados mais importantes.

Em geral, frases cujos elementos constitutivos aparecem também independentes de outros no tesauro devem ser desmembradas: "ruído do tráfego" é um bom exemplo, no caso de se precisar também dos dois termos "trafego" e "ruído". A idéia específica de "ruído do tráfego" pode ser expressa com a indexação sob **ambos** os termos da frase.

Frases que representam dois princípios de divisão diferentes (isto é, "características") devem ser sempre desmembradas. "Livros infantis ilustrados" encerra duas características diferentes: público destinatário e formato físico.

Termos que parecem representar uma espécie, mas na verdade expressam outro tipo de relação, devem ser desmembrados. "Corrosão de turbinas a gás" não é mera espécie (tipo) de corrosão (como o é, por exemplo, corrosão por fadiga), a turbina a gás é apenas o local onde ocorre a corrosão.

Compostos que, em princípio, não devem ser desmembrados:

- . os em que uma palavra têm apenas um contexto em um determinado tesouro - num tesouro sobre Biblioteconomia, "acidez" só deve existir associada a papel;
- . aqueles em que o significado do termo no composto é diferente do que ele tem quando isolado: lírio-do-mar não tem nada a ver com o vegetal lírio;
- . os cujos elementos constitutivos, tomados isoladamente, podem causar ambigüidade: "lâmina" e "vidro" podem representar tanto lâminas de vidro quanto vidro laminado;
- . os que se apresentam como membros de um gênero bem conhecido, mas querem dizer algo muito diferente: cavalos-marinhos não são, na verdade, uma raça de cavalos.
- . os que encerram um nome próprio;
- . os que ocorrem com tanta frequência numa determinada área do conhecimento, que o desmembramento do composto, embora possível, poderá provocar irritação nos usuários. É o caso de "recipientes de pressão" em engenharia mecânica.

UNIDADE 8. VOCABULÁRIO DE ENTRADA

BIBLIOTECA
DO
L. B. I. C. T.

Objetivos

1. Mostrar como se usa a remissiva USE;
2. mostrar como usar essa remissiva para controlar sinônimos e quase-sinônimos;
3. mostrar como essa remissiva pode ser empregada para remeter de um termo específico, não adotado na indexação, para outro mais genérico.

Ilustração 40

Um tesouro inclui remissivas de certos termos não adotados na indexação ou na busca (**não-descritores**) para os termos adotados em seu lugar (**descritores**). Os não-descritores encaminham para os descritores. Eles fornecem acesso adicional, ou pontos de entrada no vocabulário. Tomados em conjunto, podem ser denominados **vocabulários de entrada**. Há dois grandes tipos de **termos de entrada**, de acordo com a ilustração.

Ilustração 41

Sinônimos perfeito devem ser controlados por meio da remissiva USE. Um tipo de sinônimo é uma abreviatura ou uma sigla. Outro tipo é o que reflete convenções nacionais divergentes (por exemplo: a praxe brasileira *versus* a portuguesa). Deve-se, via de regra, adotar como descritor o termo com maior probabilidade de ocorrer à mente do público a que se destina. Para o público brasileiro o termo "trem" deve ser preferido ao termo "comboio"; o público comum provavelmente procurará sob o termo **coleção de selos**, e não sob **filatelia**.

Deve-se procurar ser consistente. Causa confusão o fato de abreviaturas e as siglas serem adotadas numa circunstância e não em outras. Contudo, uma sigla deve ser preferida sempre que a forma por extenso for pouco conhecida, como no caso de lasers.

BIBLIOTECA
DO
L. B. I. C. T.

Os quase-sinônimos ocorrem mais frequentemente do que os sinônimos verdadeiros. Um termo, A, pode ser considerado **sinônimo** de outro, B, quando a distinção entre ambos é imperceptível, ou não vale a pena fazê-la num determinado contexto. "Som" não é exatamente sinônimo de "acústica", mas em alguns tesouros esses dois termos podem ser tomados como sinônimos.

Não há necessidade de fazer remissiva das formas invertidas para as diretas, porque o ponto de acesso adicional será fornecido de outras maneiras. Por exemplo: a entrada Livros, encadernação USE Encadernação de livros, é redundante, uma vez que o termo "Encadernação" listará os descritores para todos os tipos de encadernação como seus TEs. Uma possível exceção é o caso de frases preposicionais (por exemplo: **Equações de movimento**).

Termos que representam os extremos opostos de um contínuo de valores são normalmente tratados como se fossem sinônimos. Por exemplo: um artigo sobre a elasticidade de algum material pode ser considerado como se tratasse também de sua inelasticidade. Esses pares são às vezes denominados "quase-sinônimos".

O mapeamento pode ser "um para muitos" ou "um a um". Na Ilustração 7, a expressão "arquitetura residencial", deve ser indexada e procurada sob ambos os termos, "arquitetura" e "residências".

Termos considerados demasiado específicos para adoção como termos de indexação podem ser mapeados com o termo mais geral e mais adequado. Por exemplo: há garantia literária para "ruído branco", mas determinado tesouro pode não exigir tamanho grau de especificidade.

Ilustração 42

A remissiva USE deve ter sua recíproca. Se existir a remissiva A USE B, a recíproca deve existir também: B UP A, onde UP significa "usado para". No caso do mapeamento "um para muitos", algum símbolo é

necessário para indicar que um termo é apenas um de dois ou mais termos referidos. Nesse exemplo, em "residências", o sinal + depois de UP mostra que o termo "residências" é um de dois ou mais termos referidos em "arquitetura residencial".

UNIDADE 9. NOTAS EXPLICATIVAS E IDENTIFICADORES

Objetivos

1. Ilustrar os tipos de notas explicativas e mostrar sua finalidade;
2. mostrar como se distinguem os homógrafos num tesauro;
3. distinguir um **descriptor** de um **identificador**.

Ilustração 43

Quando o significado de um termo pode parecer confuso, ele pode receber uma nota explicativa (NE) no tesauro. Pode ser ela uma verdadeira definição de dicionário, como no exemplo de "permafrost". Como alternativa, uma NE pode simplesmente indicar como um termo deve ser interpretado no contexto de determinado tesauro e/ou distingui-lo de outros termos com os quais pode vir a ser confundido.

Ilustração 44

Algumas palavras podem ter mais de um significado, mesmo no contexto de um mesmo tesauro; isto quer dizer que elas são **homógrafas**. Frequentemente é possível distinguir significados diferentes com o emprego de um qualificador entre parênteses, sem necessidade de toda uma nota explicativa. Ao contrário da nota explicativa, o qualificador entre parênteses é parte integrante do descriptor. Por exemplo: o termo "tanques" não existe por si só. Ele foi desmembrado em dois termos distintos: TANQUES (RECIPIENTES) e TANQUES (ARMAS). Sem dúvida alguma, existem outros métodos para dirimir tais ambigüidades, que podem ser adotados em alguns tesauros.

Ilustração 45

Os descritores organizados em um tesouro tendem a ser substantivos ou locuções substantivas, mas podem não incluir muitos nomes próprios. Todavia, em muitas disciplinas, nomes de pessoas, lugares, organizações, ou objetos serão necessários para representar o assunto dos documentos. Em vez de expandir enormemente a estrutura hierárquica do tesouro, esses nomes devem ser "controlados" (isto é, normalizados) em um vocabulário à parte. Os nomes tratados dessa forma são frequentemente denominados "identificadores", e a lista empregada para controlá-los será simplesmente isso - uma lista, talvez com algumas remissivas, mas não com uma verdadeira "estrutura".

São apresentados aqui exemplos de tipos de termos tratados como identificadores em diversos serviços de informação.

Não quer dizer que todos os nomes próprios devam ser tratados como identificadores em todos os tesouros. Sem dúvida alguma um tesouro sobre Ciência Política necessitará de nomes de países como elemento integrante da estrutura do tesouro. Por outro lado, um tesouro sobre Educação tem menos necessidade de descritores geográficos: nomes de países podem ter seu tratamento através da lista de identificadores. De fato, em uma disciplina em que esses nomes possam incidentalmente vir a ser necessários, a melhor coisa a fazer, talvez, é adotar alguma lista oficial (autorizada) já existente. No Brasil, as fontes óbvias a pesquisar são as publicações do Instituto Brasileiro de Geografia e Estatística (IBGE).

Em um tesouro, por exemplo, sobre Literatura Brasileira, os nomes dos autores serão obviamente necessários. Mas não há sentido prático em tentar incluir todos esses nomes na estrutura de um tesouro. Basta incluir os nomes dos autores mais significativos como descritores, tratando os demais como identificadores. Nesse caso, um ensaísta relativamente insignificante do século XIX pode ser indexado sob descritores relevantes - ENSAÍSTAS e SÉCULO XIX - bem como sob seu próprio nome, tirado da lista de identificadores.

UNIDADE 10. FORMATO E APRESENTAÇÃO DO TESAURO

Objetivos

1. Ilustrar os componentes principais da entrada de um tesauro;
2. ilustrar modos alternativos de apresentação de um tesauro;
3. mostrar, com exemplos, os esquemas gráficos e o tesauro. facetado.

Ilustração 46

Se a entrada sob um termo contiver todos os elementos habituais, ela terá essa aparência. Os elementos são apresentados na sequência recomendada pelas normas. Em cada lista de termos com remissivas TE, TR etc., os termos são listados alfabeticamente. A maioria das entradas em um tesauro não contém todos os elementos; as notas explicativas, particularmente, aparecem como exceção, não como regra.

Ilustração 47

Conforme descrito anteriormente (Unidade 2), a organização alfabética dos termos num tesauro deve conter um verdadeiro esquema de classificação hierárquica. Entretanto, certos tesouros incluem alguma forma de esquema de classificação ostensiva como um recurso secundário, um complemento ao arranjo alfabético. São também possíveis outras formas de arranjo.

O tesauro UNBIS, usado pela biblioteca Dag Hammarskjold, das Nações Unidas, em Nova Iorque, incorpora quatro arranjos complementares de termos. O arranjo alfabético, que esta ilustração apresenta, é convencional, exceto no que diz respeito à inclusão de TSs (cuja explicação se encontra na Ilustração 32) e de números para as categorias (por exemplo: 04.02.01 em IRRIGATION é a categoria na qual esse termo aparece).

Ilustração 48

Os números das categorias fornecem pontos de entrada para uma lista de termos por categoria. Isso pode ser considerado um esquema de classificação muito geral. Dentro de cada uma das categorias principais (por exemplo, CROPS) entretanto, os termos são listados alfabeticamente, em vez de hierarquicamente. Da forma como mostrado aqui, o arranjo inclui remissivas tipo USE e notas explicativas (NE).

Ilustração 49

O terceiro arranjo é uma "lista hierárquica". Aqui cada "termo superior" é arranjado alfabeticamente, e abaixo dele aparece a hierarquia completa, empregando-se recolhido para indicar os diferentes níveis.

Muitos tesouros poderão incluir alguma forma de arranjo sistemático para complementar o alfabético - algo semelhante às ilustrações 48 ou 49. A inclusão de dois arranjos dessa natureza é menos comum.

Ilustração 50

A maioria dos tesouros, como o tesouro UNBIS, inclui algum tipo de arranjo permutado das palavras, seja KWIC (palavra-chave no contexto), seja KWOC (palavra-chave fora do contexto). Esse tipo de índice mostra, para cada palavra, todas as suas ocorrências nos descritores. É um suplemento útil, enquanto revela ocorrências de palavras (e por isso ocorrências do tema) que poderiam não ser facilmente encontráveis no arranjo alfabético ou em outros tipos de arranjo.

Ilustração 51

Outro método alternativo de apresentar um tesouro é com o emprego de "esquemas gráficos". O tipo usado como ilustração, retirado das **Diretrizes** da UNESCO - é conhecido como um "gráfico de setas". Ele é, no fundo, um arranjo hierárquico apresentado por meio de gráficos. O termo superior da hierarquia, **câmaras**, é o centro do gráfico; os demais

giram em torno dele para mostrar os níveis de hierarquia. Dessa forma, **Câmaras de 35 mm** é subordinado a **Minicâmaras**, por sua vez subordinado a **Câmaras Fotográficas**, que é uma subdivisão de **Câmaras**. Linhas pontilhadas unem os termos a outros elementos que lhes são associados. Quando há uma certa quantidade desses gráficos de setas, eles se assemelham a um atlas rodoviário, onde um mapa remete para outro. Os defensores de arranjos que utilizam esquemas gráficos insistem em que, uma vez que alguém tenha se familiarizado com esse formato, será muito mais fácil examinar uma família inteira de termos do que em qualquer outro formato.

Ilustração 52

O gráfico de setas, basicamente hierárquico, constitui o elemento classificado do tesouro. É evidente que se faz necessária alguma abordagem alfabética. Poderá ser simplesmente um índice alfabético. Neste exemplo, entretanto, o elemento alfabético encontra-se, de fato, na forma de um tesouro convencional, com um acréscimo: a posição de cada termo no gráfico de setas — o número do mapa e a referência ao quadrante — é indicada.

Ilustração 53

O tesouro SPINES, da UNESCO inclui um tipo completamente diferente de arranjo que também pode ser considerado do tipo "gráfico". Este tesouro incorpora uma certa quantidade de características próprias. O componente alfabético identifica explicitamente todos os níveis de TG e TE para cada descritor, ao mesmo tempo em que provê notas especiais, remissivas TRs e USE. É feita uma distinção entre as remissivas USE e VER, em que esta última oferece uma oportunidade de escolher descritores em substituição aos não-descritores (termos de entrada). A inclusão desses dois tipos de referência (USE e VER) não se tornou convenção, e não pe recomendada pelas Diretrizes da UNESCO. É aconselhável que os alunos desconsiderem essa distinção.

Ilustração 54

A forma de apresentação com esquemas gráficos usa polígonos, subpolígonos e subsubpolígonos na tentativa de tornar evidentes diferentes níveis de hierarquia. Polígonos relacionados são ligados por meio de uma "relação associativa", um tipo de remissiva TR. Uma página completa com esse tipo de apresentação que utiliza esquemas gráficos ("mapa") pode incluir muitos polígonos inter-relacionados, embora o exemplo escolhido mostre apenas um polígono completo.

Ilustração 55

Como já se enfatizou anteriormente, todos os tesouros contêm um componente classificado - ao qual se chega através da adoção da estrutura TG/TE, um dispositivo gráfico que indica hierarquia ou outro recurso qualquer. O **tesauro facetado** é um tipo de vocabulário controlado que combina o tesauro alfabético com uma classificação hierárquica facetada, inclusive com notação.

São mostradas aqui algumas entradas de um **tesauro facetado** hipotético. Convém notar que essas duas partes se completam. O componente facetado se encarrega da relação hierárquica (isto é, a estrutura TG/TE), enquanto todas as outras relações aparecem no tesauro propriamente dito. O tesauro fornece a notação para cada termo, de tal forma que o usuário pode procurar o componente facetado para ver a relação hierárquica. Na classificação facetada um termo aparece apenas uma vez. Entretanto, se um termo puder legitimamente aparecer em mais de uma hierarquia, as relações secundárias serão fornecidas pelo tesauro, usando-se a convenção TG (A), que significa "termo genérico adicional". O tesauro facetado oferece vantagens óbvias sobre outras formas de tesauro: pode ser usado para arrumar livros nas estantes de uma biblioteca especializada, bem como para indexação dos itens numa base de dados. Além disso, o arranjo nas estantes e o da base de dados serão assim inteiramente compatíveis.

Ilustração 56

O UNESCO Tesauro pode ser considerado um tipo de tesauro facetado. Aí temos o elemento facetado onde cada faceta é **explicitamente** identificada ("by crimes against property", e ("by crimes against public order").

Ilustração 57

Neste tesauro porém, os elementos alfabético e facetado não são inteiramente complementares, conforme mostrado no exemplo hipotético (Ilustração 55). Ao mesmo tempo em que o elemento alfabético fornece muita informação que não se encontra no elemento facetado, ele repete algumas das informações ali contidas (por exemplo: BT (TG), TT (TS), UF (UP)).

Ilustração 58

O ROOT Tesauro, da British Standards Institution, é, talvez, o mais sofisticado dos instrumentos que combinam uma classificação facetada com um arranjo alfabético dos termos. As facetas pelas quais um assunto é dividido estão clara e explicitamente indicadas, por propriedade ("by property"), por aplicação ("by application") etc. O tesauro adota algumas convenções especiais, como:

- = indica um sinônimo não-preferido
- * < indica um termo genérico adicional retirado de outra parte do tesauro (não há exemplo aqui)
- * > indica um termo específico adicional retirado de outra parte do tesauro (por exemplo: "drainpipes" aparece em RFM.N, ao mesmo tempo em que é considerado um tipo de cano em NJV/NJX; "rigid pipes", por outro lado, em NJV.B, não aparece em mais lugar algum na classificação).
- *- indica um termo relacionado adicional retirado de outra parte do tesauro.

Ilustração 59

A seção alfabética fornece apenas um nível de hierarquia. São adotadas as seguintes convenções:

- = sinônimo não-preferido
 - < termo genérico na mesma parte do tesouro (por exemplo "pipes" é mostrado como um termo específico de "fluid equipment").
 - > termo específico na mesma parte do tesouro (por exemplo: "rigid pipes" como específico de "pipes").
- termo relacionado na mesma parte do tesouro.
- * < termo genérico em outra parte do tesouro.
- * > termo específico em outra parte do tesouro ("drainpipes", foi o exemplo usado anteriormente).
- *- termo relacionado em outra parte do tesouro.

Não-descritores são tratado de duas maneiras, como nesses exemplos:

Pipe threads

→ Threads NWC/NWD

indica que o tópico anterior deve ser indexado sob o outro termo, enquanto que

* Asbestos cement pipes NJW.S

→ Pipes

+ Asbestos cement

indica um "termo sintetizado": o termo "asbestos cement pipes" não deve ser usado; esse conceito deve ser indexado em "pipes" e em "asbestos cement" (observe-se que a entrada em "pipes" mostra que esse termo + "asbestos cement" deve ser usado para representar "asbestos cement pipes"). O símbolos especiais usados em lugar de BT (TG), NT (TE), RT (TR) e USE tornam a estrutura deste tesouro independente da língua. Todavia o emprego de tais símbolos não é preconizado pelas normas internacionais.

UNIDADE 11. CRESCIMENTO E ATUALIZAÇÃO

Objetivos

1. Mostrar como se pode aplicar um teste simples para verificar o grau de "completeza" da versão preliminar de um tesouro;
2. definir as formas como um tesouro pode ser mantido atualizado.

Ilustração 60

Um tesouro não é algo estático; ele continuará a crescer à medida que novos termos forem necessários para acomodar assuntos não encontrados antes. Geralmente isto significa desenvolver as hierarquias atuais com maior detalhe, tornando assim o vocabulário cada vez mais específico.

Apresenta-se uma curva hipotética de crescimento. Partindo do pressuposto de que o tesouro é construído a partir da coleta de termos na literatura, o crescimento será muito rápido no início - virtualmente todo item novo (resumo, por exemplo) vai gerar muitos termos novos. Posteriormente, entretanto, ele começará a atingir um ponto de equilíbrio, e haverá aquele momento em que muitos novos documentos deverão ser examinados para que se encontre algum termo novo. Neste ponto a coleta de termos pode ser suspensa: os já coletados devem ser organizados para produzir o primeiro esboço de um tesouro.

Ilustração 61

Seria conveniente que a versão preliminar fosse revista, de preferência por especialistas no assunto, bem como pelos especialistas em informação que a utilizarão. Ante de começar a ser utilizado na prática, é aconselhável testar sua adequação aos objetivos.

A ilustração mostra um teste simples. Tome-se uma amostra aleatória de 100 artigos ou relatórios publicados recentemente, do tipo a ser indexado na base de dados. Um ou vários especialistas em informação

indexam esse material em duas etapas. Primeiro, o indexador anota os resultados de sua "análise conceitual" com uma série de palavras ou frases que representem os tópicos que ele julga devam ser cobertos na indexação. Essa análise conceitual é feita para todos os 100 documentos da amostra. Em seguida, consulta a versão preliminar do tesouro e faz uma tentativa de traduzir a análise conceitual para o vocabulário controlado.

Ilustração 62

Mostra-se aqui um exemplo hipotético. Neste caso, todos os conceitos listados pelo indexador podem ser traduzidos para os termos do tesouro, embora não seja sempre possível uma equivalência unívoca. Por exemplo: dois "conceitos" são traduzidos por um termo ou um conceito por dois termos. Em apenas um caso há variação na especificidade: o tesouro oferece "água residual", mas na verdade "água de esgoto" é um tipo especial daquela.

Ilustração 63

Os resultados do teste em todos os documentos da amostra podem ser expressos estatisticamente. Neste caso hipotético os resultados parecem bons. Muito poucos conceitos não são traduzíveis. Quase todos, na verdade, podem ser traduzidos no nível de especificidade desejado. Num teste dessa natureza podem-se identificar algumas deficiências específicas que, no entanto, podem ser corrigidas - talvez sejam necessários termos mais específicos para tipos de água utilizável em irrigação. Uma vez feitas as correções, o tesouro parece pronto para implantação.

Ilustração 64

Uma vez implantado, o tesouro continuará naturalmente a crescer. Ele deve ser mantido atualizado com o acréscimo de novos termos. O serviço de informação a quem ele serve de instrumento deve estabelecer procedimentos, baseados nos quais os indexadores possam acrescentar novos descritores na medida das necessidades, ao menos de uma forma

provisória. No caso de um serviço de informação de grande porte, algum "grupo de revisão do vocabulário" pode ser formado para decidir se esses descritores provisórios devem ser incorporados ao tesouro, e, em caso afirmativo, se devem aparecer na estrutura.

Os termos que ocorrem nas consultas formuladas pelos usuários devem sofrer o mesmo processo de acompanhamento. Isso se faz necessário especialmente para garantir que o vocabulário permanecerá suficientemente específico satisfazer às necessidades de informação dos usuários.

UNIDADE 12. UTILIZAÇÃO DO COMPUTADOR

Objetivos

1. Mostrar como os processos computadorizados podem ser empregados na construção e manutenção de tesouros;
2. mostrar o papel importante que os tesouros computadorizados podem desempenhar nas atividades relacionadas com a recuperação da informação.

Ilustração 65

Os computadores podem desempenhar um papel muito importante na construção e na utilização de tesouros. A possibilidade de adoção dos computadores na coleta dos termos foi mencionada anteriormente (Ilustração 15). Os termos podem ser extraídos de bases de dados em computador, produzindo-se listas ordenadas de acordo com a frequência de sua ocorrência. Mesmo que os termos não sejam extraídos de bases de dados em computador, é mais do que aconselhável lançá-los em suportes informáticos, para permitir sua ordenação e contagem, e, possivelmente, para produzir tabelas de co-ocorrência.

Devido ao fato de a construção de um tesouro convencional ser essencialmente uma atividade intelectual, o processamento de termos

por meio do computador pode para tal contribuir apenas em escala relativamente reduzida. Por exemplo: ordenando e imprimindo todos os termos pela palavra de comando (soldagem, soldagem a gás, soldagem a arco blindado) ou mesmo pelos radicais dessas palavras, o computador pode ajudar na identificação de facetas ou hierarquias.

Tendo o compilador do tesouro organizado o vocabulário em facetas e hierarquias, o processamento por computador pode ser de extrema valia, para não dizer absolutamente essencial. Como exemplo, considere-se a hierarquia parcial da Ilustração 8, onde uma notação simples é utilizada para representar os níveis hierárquicos. Com esses dados o computador pode imprimir um arranjo hierárquico igual ao mostrado na Ilustração 7, um arranjo de palavras permutadas, como o da Ilustração 50, e (contanto que cada hierarquia tenha sido atribuída a uma determinada categoria de assunto) um arranjo por categoria, como o da Ilustração 48.

O que é mais importante, contudo, é que pode ser usado um programa de computador para gerar um arranjo alfabético perfeitamente consistente (como na Ilustração 10), com todas as referências TG, TE, TR e USE em perfeita relação de reciprocidade. Mais ainda, os dados do tesouro em forma legível por computador podem dar entrada numa máquina de fotocomposição, de modo que a impressão de todos os arranjos pode ter a mais alta qualidade tipográfica.

Além de produzir tesouros e apresentá-los numa grande variedade de formatos de saída, os programas de computador podem ser utilizados na manutenção do tesouro. Termos novos podem ser acrescentados a qualquer momento. O acréscimo de um termo na forma apropriada fará com que o mesmo seja incorporado à estrutura do tesouro na posição correta, ao mesmo tempo em que são estabelecidas as remissivas de reciprocidade. Quando um termo for eliminado do tesouro, os programas checarão todos os termos em relação de reciprocidade com ele e eliminarão automaticamente todas as remissivas ao mesmo. Se a grafia de um termo foi alterada, a alteração será processada pelos programas em todos os lugares onde o mesmo ocorrer. Se forem introduzidas referências cruzadas adicionais a um termo, os programas gerarão automaticamente

as relações recíprocas dessas remissivas. Semelhantemente, o computador procederá às alterações necessárias no tesouro toda vez que uma nota explicativa for alterada, ou quando um termo for deslocado de uma categoria de assunto para outra.

Ilustração 66

O computador pode desempenhar papel relevante na geração de tesouros impressos. O tesouro computadorizado pode também ter importantes funções a desempenhar na operacionalização de um sistema de recuperação de informação.

O tesouro computadorizado é um registro autorizado do vocabulário do sistema, inteiramente atualizado. Podemos chamá-lo de "arquivo-mestre do vocabulário" ou AMV. Além de fornecer listagens do tesouro de diversos tipos, o AMV tem as seguintes funções em um sistema informatizado de recuperação da informação: ele checa a consistência e aceitabilidade dos termos empregados por indexadores e usuários do sistema. No caso de um indexador ou um usuário empregar um termo não reconhecido pelo AMV, o registro de entrada (ou entrada de busca) será rejeitado, com uma notificação apropriada ao indexador ou usuário.

Em alguns sistemas o AMV pode realizar certas atividades de mapeamento automático. Por exemplo, o indexador ou o usuário podem ser autorizados a usar um qualquer de dois sinônimos reconhecidos pelo tesouro, com o AMV procedendo automaticamente à substituição do termo preferido. Somente quando o termo empregado pelo indexador ou pelo usuário for inteiramente estranho ao sistema, sofrerá ele rejeição.

O AMV mantém certas informações de natureza estatística úteis que seriam de difícil ou impossível manutenção por processos manuais. Em primeiríssimo lugar, ele mantém um registro do número de referências bibliográficas relacionadas com cada termo, isto é, o número de vezes que ele foi empregado na indexação. Esses dados podem ser apresentados em forma impressa ou em terminal linha. A informação sobre a frequência com que um termo foi empregado na indexação é de utilidade

nas operações de busca, porque ela permite ao pesquisador calcular quantas referências serão recuperadas em resposta a uma determinada estratégia de busca, ou, na pior das hipóteses, mostrar o número máximo de referências que poderiam, em princípio, ser recuperadas.

Informações de cunho estatístico sobre a frequência de uso de termos na indexação são relevantes para as atividades de controle de vocabulário. Um descritor usado com muito pouca frequência durante os últimos 12 meses pode ser um forte candidato a ser eliminado do vocabulário.

Dados estatísticos sobre frequências de uso de termos na busca também são de importância capital para as atividades de controle de vocabulário (talvez até mais importante do que as informações estatísticas sobre o uso na indexação); essas informações são usadas com menos frequência nos sistemas automatizados, embora não ofereçam particular dificuldade de coleta.

Um registro da estrutura hierárquica completa do vocabulário deve aparecer no AMV. Isso torna mais fácil a execução de buscas genéricas. Apenas o termo mais genérico precisa ser especificado na busca; os seus descendentes sendo substituídos automaticamente no AMV. Observando mais uma vez a Ilustração 49, deve existir algum meio de incluir numa busca todos os termos de uma determinada hierarquia — por exemplo, todos os termos relacionados com "veículos" ou todos os termos relacionados com "veículos motorizados".

Empregando informações de natureza estatística sobre a frequência de uso de termos na indexação, o AMV pode ser adotado na otimização automática da estratégia de busca. Tomemos como exemplo a estratégia booleana (A ou B) e (L ou M ou N) e Y. Para satisfazer a essa fórmula, uma referência deve ter sido indexada por qualquer dos termos A ou B e (L ou M ou N) e, além disso, pelo termo Y. Suponhamos que Y é um termo muito genérico e foi usado 10 000 vezes: que A e B foram, juntos, empregados 750 vezes, e que L, M e N juntos, foram empregados apenas 84 vezes. Ao fazer a busca no sistema será mais proveitoso ler

antes a lista menor de números de documentos, comparando-a, em seguida, com a segunda mais curta, deixando por último no processo a comparação com a lista mais extensa. A partir das informações estatísticas sobre as referências no AMV, um programa eficiente de busca pode otimizar automaticamente a estratégia, isto é, identificar a lista com menor número de referências, em seguida a que ocupa o segundo lugar, e assim por diante, organizando a estratégia de busca em ordem ascendente de número de referências e manipulando os arquivos nessa seqüência, para economizar o tempo do computador.

O AMV pode ser empregado para gerar automaticamente remissivas VER e VER TAMBÉM nos índices impressos. Por exemplo, quando uma entrada tem que aparecer no índice impresso sob o termo EVAPORAÇÃO, a remissiva vaporização VER evaporação é gerada automaticamente, e assim aparecerá no índice impresso. Da mesma forma, se os termos DEMODULADORES e MODULADORES forem, ambos, adotados como cabeçalhos na edição de um índice impresso, devido ao fato de estarem ligados pelos indicadores TR no tesauro computadorizado, as remissivas impressas VER TAMBÉM serão geradas automaticamente.

Ilustração 67

Num sistema de informação em linha tanto o indexador quanto o usuário precisam consultar uma lista alfabética do vocabulário. Quer dizer, ao menos em algumas ocasiões eles precisarão se certificar de que determinado termo existe na base de dados. A capacidade de apresentar um vocabulário em ordem alfabética é relativamente freqüente nos atuais sistemas em linha. Em resposta a um comando NEIGHBOUR ou EXPAND, para qualquer termo utilizado, o sistema mostrará esse termo juntamente com os que o precedem na seqüência alfabética, ou com os que se lhe seguem. À direita de cada termo aparece um número que indica a quantidade de referências bibliográficas sob aquele termo. Convém notar que o termo pedido pelo usuário não apresenta referências relacionadas com ele. Significa que esse termo específico não se encontra na base de dados. Não obstante isso, o sistema apresenta aqueles termos que se encontram mais próximos dele na ordem alfabética. Este recurso pode

ajudar o pesquisador a reconhecer o termo que ele precisa de fato usar (nesse caso, talvez, LINGUAGEM DE INDEXAÇÃO) e é, também, um recurso que ajudará a contornar os problemas de erro de grafia.

É evidente que uma capacidade de arranjo de vocabulário dessa natureza é uma característica essencial de um sistema em linha, quer adote a base de dados a ser pesquisada um vocabulário controlado, quer um não controlado. Esse recurso pode ser usado não apenas para verificar a existência de um termo, mas ele pode conduzir o consulente à descoberta de um termo alternativo, mais adequado do que o inicialmente escolhido, devido ao fato de que a proximidade alfabética apresenta juntos os termos relacionados. Neste exemplo o consulente pode decidir que LINGÜÍSTICA APLICADA é o termo mais apropriado à sua necessidade de informação. O sistema em linha pode oferecer ao usuário a possibilidade de buscar termos subindo e descendo na ordem alfabética. Ele poderia, na verdade, dessa forma, verificar o vocabulário inteiro do sistema, embora essa atividade resultasse em emprego demasiado ineficiente e dispendioso do tempo de conexão em linha.

É importante que, quando um vocabulário for mostrado em linha, o usuário seja capaz de escolher à simples leitura na tela, sem necessidade de entrar os termos no teclado. Pode-se conseguir a seleção de termos da lista apresentada usando os números identificadores das linhas na tela. Esse recurso economiza o tempo do usuário e reduz a probabilidade de erro.

Ilustração 68

Uma simples listagem alfabética de termos, embora necessária por várias razões, presta muito pouca assistência na formulação de uma estratégia de busca abrangente. O usuário necessita também de um recurso capaz de lhe permitir visualizar os termos que guardem uma maior relação semântica entre si. Para qualquer termo que se coloque no terminal, o sistema deveria ser capaz de apresentar todos os outros termos que, segundo o vocabulário, tenham com ele relações paradigmáticas e sintagmáticas. Há uma variedade de alternativas para

esse tipo de apresentação. Poderia ser a apresentação convencional de um tesouro, indicando para o termo inicial o termo ou termos que são hierarquicamente mais genéricos (TG), o termo ou termos que são mais específicos (TE) e o termo ou termos com ele relacionados de forma não-hierárquica (isto é, os termos relacionados ou TRs). Mais uma vez, o usuário deveria ser capaz de escolher qualquer termo dessa apresentação, sem ter que entrá-lo por extenso.

Melhor, talvez, do que a apresentação em forma de tesouro, é a em forma de classificação hierárquica, como as exemplificadas nas ilustrações 7 ou 8.

Um sistema em linha deveria ser capaz de não apenas apresentar as hierarquias de termos, mas também de permitir ao usuário incorporar uma hierarquia inteira de termos na estratégia de busca por meio de algum comando simples. Na Ilustração 8, por exemplo, isso poderia ser feito apenas com o pedido de todos os termos "B". No caso de um arranjo do tipo mostrado na Ilustração 68, toda a hierarquia de termos pode ser incorporada através do uso de todos os números apropriados (neste exemplo, talvez 1, 3 -7).

UNIDADE 13. FATORES DO VOCABULÁRIO QUE AFETAM O DESEMPENHO DE SISTEMAS DE INFORMAÇÃO

Objetivos

1. Mostrar os objetivos de um eficiente sistema de recuperação de informação;
2. definir **revocação** e **precisão**;
3. mostrar a importância da especificidade do vocabulário;
4. mostrar a importância da estrutura hierárquica adequada;
5. definir **coordenação falsa** e **relação incorreta entre os termos**;
6. distinguir entre sistemas pré-coordenados e pós-coordenados.

Ilustração 69

Em operações de recuperação da informação pode ser desejável extrair da base de dados aqueles registros que podem ser considerados **relevantes** a alguma necessidade de informação. Ao mesmo tempo, é desejável evitar a recuperação de quaisquer registros que não sejam relevantes.

O retângulo representa uma base de dados hipotética. Cada "+" representa um registro que é relevante para uma determinada necessidade de informação, e cada "-" um documento não-relevante. Executando uma busca, então, é desejável recuperar o maior número possível de itens "+" e o menor número possível de itens "-".

É apresentado o resultado de duas buscas. Em A, três itens relevantes foram recuperados, e nenhum irrelevante. Essa pode ser considerada uma busca de **alta precisão**. Ao mesmo tempo, é uma busca de **baixa revocação**, pois a maioria dos itens relevantes ficou de fora.

A busca B, por outro lado, recupera a maioria dos itens relevantes. Ao mesmo tempo, entretanto, recupera muitos que não são relevantes.

- **Revocação** refere-se à capacidade de recuperar itens relevantes; **precisão**, à capacidade de recuperar **apenas** itens relevantes (ou, em outras palavras, a capacidade de não recuperar itens irrelevantes).

Em termos de revocação e precisão, os resultados das duas buscas podem ser expressos estatisticamente por meio de proporções ou coeficientes.

Em A, o coeficiente de revocação é de apenas 3/10 (30%), mas o coeficiente de precisão é perfeito: 3/3 (100%). Em B, o coeficiente de revocação é muito alto, 9/10 (90%), mas a precisão caiu para 9/15 (60%). Os dois conjuntos de resultados expressam a realidade, no sentido de que uma busca muito genérica tenderá a recuperar muitos itens (a revocação pode ser alta, mas a precisão será baixa), enquanto

uma busca muito específica recuperará muitos poucos itens (portanto, a precisão pode ser alta, mas a revocação baixa). Para conseguir alta revocação pode ser necessário sacrificar a precisão; para conseguir alta precisão pode ser preciso sacrificar a revocação.

Ilustração 70

Há inúmeros fatores que influenciam no desempenho de um sistema de recuperação quando medido em termos de revocação e precisão. Os mais importantes estão listados aqui. Em certo sentido, o primeiro é o mais importante de todos: se um usuário do sistema fizer uma consulta que reflita de forma muito infeliz suas verdadeiras necessidades de informação, será muito improvável que os resultados da busca sejam satisfatórios, por melhores que sejam a indexação, as estratégias de busca e o vocabulário.

Falhas na busca podem ocorrer devido a problemas na indexação: um indexador pode não incluir um termo que deveria ter sido usado, pode usar um que não é apropriado, ou pode usar um que não é suficientemente específico.

A pessoa que consulta a base de dados pode fazê-lo de uma forma demasiado geral, ou demasiado específica: pode usar um termo incorreto, ou pode omitir inadvertidamente termos importantes que deveriam ser incorporados.

Uma discussão detalhada sobre esses fatores foge ao escopo deste curso. Fatores, entretanto, que afetam o desempenho de um sistema de recuperação, e que são diretamente imputáveis ao vocabulário, merecem alguma explicação.

Ilustração 71

A especificidade dos termos é a mais importante característica de um vocabulário controlado. O diagrama ilustra três níveis possíveis de especificidade. Se o termo específico **limões** existe em um tesuro, os

artigos sobre essa fruta podem ser indexados com precisão. Se o termo específico não existe, esses artigos teriam de ser indexados com muito menor precisão - sob **frutas cítricas**, ou, mesmo, só sob **frutas**.

A especificidade controla a **precisão** que pode ser conseguida na busca. Se alguém procura recuperar todos os itens numa base de dados sobre **frutas cítricas**, quando na verdade interessam apenas itens sobre limões, a taxa de precisão da busca será presumivelmente muito baixa. A situação seria ainda mais grave se o termo **frutas** fosse o de maior especificidade disponível nessa hierarquia.

A especificidade pode estar relacionada com as necessidades dos usuários do serviço de informação, razão por que os princípios da **garantia literária** e da **garantia de uso** são tão importantes. Num tesouro sobre Agricultura precisar-se-á, presumivelmente, de termos específicos para cada fruta, enquanto em outra área, como Economia, ou Comércio, por exemplo, o termo genérico **frutas** pode ser perfeitamente satisfatório. Na dúvida, é melhor ser específico demais do que de menos.

Deve-se reconhecer que um vocabulário não-específico, embora dificulte alcançar-se alta precisão, tende a produzir maior revocação. Se formos forçados a recuperar tudo sobre frutas, quando apenas itens sobre limões interessam, a revocação poderá ser bem mais alta do que seria se o termo específico existisse. Por exemplo: alguns itens sobre "limões" podem não ter sido indexados sob o termo específico por erro de indexação, e alguns dos documentos mais gerais indexados sob **frutas cítricas** podem conter informação substancial sobre limões.

Ilustração 72

Um segundo e importante requisito de um vocabulário controlado é que sua estrutura deve ser tal que proporcione a maior ajuda possível aos indexadores e aos pesquisadores.

O diagrama mostra um conjunto muito incompleto de termos relacionados com poluição. Os termos são unidos por meio de setas. Setas contínuas representam relações hierárquicas; setas interrompidas representam relações que cruzam as hierarquias.

O tesouro deve apresentar de forma explícita todas as relações indicadas (e quaisquer outras que pareçam úteis) pela sua estrutura de TG/TE/TR, ou através de outros recursos. Assim fazendo, ele ajuda o indexador a escolher o termo mais adequado para representar determinado conceito. Por exemplo: ele o conduz de **poluição da água** a **poluição térmica**, permitindo-lhe indexar mais corretamente um relatório sobre "a destruição da vida nos cursos d'água em consequência de despejos aquecidos".

Ele desempenha a mesma função em relação ao usuário, e, além do mais, desempenha outra função de importância ainda maior: mostra todos os termos necessários a uma consulta exaustiva sobre determinado assunto. Para realizar uma ampla consulta sobre poluição, diz-nos o tesouro, precisaremos incluir termos sobre vários tipos de poluição, e, talvez, também, ir buscar termos em outras hierarquias: poluentes, resíduos etc.

Se a estrutura do tesouro for imperfeita ou inadequada, isso poderá levar a erros na indexação ou no desempenho de uma busca.

Ilustração 73

Falhas na precisão podem ocorrer devido ao fato de o vocabulário (ou a maneira como ele é usado) padecer da falta de sintaxe explícita. Para compreender isso é preciso distinguir entre sistemas pré-coordenados e sistemas pós-coordenados.

Consideremos um relatório hipotético indexado sob os termos listados. Ele trata de alguns aspectos do comércio atual no Brasil, particularmente suas relações de importação/exportação com a Argentina e o México.

Visivelmente algumas das relações sugeridas pela lista são "válidas", enquanto outras são "espúrias". O relatório, por exemplo, discute as licenças de exportação brasileiras, não as argentinas, e não há relação direta alguma entre o México e a Argentina.

Ilustração 74

Aqui diversos termos foram combinados para formar entradas da forma como poderiam aparecer em um catálogo convencional em fichas ou num índice impresso. As relações espúrias são evitadas — apenas termos diretamente relacionados são postos lado a lado numa entrada, e a sequência de termos ajuda a tornar a relação explícita e inequívoca (por exemplo: o Brasil exportando para a Argentina, e não Argentina para o Brasil). Todavia ainda são possíveis interpretações diferentes. Para afastá-las por completo pode ser necessário introduzir mais sintaxe explícita.

Isso exemplifica um índice pré-coordenado: os termos são coordenados (combinados) explicitamente, e essas combinações são introduzidas no índice. Tem isso a vantagem de reduzir ambigüidades, mas traz consigo suas próprias desvantagens.

Ilustração 75

Num sistema pós-coordenado, ao contrário, todos os termos atribuídos a um documento por um indexador estão, por assim dizer, atados ao registro bibliográfico daquele documento, mas os termos não estão combinados em grupos como na Ilustração 74. Na verdade, nesse tipo de sistema, o documento de nossa hipótese poderia ser recuperado por combinação — de dois termos, três termos etc. Algumas das relações possíveis entre os oito termos atribuídos ao documento aparecem neste diagrama indicadas por meio de linhas que os unem.

Ilustração 76

Uma **falsa coordenação** ocorre quando os termos responsáveis pela recuperação do documento não se encontram diretamente relacionados —

naquele documento. Por exemplo: com base na indexação mostrada na Ilustração 73, este documento seria recuperado em resposta à consulta sobre exportação da Argentina para o México. Na verdade, nenhuma relação entre a Argentina e o México é analisada no documento; isto é, os dois termos não apresentam a menor relação.

Falsa coordenação dessa natureza pode ser evitada. A maneira mais óbvia é formar grupos com os termos de indexação, conforme mostrado, de tal forma que todos os termos de um grupo estejam inter-relacionados. Alguma letra ou número é usado para identificar cada grupo, e esses símbolos são usados para manter a identidade própria de cada grupo numa base de dados. O recurso é denominado **elo**, porque os termos inter-relacionados são, por assim dizer, encadeados uns com os outros.

Por meio desse encadeamento, a falsa coordenação mencionada anteriormente pode ser evitada, uma vez que México e Argentina aparecem em grupos diferentes. Muitas outras falsas coordenações possíveis são também evitadas: qualquer um dos termos relacionados com comércio no grupo 3 combinados com México ou Argentina.

Ilustração 77

Uma situação mais complexa, que os elos não evitam facilmente, é a de **relação incorreta com o termo**. Mesmo com a adoção dos elos da Ilustração 76, este documento poderia ser recuperado numa busca sobre exportações da Argentina para o Brasil. Na verdade, o documento trata da situação oposta: exportações do Brasil para a Argentina (importações argentinas). De novo, o problema pode ser resolvido — dessa vez por meio do uso de **indicadores de função**. No exemplo, duas funções simples foram, por assim dizer, acrescentadas ao vocabulário. Usando-as como modificadores dos descritores no registro bibliográfico, estamos agora aptos a indicar que o Brasil é o exportador e que a Argentina importa.

Um número relativamente pequeno de indicadores de função, 10 ou 12, talvez, expressando relações que chamaríamos de básicas, quando escolhidas para atender às necessidades de uma determinada área de assunto, resolverá a maioria desses problemas.

Ilustração 78

Num vocabulário formado principalmente por palavras isoladas, falsa coordenação e incorreto relacionamento de termos podem acontecer com alguma frequência. São menos prováveis de acontecer num vocabulário que adote uma razoável dose de termos compostos (ver Unidade 7). Por exemplo: se um documento for indexado sob o termo **mulheres** e o termo **poesia** ele poderá estar tratando de assuntos bem diferentes: as mulheres representadas na poesia, as mulheres como leitoras de poesia, as mulheres como autoras de poesia, e assim por diante. De igual forma, **computadores** e **desenho** poderia referir-se a desenho de computadores ou a desenho de alguma coisa (aeronaves, por exemplo) utilizando computadores como instrumento.

Muitas dessas ambigüidades são eliminadas com o maior emprego de termos compostos no vocabulário; por exemplo, por meio da inclusão de termos como **mulheres na poesia**, **desenho de computadores**, **desenho de aeronaves** e **desenho com a ajuda de computador**.

Outra solução possível envolve o uso de subcabçalhos. Um número relativamente pequeno de subcabçalhos cuidadosamente selecionados, incorporados como um elemento especial do vocabulário, pode evitar muitas ambigüidades - como nos exemplos **computador/desenho**, **aeronaves/desenho**. Em um centro de informação sobre atividades comerciais, **exportação** e **importação** seriam ótimos candidatos à adoção como subcabçalhos. Assim, **Brasil/exportações** e **Argentina/importações** oferecem uma representação sem ambigüidades.

Deve-se observar, entretanto, que nem todas essas soluções são compatíveis com as diretrizes internacionais.

Ilustração 79

O emprego de indicadores de função, ou de subcabçalhos, na verdade aumenta o tamanho e a especificidade do vocabulário. Em vez de ter-se um único termo "Brasil", temos agora diversos: Brasil na função A, na B, ou Brasil/importações, Brasil/exportações, e assim por diante.

Os subcabçalhos oferecem vantagens óbvias sobre os indicadores de função: eles são "linguagem natural" e não simbólica, e podem ser usados para subdividir entradas nos índices impressos.

Deve-se reconhecer que o emprego de elos e funções acrescenta complexidade à indexação e à busca, acarretando, dessa forma, aumento de custos. Pode-se dizer o mesmo, em escala menor, da adoção de um conjunto simples de subcabçalhos. Além do mais, o exemplo usado na ilustração original (a de número 73) é um caso extremo. Em geral, a indexação é muito menos ambígua que esta. Embora falsas coordenações e incorreto relacionamento de termos ocorram em todos os sistemas pós-coordenados, é preciso que eles ocorram com frequência razoavelmente baixa para que sejam "toleráveis" sem o uso de dispositivos adicionais, especialmente se o vocabulário contiver um razoável nível de termos compostos.

UNIDADE 14. SISTEMAS DE LINGUAGEM NATURAL

Objetivos

1. Ilustrar, mais uma vez, os principais problemas relacionados com a operacionalização de um sistema sem controle de vocabulário;
2. ilustrar técnicas que podem ser adotadas para contornar o problema da falta de controle de vocabulário nos sistemas computadorizados;
3. introduzir e explicar o vocabulário pós-controlado.

Ilustração 80

Este curso versa sobre **vocabulários controlados**, especialmente tesouros. Com recursos exclusivamente manuais é virtualmente impossível operar com êxito um sistema de recuperação de informação sem controle de vocabulário. Quando, entretanto, se emprega o computador, torna-se mais viável trabalhar sem controle de vocabulário (isto é, trabalhar com "linguagem natural"). Por exemplo: as palavras que aparecem em

títulos e resumos, armazenados numa base de dados, podem ser usadas no lugar de termos da indexação para representar o assunto dos documentos. Essas palavras podem ser procuradas em quaisquer combinações. Pode-se, portanto, pedir ao computador para localizar quaisquer resumos que contenham a palavra "corrosão" e a palavra "cobre".

O problema com um sistema que empregue inteiramente linguagem natural é que não oferece ajuda alguma ao usuário. Para recuperar todas as informações sobre "resíduos", por exemplo, teremos provavelmente que pensar não apenas nessa palavra, mas em outras como "descargas" e "efluentes". Um bom tesouro uniria esses termos de alguma forma, ou usaria um termo para representar todo o conjunto.

Da mesma forma, um tesouro nos mostraria os muitos termos diferentes necessários a uma busca exaustiva sobre "magnetismo".

O usuário de um sistema que adota linguagem natural deve compensar a falta de controle de vocabulário usando sua engenhosidade.

Ilustração 81

↳ Não obstante, existem algumas técnicas de busca em sistemas de linguagem natural que ajudam a compensar a falta de um vocabulário controlado. A mais poderosa é o **truncamento**, ou **busca apoiada em fragmentos de palavras**, que permite fazer a busca em partes das palavras, bem como em palavras completas. Existem três principais tipos de truncamento:

1. Busca baseada nos caracteres iniciais da palavra (**truncamento à direita**): todas as palavras que começam com "magnet" poderiam ser recuperadas;
2. busca baseada nos caracteres finais (**truncamento à esquerda**): todas as palavras que terminam com "magneto" poderiam ser recuperadas;
3. busca baseada na cadeia de caracteres que compõem o conjunto "magnet", onde quer que ela ocorra numa palavra (truncamento à direita e à esquerda ao mesmo tempo). Além de recuperar palavras que começam e terminam com "magnet", seriam recuperadas também as palavras com "magnet" no meio.

Neste exemplo, então, a busca baseada em fragmentos de palavras seria um recurso poderoso, porque o fragmento "magnet" muito provavelmente permitiria recuperar a maioria dos documentos que têm algo a ver com magnetismo.

Ilustração 82

É evidente que podem ocorrer falsas coordenações na busca em resumos; quanto mais longo o resumo, maior a probabilidade. Alguns oferecem ao usuário a oportunidade de reduzir essa probabilidade permitindo-lhe especificar o grau de proximidade com que as duas palavras devem aparecer no texto (por exemplo: com interposição de não mais de 10 palavras). Quanto mais próximas duas palavras aparecem num resumo, tanto maior a chance de que elas estejam diretamente inter-relacionadas. Todos os sistemas deveriam permitir pelo menos a "adjacência", por exemplo, a palavra "ligas" seguida imediatamente de "cobre".

Ilustração 83

É possível ter-se um sistema de recuperação "híbrido", no sentido de que adota ao mesmo tempo termos controlados e linguagem natural. Por exemplo, algumas bases de dados incorporam descritores e palavras-chave nelas introduzidos por processos manuais (não-controlados), enquanto outras combinam descritores com o texto dos resumos.

É também possível sobrepor a um sistema de linguagem natural um tipo de "superestrutura" de vocabulário controlado — um conjunto relativamente pequeno de códigos muito gerais atribuídos por indexadores. Isso ajuda a reduzir a ambigüidade (por exemplo, quando a palavra "piquete" ocorre em conexão com o código para "trabalho", é muito provável que ela signifique algo muito diferente de quando ela ocorre associada a "atividades bélicas", e facilita enormemente o desenvolvimento de uma busca ampla (por exemplo: todos os aspectos de atividades bélicas relacionadas com determinada região).

Ilustração 84

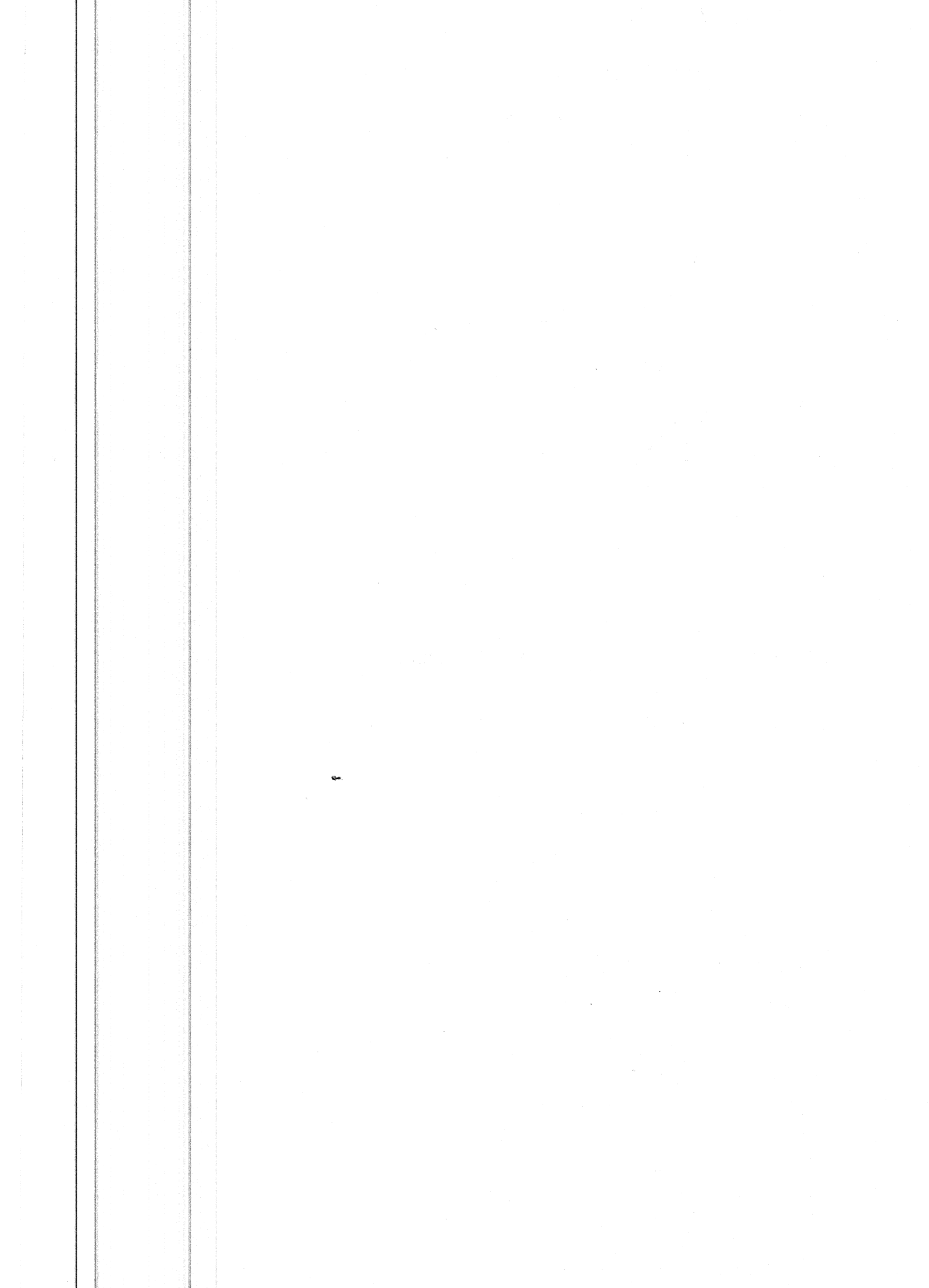
Para ajudar na busca em bases de dados que adotam linguagem natural, é possível, também, construir (e armazenar em computador) tabelas de sinônimos, de quase-sinônimos ou outros termos conceitualmente inter-relacionados. Eles podem incluir tanto fragmentos de palavras quanto palavras inteiras. São dados aqui dois exemplos, onde o sinal * significa truncamento. A primeira tabela é destinada a recuperar documentos que tenham algo a ver com "resíduos", enquanto a segunda tem probabilidade de recuperar a maior parte dos resumos que tratam de alguma forma de antibióticos.

Essas tabelas recebem nomes ou números que as identificam. Elas podem ser mostradas em linha ou incluídas numa estratégia de busca, exatamente da mesma forma como uma hierarquia inteira extraída de um tesouro.

É evidente que essas tabelas atendem a algumas das mesmas finalidades as quais os tesouros se destinam. Na verdade elas podem ser consideradas como um tipo bem diferente de "vocabulário controlado"— um **vocabulário pós-controlado**, destinado a ajudar na busca em bases de dados, mas não a impor controle algum sobre a terminologia introduzida no sistema.

À medida que o custo do processamento intelectual humano continua a crescer com relação ao custo do processamento por meio do computador, é provável que os serviços de informação se voltem cada vez mais para a linguagem natural, afastando-se dos vocabulários controlados tradicionais. Em decorrência, maior atenção será dada à construção de **vocabulários pós-controlados**. Poderá então surgir um tipo de tesouro bem diferente, muito menos estruturado.

ILUSTRAÇÕES



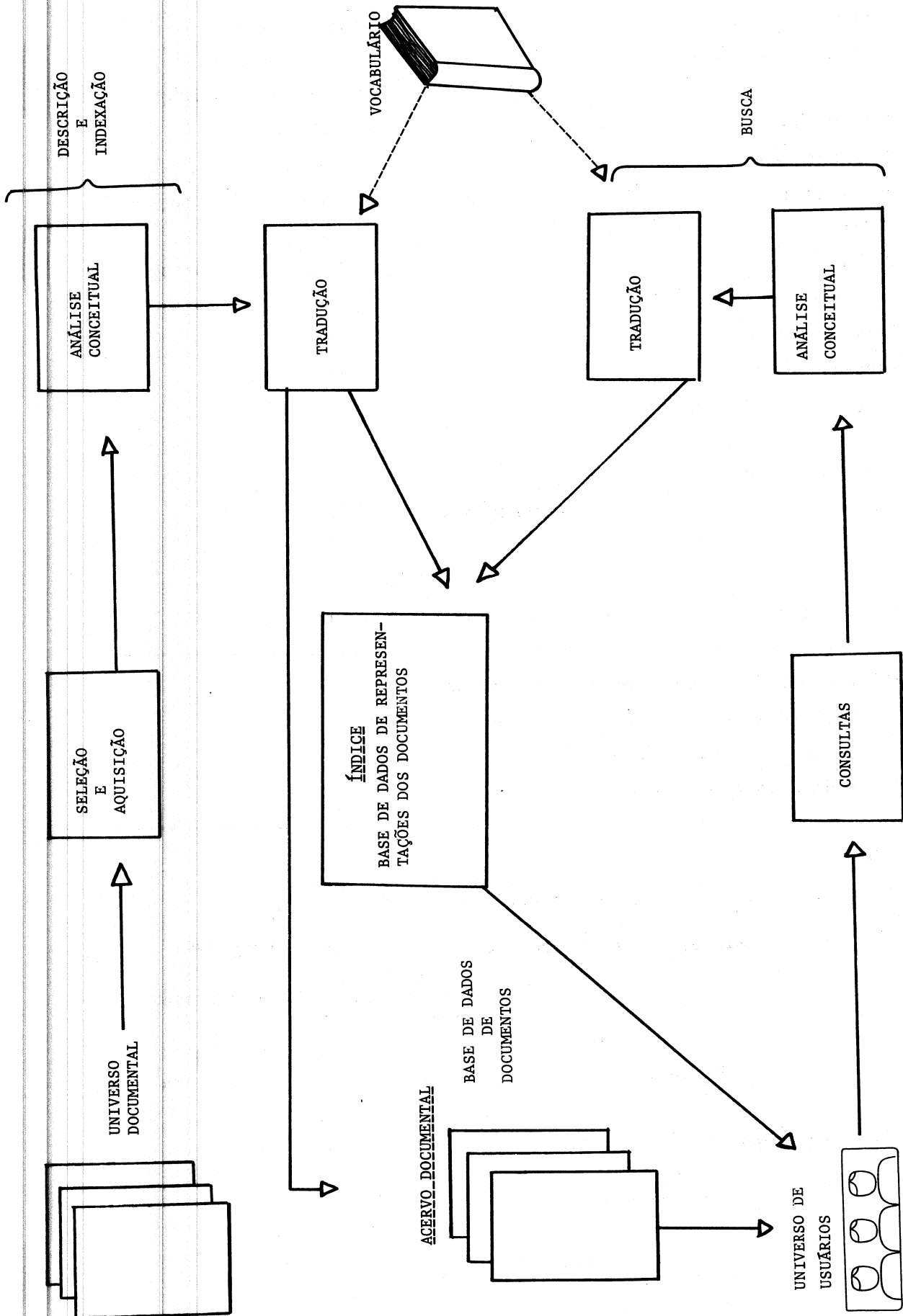


Ilustração 1

Corrosão
 Corrosão por fadiga
 Defeito
 Deformação
 Deterioração
 Ductilidade
 Fadiga
 Falha
 Fendimento
 Fragilidade
 Fragilidade cáustica
 Fragilidade da têmpera

Fratura
 Fratura celular
 Fratura quebradiça
 Imperfeições
 Resistência
 Resistência de tração
 Ruptura
 Ruptura por deformação
 Ruptura por fadiga
 Tensão

Ilustração 2

(a) Defeito \longrightarrow Falha

(b) Ruptura (fratura; por ex.: dos metais)
 Ruptura (violação ou infração de um contrato)
 Fadiga (nos seres humanos)
 Fadiga (nos materiais)

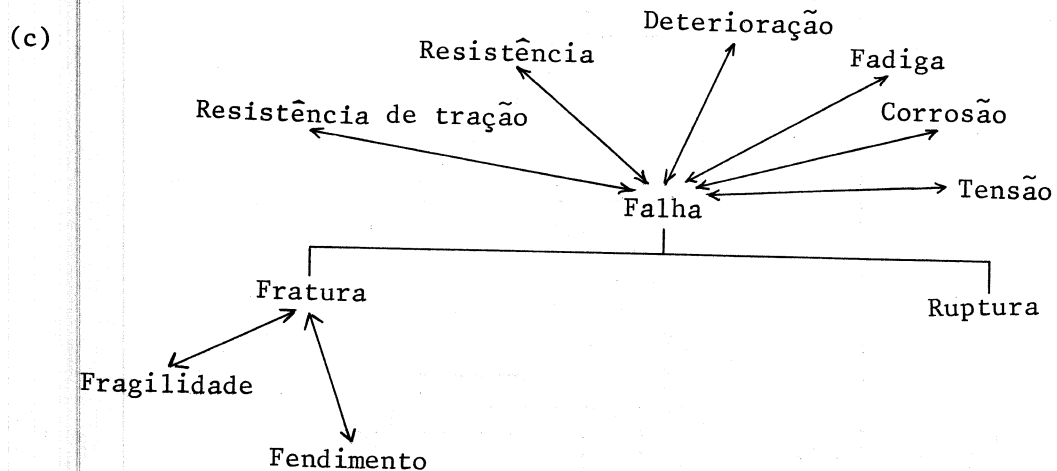


Ilustração 3

1. Promover representações consistentes

- . na entrada (indexação)
- . na saída (busca)

através de:

- . controle de sinônimos
- . distinção dos homógrafos

2. Facilitar buscas exaustivas com a associação semântica dos termos

inter-relacionados:

- . relação gênero/espécie
- . outros tipos

Ilustração 4

Componentes complementares

- (a) alfabéticos
- (b) classificados

Podem ser:

- . ostensivamente classificados, com índice alfabético
- . ostensivamente alfabéticos, com classificação implícita

Ilustração 5

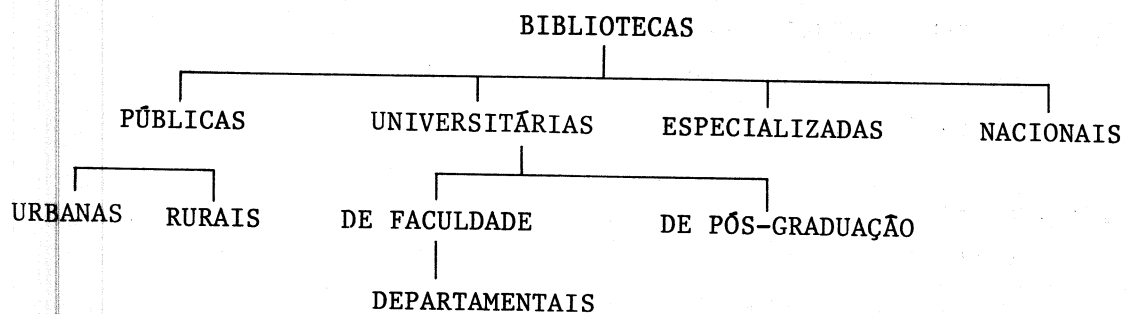


Ilustração 6

Bibliotecas

Bibliotecas especializadas

Bibliotecas nacionais

Bibliotecas públicas

 Bibliotecas públicas rurais

 Bibliotecas públicas urbanas

Bibliotecas universitárias

 Bibliotecas de faculdade

 Bibliotecas departamentais

 Bibliotecas de pós-graduação

Ilustração 7

B	Bibliotecas
Bc	Bibliotecas públicas
Bcc	Bibliotecas públicas urbanas
Bcr	Bibliotecas públicas rurais
Be	Bibliotecas universitárias
Bec	Bibliotecas de faculdade
Becj	Bibliotecas departamentais
Beu	Bibliotecas de pós-graduação
Bg	Bibliotecas especializadas
Bi	Bibliotecas nacionais

Ilustração 8

Bibliotecas	B
Bibliotecas departamentais	Becj
Bibliotecas especializadas	Bg
Bibliotecas de faculdade	Bec
Bibliotecas nacionais	Bi
Bibliotecas de pós-graduação	Beu
Bibliotecas públicas	Bc
Bibliotecas públicas rurais	Bcr
Bibliotecas públicas urbanas	Bcc
Bibliotecas universitárias	Be

Ilustração 9

Bibliotecas

TE Bibliotecas especializadas
Bibliotecas nacionais
Bibliotecas públicas
Bibliotecas universitárias

Bibliotecas departamentais

TG Bibliotecas de faculdade

Bibliotecas especializadas

TG Bibliotecas

Bibliotecas de faculdade

TE Bibliotecas departamentais

TG Bibliotecas universitárias

Bibliotecas nacionais

TG Bibliotecas

Bibliotecas de pós-graduação

TG Bibliotecas universitárias

Bibliotecas públicas

TG Bibliotecas

Bibliotecas públicas rurais

TE Bibliotecas públicas urbanas

Bibliotecas públicas urbanas

TG Bibliotecas públicas

Bibliotecas universitárias

TG Bibliotecas

TE Bibliotecas de faculdade

Bibliotecas de pós-graduação

Ilustração 10

(a) Bibliotecas

Bibliotecas públicas

Bibliotecas públicas urbanas —————> Governo municipal

Bibliotecas públicas rurais

Bibliotecas universitárias —————> Educação superior

Bibliotecas de faculdade

Bibliotecas de pós-graduação

(b) Bibliotecas universitárias

TG Bibliotecas

TE Bibliotecas de faculdade

Bibliotecas de pós-graduação

TR Educação superior

Bibliotecas públicas urbanas

TG Bibliotecas públicas

TR Governo municipal

Ilustração 11

1. Adotar um tesouro já existente
2. Modificar (por exemplo, expandir) um tesouro já existente
3. Traduzir um tesouro já existente
4. Construir um tesouro novo

Ilustração 12

1. Coleta de termos
2. Ordenação dos termos
3. Produção de uma estrutura final
4. Impressão e apresentação

Ilustração 13

1. Abordagem "teórica" de cima para baixo (dedutiva)
2. Abordagem empírica, de baixo para cima (indutiva)
Garantia literária
Garantia de uso
Combinação das garantias literária e de uso

Ilustração 14

Garantia literária

Obras de referências: dicionários, glossários, enciclopédias, manuais, livros-texto, índices

Resumos

Listagens impressas de bases de dados

De ocorrência de termos

----- 1172

----- 1107

----- 1049

----- 1002

De co-ocorrência de termos

A ----- 0

----- R

----- S

Ilustração 15

Garantia de uso

1. Pedidos de busca
2. Questionário a uma amostra de usuários
 - Descrição de tarefas
 - Interesses na pesquisa
 - Títulos/resumos de publicações
3. Os usuários escolhem os termos na literatura

Ilustração 16

Experiência na construção de tesouros

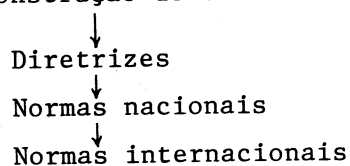
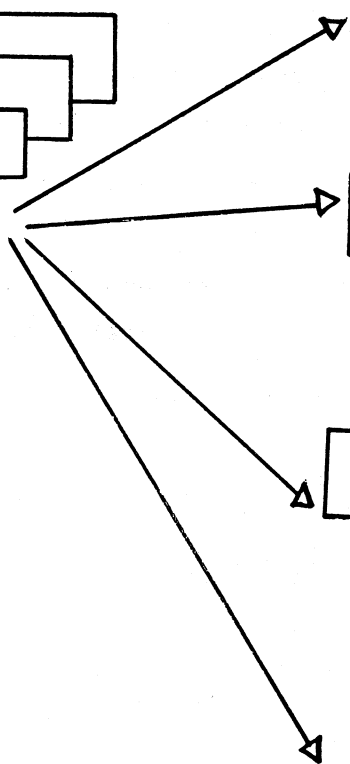
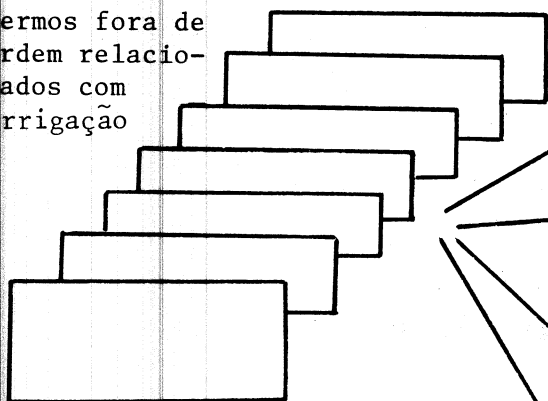


Ilustração 17

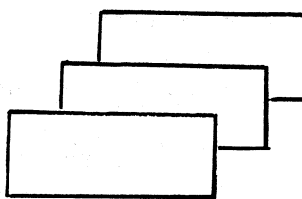
1. Coletar os termos
2. Identificar as principais categorias (classes) de termos
3. Estabelecer hierarquias dentro das categorias
4. Estabelecer relações úteis entre os termos de diferentes hierarquias
5. Converter para a ordem alfabética

Ilustração 18

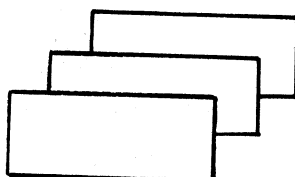
Termos fora de
ordem relacio-
nados com
irrigação



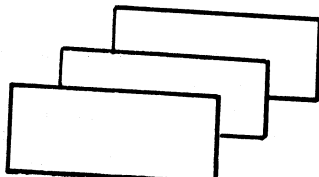
Tipos de
irrigação



Tipos de
produtos
agrícolas



Tipos de
solo



Tipos de
clima

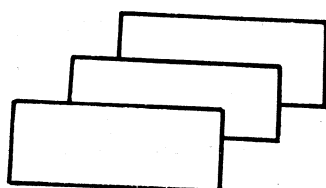


Ilustração 19

Produtos agrícolas

Frutas

Frutas cítricas

Laranjas

Limas

Limões

Toranjas

Legumes

Cereais

Ilustração 20

Cereais

TG Produtos agrícolas

Frutas

TG Produtos agrícolas

TE Frutas cítricas

Frutas cítricas

TG Frutas

TE Laranjas

Limas

Limões

Toranjas

Laranjas

TG Frutas cítricas

Legumes

TG Produtos agrícolas

Limas

TG Frutas cítricas

Limões

TG Frutas cítricas

Produtos agrícolas

TE Frutas

Toranjas

TG Frutas cítricas

Ilustração 21

Produtos agrícolas

TE Cereais
Frutas
Laranjas
Legumes
Limas
Limões
Toranjas

Produtos agrícolas

TE Cereias
Frutas
Frutas cítricas
Laranjas
Limas
Limões
Toranjas
Legumes

Ilustração 22

Produtos Agrícolas

TE1 Frutas
TE2 Frutas cítricas
TE3 Laranjas
Limas
Limões
Toranjas
TE1 Cereais
TE1 Legumes

Ilustração 23

1. Relação genérica
2. Relação partitiva
3. Relação exemplar

Ilustração 24

América do Sul

Venezuela

Caracas

Sistema cardiovascular

Coração

Aorta

Ilustração 25

Navios

TEG

Cargueiros

Navios de guerra

Navios de passageiros

TEP

Casco

Proa

Tombadilho

Ilustração 26

Catedrais

TE

Catedral de Brasília

Catedral do Rio de Janeiro

Catedral de São Paulo

Ilustração 27

Arquitetura

[Por estilo]

Barroca

Gótica

Românica

[Por aplicação]

Residencial

Religiosa

Naval

[Por país]

Alemã

Norte-americana

Francesa

Ilustração 28

Arquitetura

TE

Arquitetura Barroca

Arquitetura Francesa

Arquitetura Gótica

Arquitetura Norte-americana

Arquitetura Religiosa

Arquitetura Residencial

etc.

Ilustração 29

Arquitetura

TE (por aplicação)

Arquitetura Naval
Arquitetura Religiosa
Arquitetura Residencial

TE (por estilo)

Arquitetura Barroca
Arquitetura Gótica
Arquitetura Românica

Ilustração 30

Diamantes

TG Instrumentos de corte

Diamantes

TG Pedras
Preciosas

Ilustração 31

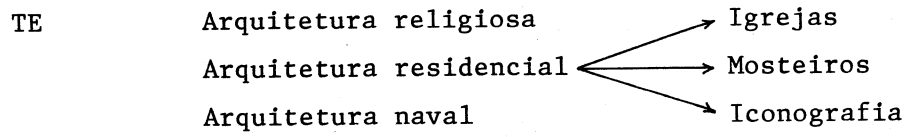
Laranjas

TS Produtos agrícolas

TG Frutas cítricas

Ilustração 32

Arquitetura



Arquitetura religiosa

- | | |
|----|-------------|
| TG | Arquitetura |
| TR | Iconografia |
| | Igrejas |
| | Mosteiros |

Ilustração 33

Coisa/aplicação

Abastecimento de água
TR Irrigação

Efeito/causa

Febre
TR Infecção

Coisa/propriedade

Criança superdotada
TR Inteligência

Matéria-prima/produto

Bauxita
TR Alumínio

Atividades complementares

Compra
TR Venda

Opostos

Vida
TR Morte

Atividade/propriedade

Corte
TR Usinabilidade

Atividade/agente

Tabagismo
TR Fumo

Atividade/produto

Tear
TR Tecido

Todo/parte

Aeronaves
TR Asas

Ilustração 34

Substantivos ou locuções

Gravidade (substantivo)

Gravidade específica (locução com ajuda de adjetivo)

Centro de gravidade (locução com ajuda de preposição)

Ordem direta, não-invertida

Impressão de livros

Não: Livros, impressão

Venda de livros

Não: Livros, venda

Ilustração 35

Plural

Substantivos que encerram
a idéia de quantidade
("contáveis")

Quantos (as)?

Cães

Ferrovias

Ratos

Singular

Substantivos que encerram
a idéia de volume

Que quantidade de?

Farinha

Milho

Papel

Ilustração 36

Cooperação (entre bibliotecas)
Arquitetura (de bibliotecas)
Prédios (de bibliotecas)

Ilustração 37

- | | | |
|------------------------------------|---------|------------------|
| 1. Ruído dos motores de automóveis | | |
| 2. Motores de automóveis | | Ruído |
| 3. Automóveis | | Ruído de motores |
| 4. Automóveis | Motores | Ruído |

Ilustração 38

Separar o composto

1. Ruído do tráfego

em

Tráfego

Ruído

2. Livros infantis ilustrados

em

Livros ilustrados

Livros infantis

3. Corrosão de palheta de turbina

em

Corrosão

Palheta de turbina

Manter o composto

4. Acidez do papel
5. Lírios-do-mar
6. Lâmina de vidro
7. Cavalos-marinhos
8. Doutrina de Monroe
9. Recipientes de pressão

Ilustração 39

Vocabulário de entrada

1. Sinônimos e quase-sinônimos
2. Termos desnecessariamente considerados específicos no contexto do tesouro

Ilustração 40

1. Carros USE Automóveis
2. PNB USE Produto Nacional Bruto
3. Portagem USE Pedágio
4. Acústica USE Som
5. Movimento, equações do USE Equações do movimento
6. Inelasticidade USE Elasticidade
7. Arquitetura residencial USE Arquitetura e Residências
8. Ruído branco, USE Ruído

Ilustração 41

Carros USE Automóveis

Automóveis UP Carros

Arquitetura residencial USE Arquitetura e Residências

Arquitetura

UP+ Arquitetura residencial

Residências

UP+ Arquitetura residencial

Ilustração 42

PERMAFROST

NE Terreno ou rocha perene ou permanentemente gelados

GATOS

NE Usar apenas em estudos sobre a família do gato em geral, incluindo-se aí o gato doméstico e o selvagem. Sempre que mais indicado, preferir os termos mais específicos, como GATOS DOMÉSTICOS, LEÕES, TIGRES.

OBSOLESCÊNCIA

NE A obsolescência do material bibliográfico, medido em termos de declínio do uso com a passar dos anos. Para declínio físico do material bibliográfico USE DETERIO-RAÇÃO.

Ilustração 43

Tanques (armas)

Tanques (recipientes)

Raios (elementos

estruturais

Raios (eletrônica)

Ilustração 44

Identificadores

1. Nomes de armas e outros equipamentos
2. Nomes próprios (por exemplo: autores, personagens históricos)
3. Nomes geográficos
4. Compostos químicos
5. Instituições

Ilustração 45

Instrumentos para localização de posição

NE	Instrumentos empregados na localização de aeronaves, navios ou outros objetos em relação a um ponto ou pontos de referência
UP	Indicadores de posição
TG	Instrumentos
TE	Bússolas Indicadores de posição de um alvo Indicadores de posição no plano Indicadores de posição no solo Sextantes
TR	Deteção Rastreamento

Ilustração 46

IRRIGATION

IRRIGATION (cont.)
 RT: DAMS
 DRAINAGE
 IRRIGATION CANALS
 IRRIGATION FARMING
 IRRIGATION RESEARCH
 SPRAYING EQUIPMENT
 WATER STORAGE
 TT: HYDRAULICS
 SCIENCE AND TECHNOLOGY

IRRIGATION CANALS
 CANAUX D'IRRIGATION
 CANALES DE RIEGO
 04.02.01
 RT: CANALS
 IRRIGATION

IRRIGATION ENGINEERING
 04.02.01
 USE: IRRIGATION

IRRIGATION FARMING
 CULTURE IRRIGUEE
 AGRICULTURA DE RIEGO
 04.02.01
 BT: CULTIVATION SYSTEMS
 RT: IRRIGATION
 TT: CULTIVATION SYSTEMS

IRRIGATION LAW
 LEGISLATION DE L'IRRIGATION
 LEGISLACION DEL RIEGO
 04.01.01
 BT: WATER LAW
 RT: WATER RIGHTS
 TT: LAW

IRRIGATION RESEARCH
 RECHERCHE SUR L'IRRIGATION
 INVESTIGACION SOBRE EL RIEGO
 04.01.02
 BT: AGRICULTURAL RESEARCH
 RT: IRRIGATION
 TT: RESEARCH

ISLAM
 ISLAM
 ISLAM
 15.03.00
 UF: MOHAMMEDANISM
 BT: RELIGION
 RT: ISLAMIC ART
 ISLAMIC CIVILIZATION
 ISLAMIC COUNTRIES
 ISLAMIC EPISTEMOLOGY
 ISLAMIC ETHICS
 ISLAMIC PRAYERS
 ISLAMIC SOCIOLOGY
 ISLAMIC THEOLOGY
 KORAN
 MUSLIMS
 PANISLAMISM
 TT: RELIGION

ISLAMIC ANTIQUITIES
 ANTIQUITES ISLAMIQUES
 ANTIQUIDADES ISLAMICAS
 15.01.02
 RT: ARCHAEOLOGY
 CHRISTIAN ANTIQUITIES

ISLAMIC ART
 ART ISLAMIQUE
 ARTE ISLAMICO
 15.01.02
 RT: ISLAM

ISLAMIC CIVILIZATION
 CIVILISATION ISLAMIQUE
 CIVILIZACION ISLAMICA
 15.00.00
 BT: CIVILIZATION
 RT: ARABIC LITERATURE
 ISLAM
 TT: CIVILIZATION

ISLAMIC COUNTRIES
 PAYS ISLAMIQUES
 PAISES ISLAMICOS
 15.03.00

ISLAMIC COUNTRIES (cont.)
 RT: ARAB COUNTRIES
 ISLAM

ISLAMIC COURTS
 TRIBUNAUX ISLAMIQUES
 TRIBUNALES ISLAMICOS
 01.07.00
 BT: COURTS
 RT: ISLAMIC LAW
 TT: JUDICIAL SYSTEM

ISLAMIC EPISTEMOLOGY
 EPISTEMOLOGIE ISLAMIQUE
 EPISTEMOLOGIA ISLAMICA
 15.03.00
 BT: PHILOSOPHY
 RT: ISLAM
 THEORY OF KNOWLEDGE
 TT: HUMANITIES

ISLAMIC ETHICS
 MORALE ISLAMIQUE
 ETICA ISLAMICA
 15.03.00
 BT: ETHICS
 RT: ISLAM
 TT: ETHICS

ISLAMIC LAW
 DROIT ISLAMIQUE
 DERECHO ISLAMICO
 01.07.00
 BT: LAW
 RT: FATWAS
 ISLAMIC COURTS
 KORAN
 TT: LAW

ISLAMIC PRAYERS
 PRIERES ISLAMIQUES
 ORACIONES ISLAMICAS
 15.03.00
 RT: ISLAM
 KORAN

ISLAMIC SOCIOLOGY
 SOCIOLOGIE ISLAMIQUE
 SOCIOLOGIA ISLAMICA
 14.05.01
 BT: SOCIOLOGY
 RT: ISLAM
 TT: SOCIAL SCIENCES

ISLAMIC THEOLOGY
 THEOLOGIE ISLAMIQUE
 TEOLOGIA ISLAMICA
 15.03.00
 BT: THEOLOGY
 RT: ISLAM
 TT: THEOLOGY

ISLANDS
 ILES
 ISLAS
 03.01.03 01.07.03
 BT: LANDFORMS
 NT: ARCHIPELAGOS
 BARRIER ISLANDS
 CORAL REEFS AND ISLANDS
 DEVELOPING ISLAND COUNTRIES
 TT: LANDFORMS

ISOTOPE SEPARATION
 SEPARATION ISOTOPIQUE
 SEPARACION DE ISOTOPOS
 16.08.00
 BT: NUCLEAR TECHNOLOGY
 RADIOCHEMISTRY
 RT: ISOTOPES
 RADIOACTIVE MATERIALS
 RADIOACTIVE TRACERS
 RADIOISOTOPES
 TT: SCIENCE AND TECHNOLOGY

ISOTOPES
 ISOTOPES
 ISOTOPOS
 16.08.00
 NT: RADIOISOTOPES
 STRONTIUM ISOTOPES

ISOTOPES (cont.)
 RT: ATOMS
 HYDROGEN
 ISOTOPE SEPARATION
 METALS

ISRAEL
 ISRAEL
 ISRAEL
 17.05.00
 BT: MEDITERRANEAN COUNTRIES
 RT: ISRAEL-ARAB WAR (1948-1949)
 ISRAEL-ARAB WAR (1967)
 ISRAEL-ARAB WAR (1973)
 ISRAELIS
 JEWISH LAW
 JORDAN RIVER QUESTION
 MIDDLE EAST
 MIDDLE EAST SITUATION
 PALESTINE QUESTION
 PALESTINIANS
 ZIONISM
 TT: MEDITERRANEAN COUNTRIES

ISRAEL-ARAB WAR (1948-1949)
 GUERRE ISRAELO-ARABE (1948-1949)
 GUERRA ARABE-ISRAELI (1948-1949)
 01.06.01
 SN: Used for works dealing with the war itself and its immediate results. Otherwise, use "Palestine question"
 RT: ARAB COUNTRIES
 ISRAEL
 PALESTINE QUESTION

ISRAEL-ARAB WAR (1967)
 GUERRE ISRAELO-ARABE (1967)
 GUERRA ARABE-ISRAELI (1967)
 01.06.01
 SN: Used for works dealing with the war itself and its immediate results. Otherwise, use "Palestine question" or "Middle East situation". For the occupied territories, use "Territories occupied by Israel"
 RT: ARAB COUNTRIES
 ISRAEL
 MIDDLE EAST SITUATION
 PALESTINE QUESTION

ISRAEL-ARAB WAR (1973)
 GUERRE ISRAELO-ARABE (1973)
 GUERRA ARABE-ISRAELI (1973)
 01.06.01
 SN: Used for works dealing with the war itself and its immediate results. Otherwise, use "Palestine question" or "Middle East situation"
 RT: ARAB COUNTRIES
 ISRAEL
 MIDDLE EAST SITUATION
 PALESTINE QUESTION

ISRAELIS
 ISRAELIENS
 ISRAELIES
 08.03.02
 RT: ISRAEL

ITALIAN LANGUAGE
 LANGUE ITALIENNE
 LENGUA ITALIANA
 15.01.01
 BT: ROMANCE LANGUAGES
 TT: HUMANITIES

ITALIANS
 ITALIENS
 ITALIANOS
 08.03.02
 RT: ITALY

ITALY
 ITALIE
 ITALIA
 17.04.00
 BT: MEDITERRANEAN COUNTRIES
 WESTERN EUROPE
 RT: ITALIANS
 TT: EUROPE
 MEDITERRANEAN COUNTRIES

BT(Broader Term) = TG(Termo genérico)
 NT(Narrower Term) = TE(Termo específico)
 RT(Related Term) = TR(Termo relacionado)
 TT(Top Term) = TS(Termo superior)
 UF(Used For) = UP(Usado por)

Unbis Thesaurus - Categorized List of Descriptors

04.01.02 (cont.)

AGRICULTURAL TECHNOLOGY
 USE: AGRICULTURAL ENGINEERING
AGRICULTURAL TRAINING
AGROCLIMATOLOGY
 USE: AGROMETEOROLOGY
AGROMETEORODOLOGY
AGRONOMY
 SN: The science of crop and soil management
FARM MECHANIZATION
 USE: AGRICULTURAL MECHANIZATION
GENETIC IMPROVEMENT
GREEN REVOLUTION
 SN: Promotion of high yield grain crops
IRRIGATION RESEARCH
RICE RESEARCH
SPRAYING EQUIPMENT
SUGAR RESEARCH
TRACTORS
 SN: Farm tractors

04.02.01
CROP MANAGEMENT

AGRICULTURAL CROPS
 USE: CROPS
COMPOSTS
 SN: Material produced by a controlled process of degrading organic matter by microorganisms
CROP DIVERSIFICATION
CROP MANAGEMENT
CROP ROTATION
CROP SUBSTITUTION
 SN: Replacement of illicit narcotic raw materials with non-narcotic crops
CROP YIELDS
CROP ZONES
CROPPING SYSTEMS
 USE: CULTIVATION SYSTEMS
CROPS
CULTIVATION SYSTEMS
DEFOLIANTS
DRY FARMING
 SN: A system of producing crops in arid and semi-arid regions without the use of irrigation
FERTILIZER USE
FERTILIZERS
FIELD CROPS
FLORICULTURE
 SN: Ornamental gardening
FRUIT CULTURE
GREENHOUSES
HORTICULTURE
 SN: Cultivation of gardens or orchards, including the growing of vegetables, fruits, flowers and ornamental shrubs and trees
HOT HOUSES
 USE: GREENHOUSES
IRRIGATION
IRRIGATION CANALS
IRRIGATION ENGINEERING
 USE: IRRIGATION
IRRIGATION FARMING
NITROGEN FERTILIZERS
 USE: FERTILIZERS
NUCLEAR AGRICULTURE
ORGANIC FARMING
ORGANIC GARDENING
 USE: ORGANIC FARMING
PLANT BREEDING
PLANT DISEASES
PLANT GENETICS
PLANT NAMES
PLANT NUTRITION
PLANT PATHOLOGY
 USE: PLANT DISEASES
PLANT PROTECTION
RAIN FED FARMING
 USE: DRY FARMING
ROTATION OF CROPS
 USE: CROP ROTATION
SERICULTURE
SHIFTING CULTIVATION
SUGAR GROWING
 SN: Refers to problems related to growing sugar crops
TRICKLE IRRIGATION
VITICULTURE

04.02.02
CROPS

ABACA
APPLES
APRICOTS
ARTICHOKES
ASPARAGUS
BAGASSE
 SN: Fibrous residue resulting from the extraction of juice from the sugar cane
BALANITES AEGYPTIACA
BAMBOO
BANANAS
BARLEY
BEANS
BRUSSELS SPROUTS
CACAO
 USE: COCOA
CARDAMOM
CARROTS
CASSAVA
CASSAVA STARCH
 USE: TAPIOCA
CASTOR BEANS
CELERY
CEREALS
 USE: GRAINS
CHERRIES
CINNAMON
CITRUS FRUITS
CLOVES
COARSE GRAINS
COCOA
COCONUT PRODUCTS
COFFEE
COIR
COPRA
COTTON
DATE PALMS
DATES
FLAX
FLOWERS
FORAGE CROPS
 SN: Crops cultivated to provide leaves and stalks for feeding livestock. For forage plants in general, use "Pasture and forage"
FORAGE PLANTS
 USE: FORAGE CROPS
PASTURE AND FORAGE
FRUIT
FRUIT TREES
GARLIC
GRAINS
GRAPES
GRASSES
GROUNDNUTS
HARD FIBRES
HEMP
 SN: Use with respect to cannabis sativa as a fibre plant as well as the fibre obtained therefrom. For cannabis sativa as a drug plant, use "cannabis"
HENEQUEN
HORSE RADISH
JUTE
KENAF
LALOBE
 USE: BALANITES AEGYPTIACA
LEEKs
LEGUMES
MAIZE
MILLET
MOLASSES
NUTMEG
NUTS
OATS
OILSEED PLANTS
OILSEEDS
OLIVES
ONIONS
ORANGES
PEACHES
PEANUTS
 USE: GROUNDNUTS
PEARS
PEAS
PEPPER
PLANT PRODUCTS
PLANTS
PLUMS
POISONOUS PLANTS

POTATOES

PULSES
 USE: LEGUMES
RAISINS
RAMIE
RICE
ROOT CROPS
RYE
SEEDS
SISAL
SORGHUM
SOYBEANS
SPICES
STRAW
SUGAR
 SN: Use only for discussion of sugar as a commodity, or, in agricultural context when sugar beets or sugar cane cannot be identified

SUGAR BEETS
SUGAR CANE
SWEET POTATOES
TAPIOCA
TEA
TOBACCO
TRITICALES
TROPICAL CROPS
 SN: Use with respect to the cultivation of tropical crops. For tropical vegetation in general, use "tropical plants"

TROPICAL FRUITS
TUBERS
 USE: ROOT CROPS
TURMERIC
VEGETABLES
WHEAT
WINGED BEANS
YAMS

04.03.00
LIVESTOCK

AFTOSA FEVER
 USE: FOOT AND MOUTH DISEASE
ANIMAL DISEASES
ANIMAL FEED
 SN: Use with respect to processed or prepared food for livestock. For grasses or the leaves and stalks of plants on which livestock graze, use "pasture and forage"
ANIMAL FEED INDUSTRY
ANIMAL GENETICS
ANIMAL HEALTH
ANIMAL HUSBANDRY
ANIMAL INDUSTRY
 USE: LIVESTOCK INDUSTRY
ANIMAL NUTRITION
ANIMAL PRODUCTION
 USE: LIVESTOCK INDUSTRY
ANIMAL PRODUCTS
ANIMAL TRYPANOSOMIASIS
APICULTURE
BEE CULTURE
 USE: APICULTURE
BEEF
BREEDING STOCK
BUTTER
CAMELS
CATTLE
CATTLE FEED
CHEESE
DAIRY FARMS
DAIRY PRODUCTS
DAIRYING
DRAFT ANIMALS
 USE: DRAUGHT ANIMALS
DRAUGHT ANIMALS
EGGS
FEED SUPPLY
FOOT AND MOUTH DISEASE
FORAGE PLANTS
 USE: FORAGE CROPS
PASTURE AND FORAGE
GRASSLAND MANAGEMENT
 USE: RANGE MANAGEMENT
GRAZING
HONEY
HORSES
LAMB
 SN: Meat product
LIVESTOCK
LIVESTOCK INDUSTRY

SN(Scope Note)= NE(Nota Explicativa)

Ilustração 48

- SURFACE WATERS (cont.)**
- .. LAKE ITTICA
 - .. LAKE VICTORIA
 - .. MAN-MADE LAKES
 - .. OCEANS
 - .. ARCTIC OCEAN
 - .. ATLANTIC OCEAN
 - .. NORTH ATLANTIC OCEAN
 - .. BALTIC SEA
 - .. BLACK SEA
 - .. CARIBBEAN SEA
 - .. MEDITERRANEAN SEA
 - .. ADRIATIC SEA
 - .. AEGEAN SEA
 - .. NORTH SEA
 - .. SOUTH ATLANTIC OCEAN
 - .. INDIAN OCEAN
 - .. ARABIAN SEA
 - .. GULF OF ADEN
 - .. PERSIAN GULF
 - .. RED SEA
 - .. PACIFIC OCEAN
 - .. NORTH PACIFIC OCEAN
 - .. EAST CHINA SEA
 - .. SOUTH PACIFIC OCEAN
 - .. SOUTH CHINA SEA
 - .. RESERVOIRS
 - .. RIVERS
 - .. INTERNATIONAL RIVERS
 - .. AMAZON RIVER
 - .. AMUR RIVER
 - .. BRAHMAPUTRA RIVER
 - .. CONGO RIVER
 - .. DANUBE RIVER
 - .. EUHRATES RIVER
 - .. GANGES RIVER
 - .. INDUS RIVER
 - .. JORDAN RIVER
 - .. MEKONG RIVER
 - .. NIGER RIVER
 - .. NILE RIVER
 - .. ORINOCO RIVER
 - .. PARAGUAY RIVER
 - .. PARANA RIVER
 - .. RHINE RIVER
 - .. RIO GRANDE
 - .. SAINT LAWRENCE SEAWAY
 - .. SALWEEN RIVER
 - .. SHEBELI RIVER
 - .. TIGRIS RIVER
 - .. UBANGI RIVER
 - .. YUKON RIVER
 - .. ZAMBESI RIVER
 - .. WETLANDS
 - .. MANGROVE SWAMPS
- SYNTHETICS**
- .. SYNTHETIC FIBRES
 - .. SYNTHETIC RUBBER
- TAX CREDITS**
- .. INVESTMENT TAX CREDIT
- TAX DEDUCTIONS**
- .. DEPRECIATION ALLOWANCES
- TEACHING MATERIALS**
- .. TEXTBOOKS
- TEACHING METHODS**
- .. AUDIOVISUAL INSTRUCTION
 - .. PROGRAMMED INSTRUCTION
- TECHNICAL CO-OPERATION**
- .. AGRICULTURAL TECHNICAL CO-OPERATION
 - .. TECHNICAL CO-OPERATION AMONG DEVELOPING COUNTRIES
- TECHNICAL LITERATURE**
- .. COMPUTER SCIENCE LITERATURE
- TECHNOLOGY TRANSFER**
- .. DIFFUSION OF INNOVATIONS
- THEOLOGY**
- .. CATHOLIC THEOLOGY
 - .. ISLAMIC THEOLOGY
- THERAPEUTIC MATERIALS**
- .. MEDICINAL PLANTS
 - .. MINERAL WATERS
 - .. QUININE
- TOURIST FACILITIES**
- .. HOTELS
 - .. TRAILER CAMPS
 - .. YOUTH HOSTELS
- TRADE ASSOCIATIONS**
- .. CHAMBERS OF COMMERCE
- TRADE DOCUMENTS**
- .. ALIGNED TRADE DOCUMENTS
 - .. CUSTOMS DOCUMENTS
 - .. CERTIFICATES OF ORIGIN
 - .. PHYTSANITARY CERTIFICATES
 - .. EXPORT DOCUMENTS
 - .. IMPORT DOCUMENTS
- TRADE FACILITATION**
- .. TRADE DATA INTERCHANGE
- TRADE FINANCING**
- .. COMPENSATORY FINANCING
 - .. EXPORT FINANCING
 - .. EXPORT CREDITS
 - .. EXPORT SUBSIDIES
 - .. IMPORT CREDITS
- TRADE PREFERENCES**
- .. TARIFF PREFERENCES
 - .. GENERALIZED SYSTEM OF PREFERENCES
- TRADE PROMOTION**
- .. EXPORT PROMOTION
 - .. EXPORT INCENTIVES
 - .. EXPORT SUBSIDIES
- TRADE RESTRICTIONS**
- .. NONTARIFF BARRIERS
 - .. ANTIDUMPING DUTIES
 - .. COUNTERVAILING DUTIES
 - .. EXPORT RESTRICTIONS
 - .. EXPORT QUOTAS
 - .. IMPORT RESTRICTIONS
 - .. EXPORT RESTRAINTS
 - .. IMPORT CHARGES AND SURCHARGES
 - .. IMPORT DEPOSITS
 - .. IMPORT QUOTAS
 - .. TARIFFS
 - .. TRADE BOYCOTTS
- TRAINING PROGRAMMES**
- .. MINICOURSES
- TRANSIT**
- .. DUTY-FREE TRANSIT
- TRANSPORT**
- .. AIR TRANSPORT
 - .. COASTAL TRANSPORT
 - .. INLAND WATER TRANSPORT
 - .. MARITIME TRANSPORT
 - .. MULTIMODAL TRANSPORT
 - .. PASSENGER TRANSPORT
 - .. PNEUMATIC TUBE TRANSPORTATION
 - .. RAILWAY TRANSPORT
 - .. ROAD TRANSPORT
 - .. URBAN TRANSPORT
- TRANSPORT DOCUMENTS**
- .. BILLS OF LADING
- TRANSPORT FACILITATION**
- .. AIR TRANSPORT FACILITATION
 - .. MARITIME TRANSPORT FACILITATION
 - .. RAILWAY TRANSPORT FACILITATION
 - .. ROAD TRANSPORT FACILITATION
- TRANSPORT INFRASTRUCTURE**
- .. AIRPORTS
 - .. STOLPORTS
 - .. BRIDGES
 - .. CANALS
 - .. PIPELINES
 - .. PORTS
 - .. RAILWAYS
 - .. ROADS
 - .. HIGHWAYS
- TOLL ROADS**
- .. TERMINALS
 - .. CONTAINER TERMINALS
 - .. MARINE TERMINALS
 - .. TUNNELS
- TRAVEL**
- .. AIR TRAVEL
- TRAVEL DOCUMENTS**
- .. PASSPORTS
 - .. REFUGEE PASSPORTS
 - .. VISAS
- UNEMPLOYMENT**
- .. SEASONAL UNEMPLOYMENT
 - .. TECHNOLOGICAL UNEMPLOYMENT
 - .. UNDEREMPLOYMENT
- UNIVERSITIES AND COLLEGES**
- .. JUNIOR COLLEGES
- URBAN PLANNING**
- .. PHYSICAL PLANNING
 - .. URBAN RENEWAL
 - .. SLUM CLEARANCE
 - .. ZONING
- USE STUDIES**
- .. LIBRARY USE STUDIES
- VEHICLES**
- .. AIR CUSHION VEHICLES
 - .. AIRCRAFT
 - .. BOMBERS
 - .. FIGHTER PLANES
 - .. HELICOPTERS
 - .. JET AIRCRAFT
 - .. SUPERSONIC AIRCRAFT
 - .. ELECTRIC VEHICLES
 - .. ELECTRIC LOCOMOTIVES
 - .. LOCOMOTIVES
 - .. DIESEL LOCOMOTIVES
 - .. ELECTRIC LOCOMOTIVES
 - .. MOTOR VEHICLES
 - .. AUTOMOBILES
 - .. DIPLOMATIC VEHICLES
 - .. BUSES
 - .. COMMERCIAL VEHICLES
 - .. MOTORCYCLES
 - .. TRACTORS
 - .. TRUCKS
 - .. PASSENGER VEHICLES
 - .. PASSENGER SHIPS
 - .. SPECIAL TRADE PASSENGER SHIPS
 - .. SHIPS
 - .. CARGO SHIPS
 - .. CONTAINER SHIPS
 - .. TANKERS
 - .. TRAMP STEAMERS
 - .. HOSPITAL SHIPS
 - .. HYDROFOIL CRAFT
 - .. NUCLEAR SHIPS
 - .. PASSENGER SHIPS
 - .. SPECIAL TRADE PASSENGER SHIPS
 - .. SUBMARINES
 - .. NUCLEAR SUBMARINES
 - .. WARSHIPS
 - .. NUCLEAR SUBMARINES
- VERIFICATION MEASURES**
- .. SATELLITE MONITORING
 - .. SEISMIC MONITORING
- VETERANS**
- .. DISABLED VETERANS
- VIOLENCE**
- .. DOMESTIC VIOLENCE
 - .. POLITICAL VIOLENCE
 - .. RIOTS
 - .. TERRORISM
 - .. HIJACKING OF AIRCRAFT
- VOCATIONAL TRAINING**
- .. AGRICULTURAL TRAINING
 - .. FORESTRY TRAINING
 - .. CO-OPERATIVE EDUCATION
 - .. RETRAINING

- EDUCATION (cont.)**
- 11.01.00 EDUCATION
 - 11.04.00 EDUCATION FOR PEACE
 - 18.01.00 EDUCATION GRANT
 - 11.03.00 **ELEMENTARY EDUCATION**
USE:
PRIMARY EDUCATION
 - 11.03.00 ENGINEERING EDUCATION
 - 11.03.00 ENVIRONMENTAL EDUCATION
 - 14.05.03 FAMILY LIFE EDUCATION
 - 08.02.00 FAMILY PLANNING EDUCATION
 - 11.03.00 **FUNDAMENTAL EDUCATION**
USE:
BASIC EDUCATION
 - 10.02.00 HEALTH EDUCATION
 - 11.03.00 HIGHER EDUCATION
 - 14.02.02 HUMAN RIGHTS EDUCATION
 - 10.02.00 **HYGIENE EDUCATION**
USE:
HEALTH EDUCATION
 - 11.03.00 INTERNATIONAL EDUCATION
 - 13.01.01 **LANGUAGE EDUCATION**
USE:
LANGUAGE INSTRUCTION
 - 15.05.00 LIBRARY EDUCATION
 - 11.04.00 **LIFE-LONG EDUCATION**
USE:
CONTINUING EDUCATION
 - 05.02.00 **MANAGEMENT EDUCATION**
USE:
MANAGEMENT DEVELOPMENT
 - 10.06.01 MEDICAL EDUCATION
 - 01.04.00 MILITARY EDUCATION
 - 11.04.00 MORAL EDUCATION
 - 11.04.00 NONFORMAL EDUCATION
 - 10.01.00 NUTRITION EDUCATION
 - 11.03.00 PARENT EDUCATION
 - 11.03.00 PHYSICAL EDUCATION
 - 08.02.00 POPULATION EDUCATION
 - 11.03.00 PRESCHOOL EDUCATION
 - 11.03.00 PRIMARY EDUCATION
 - 11.03.00 PROFESSIONAL EDUCATION
 - 11.01.00 **PSYCHOLOGY OF EDUCATION**
USE:
EDUCATIONAL PSYCHOLOGY
 - 15.03.00 RELIGIOUS EDUCATION
 - 14.02.02 RIGHT TO EDUCATION
 - 11.03.00 RURAL EDUCATION
 - 11.03.00 SCIENCE EDUCATION
 - 11.03.00 SECONDARY EDUCATION
 - 08.02.00 SEX EDUCATION
 - 11.03.00 SOCIAL WORK EDUCATION
 - 11.03.00 SPECIAL EDUCATION
 - 11.03.00 TECHNICAL EDUCATION
 - 09.02.00 TENANT EDUCATION
 - 11.03.00 VOCATIONAL EDUCATION
 - 12.04.00 WORKERS' EDUCATION
- EDUCATIONAL**
- 11.01.00 EDUCATIONAL ADMINISTRATION
 - 11.01.00 EDUCATIONAL ANTHROPOLOGY
 - 11.01.00 EDUCATIONAL ASPECTS
 - 11.01.00 EDUCATIONAL ASSISTANCE
 - 11.03.00 **EDUCATIONAL ATTENDANCE**
USE:
SCHOOL ATTENDANCE
 - 11.01.00 EDUCATIONAL COSTS
 - 11.01.00 EDUCATIONAL DEVELOPMENT
 - 11.01.00 EDUCATIONAL EXCHANGES
 - 11.02.00 EDUCATIONAL FACILITIES
 - 11.02.00 EDUCATIONAL FILMS
 - 11.01.00 EDUCATIONAL FINANCING
 - 11.01.00 EDUCATIONAL FORECASTING
 - 11.03.00 EDUCATIONAL GUIDANCE
 - 11.01.00 EDUCATIONAL INFORMATION
 - 11.01.00 EDUCATIONAL INNOVATIONS
 - 11.01.00 EDUCATIONAL INSTITUTIONS
 - 12.03.00 EDUCATIONAL LEAVE
 - 11.01.00 EDUCATIONAL LEGISLATION
 - 11.01.00 EDUCATIONAL MANAGEMENT
 - 11.05.00 **EDUCATIONAL PERSONNEL**
USE:
TEACHING PERSONNEL
 - 11.01.00 EDUCATIONAL PHILOSOPHY
 - 11.01.00 EDUCATIONAL PLANNING
 - 11.01.00 EDUCATIONAL POLICY
 - 11.01.00 EDUCATIONAL PSYCHOLOGY
 - 11.02.00 EDUCATIONAL RADIO
 - 11.01.00 EDUCATIONAL REFORM
 - 11.01.00 EDUCATIONAL RESEARCH
- 11.01.00 EDUCATIONAL SOCIOLOGY
 - 11.01.00 EDUCATIONAL STATISTICS
 - 11.01.00 EDUCATIONAL SURVEYS
 - 11.03.00 EDUCATIONAL SYSTEMS
 - 11.02.00 EDUCATIONAL TECHNOLOGY
 - 11.02.00 EDUCATIONAL TELEVISION
- EDUCATIONAL,**
- 11.02.00 EDUCATIONAL, SCIENTIFIC, CULTURAL MATERIALS
- EFFECT**
- 06.02.00 **GROUND EFFECT MACHINES**
USE:
AIR CUSHION VEHICLES
- EFFECTIVENESS**
- 02.01.00 COST EFFECTIVENESS
- EFFECTS**
- 10.04.00 RADIATION EFFECTS
 - 16.10.00 REDUCED GRAVITY EFFECTS
- EFFICIENCY**
- 03.10.00 ENERGY EFFICIENCY
 - 03.10.00 **FUEL EFFICIENCY**
USE:
ENERGY EFFICIENCY
 - 05.02.00 INDUSTRIAL EFFICIENCY
- EGGS**
- 04.03.00 EGGS
- EGYPT**
- 17.01.00 EGYPT
- EGYPTIANS**
- 08.03.02 EGYPTIANS
- ELDERLY**
- 14.03.00 **ELDERLY**
USE:
AGING PERSONS
- ELECTION**
- 01.01.00 ELECTION DISTRICTS
 - 01.07.00 ELECTION LAW
 - 18.07.01 ELECTION OF MEMBERS
 - 18.07.01 ELECTION OF OFFICERS
- ELECTIONS**
- 01.01.00 ELECTIONS
- ELECTRIC**
- 05.06.00 ELECTRIC CABLES
 - 16.09.00 ELECTRIC FIELDS
 - 06.03.00 ELECTRIC LOCOMOTIVES
 - 05.06.00 ELECTRIC MOTORS
 - 03.10.00 ELECTRIC POWER
 - 05.06.00 ELECTRIC POWER DISTRIBUTION
 - 05.06.00 ELECTRIC POWER PLANTS
 - 05.06.00 ELECTRIC POWER RATES
 - 05.06.00 **ELECTRIC POWER TRANSMISSION**
USE:
ELECTRIC POWER DISTRIBUTION
 - 05.06.00 **ELECTRIC UTILITY RATES**
USE:
ELECTRIC POWER RATES
 - 06.03.00 ELECTRIC VEHICLES
 - 05.06.00 HYDROTHERMAL ELECTRIC POWER SYSTEMS
- ELECTRICAL**
- 05.06.00 ELECTRICAL ENGINEERING
 - 05.03.00 ELECTRICAL INDUSTRY
 - 05.06.00 ELECTRICAL INSTALLATIONS
 - 05.03.00 ELECTRICAL MACHINERY
- ELECTRICITY**
- 16.09.00 ELECTRICITY
 - 05.06.00 **ELECTRICITY RATES**
USE:
ELECTRIC POWER RATES
- ELECTRIFICATION**
- 05.06.00 ELECTRIFICATION
 - 04.01.01 RURAL ELECTRIFICATION
- ELECTROACOUSTICS**
- 05.06.00 ELECTROACOUSTICS
- ELECTROCHEMISTRY**
- 16.09.00 ELECTROCHEMISTRY
- ELECTRON**
- 16.01.00 ELECTRON MICROSCOPY
- ELECTRONIC**
- 14.04.02 ELECTRONIC BEHAVIOUR CONTROL
 - 05.06.00 ELECTRONIC CIRCUITS
 - 16.02.00 **ELECTRONIC DATA PROCESSING**
USE:
DATA PROCESSING
 - 05.06.00 ELECTRONIC EQUIPMENT
 - 02.11.00 ELECTRONIC FUNDS TRANSFER
 - 05.06.00 ELECTRONIC INSTRUMENTS
 - 14.04.02 ELECTRONIC SURVEILLANCE
- ELECTRONICS**
- 05.06.00 ELECTRONICS
 - 05.06.00 ELECTRONICS INDUSTRY
- ELECTRONS**
- 16.09.00 ELECTRONS
- ELEMENTARY**
- 11.03.00 **ELEMENTARY EDUCATION**
USE:
PRIMARY EDUCATION
 - 16.09.00 ELEMENTARY PARTICLES
- ELEMENTS**
- 18.06.00 **PROJECT ELEMENTS**
USE:
PROJECT ACTIVITIES
- ELITE**
- 14.05.03 ELITE
- ELLICE**
- 17.03.02 **ELLICE ISLANDS**
USE:
TUVALU
- EMBARGO**
- 01.03.00 ARMS EMBARGO
 - 01.07.04 EMBARGO
 - 01.03.00 GRAIN EMBARGO
 - 01.03.00 OIL EMBARGO
- EMBASSIES**
- 01.02.00 EMBASSIES
- EMBEZZLEMENT**
- 14.04.02 EMBEZZLEMENT
- EMBLEMS**
- 14.06.00 EMBLEMS
 - 14.06.00 NATIONAL EMBLEMS
- EMBRYOLOGY**
- 10.06.02 EMBRYOLOGY
- EMERGENCIES**
- 10.02.00 **MEDICAL CARE IN EMERGENCIES**
USE:
EMERGENCY MEDICAL SERVICES
- EMERGENCY**
- 01.01.00 EMERGENCY LEGISLATION
 - 10.02.00 EMERGENCY MEDICAL SERVICES
 - 01.01.00 EMERGENCY POWERS
 - 13.02.00 EMERGENCY RELIEF
 - 13.02.00 EMERGENCY SHELTER
- EMIGRATION**
- 08.01.00 EMIGRATION
 - 01.07.00 **EMIGRATION AND IMMIGRATION LAW**
USE:
EMIGRATION LAW
IMMIGRATION LAW
 - 01.07.00 IMMIGRATION LAW
- EMINENT**
- 09.01.00 EMINENT DOMAIN
- EMIRATES**
- 17.05.00 UNITED ARAB EMIRATES
- EMISSIONS**
- 06.02.00 AIRCRAFT ENGINE EMISSIONS

E417

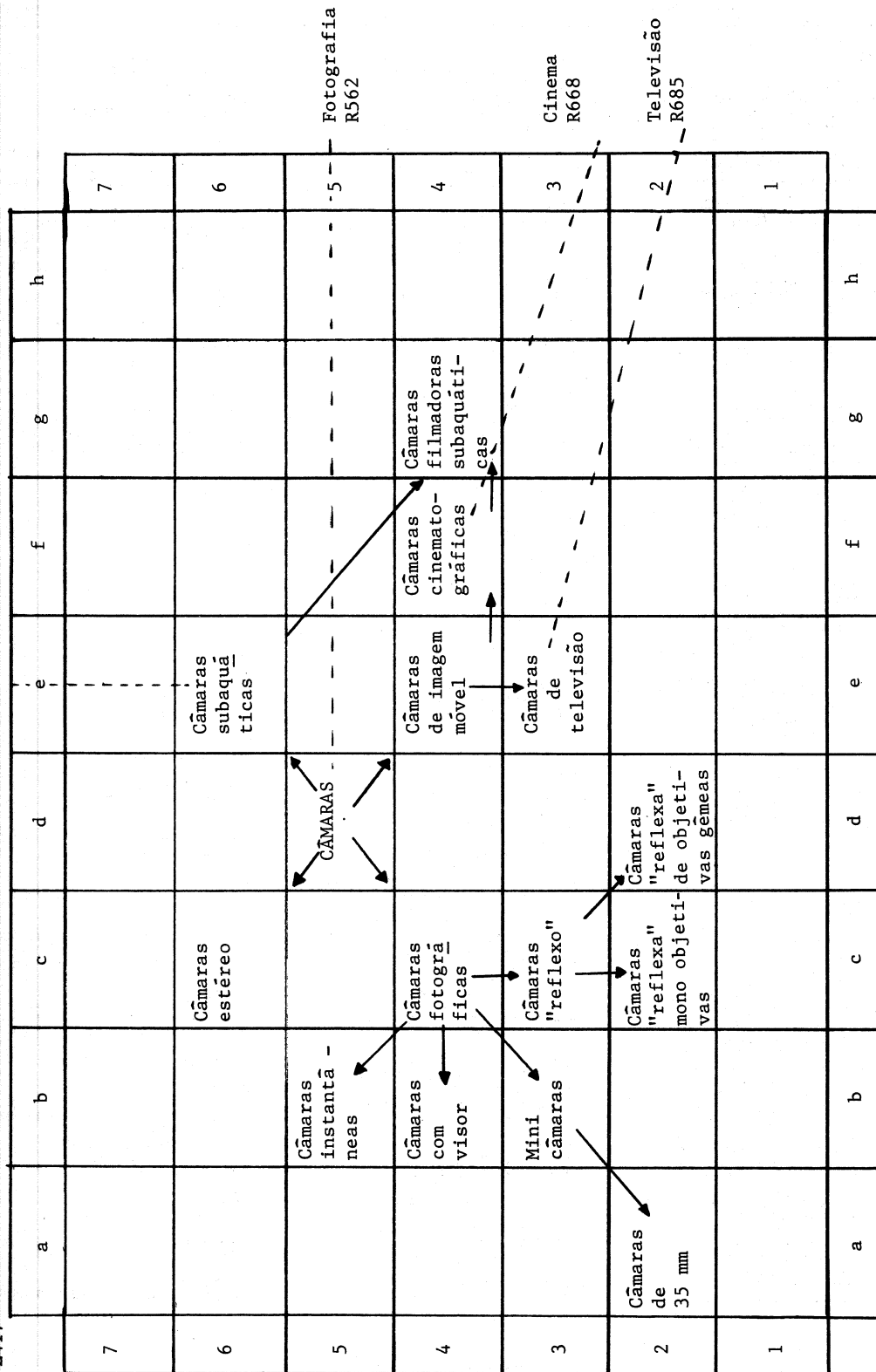


Ilustração 51

35mm CAMERAS E417.a2
BT: Miniature cameras

CAMERAS E417.d5
RT: Photography R562

CINE CAMERAS E417.f4
BT: Moving picture cameras
NT: Underwater cine cameras
RT: Cinema R668

CINEMA R668.d5
RT: Cine cameras E417

DIVING T473.g5
RT: Underwater cameras E417

INSTANT PICTURE CAMERAS E417.b5
SN: Cameras which produce a
finished print directly
BT: Still cameras

Land cameras *USE* VIEW CAMERAS

MINIATURE CAMERAS E417.b3
BT: Still cameras
NT: 35mm cameras

MOVING PICTURE CAMERAS E417.e4
BT: Cameras
NT: Cine cameras
Television cameras

PHOTOGRAPHY R562.d5
RT: Cameras E417

REFLEX CAMERAS E417.c3
BT: Still cameras
NT: Single lens reflex cameras
Twin lens reflex cameras

SINGLE LENS REFLEX CAMERAS E417.c2
BT: Reflex cameras

STEREO CAMERAS E417.c6
BT: Cameras

STILL CAMERAS E417.c4
BT: Cameras
NT: Instant picture cameras
Miniature cameras
Reflex cameras
View cameras

TELEVISION R685.d5
RT: Television cameras E417

TELEVISION CAMERAS E417.e3
BT: Moving picture cameras
RT: Television R685

TWIN LENS REFLEX CAMERAS E417.d2
BT: Reflex cameras

UNDERWATER CAMERAS E417.e6
BT: Cameras
NT: Underwater cine cameras
RT: Diving T473

UNDERWATER CINE CAMERAS E417.g4
BT: Cine cameras
Underwater cameras

VIEW CAMERAS E417.b4
SN: Cameras with through-the-lens
focusing and a range of
movements of the lens plane
relative to the film plane
UF: Land cameras
BT: Still cameras

Ilustração 52

Entrada do descritor superior

- A. Descritor _____.
- B. Etiqueta numérica....5757 P&D.
 (Abreviatura de PESQUISA E DESENVOLVIMENTO EXPERIMENTAL)
- C. Usado por _____ up Pesquisa e Desenvolvimento _____
- D. Indica uso de mais de um descritor up pd & e _____
- E. Visto em _____ ver processos inovadores _____
- F.

{	Termos específicos	TE	DESENVOLVIMENTO EXPERIMENTAL	}	
	-- nível 1	TE1	PESQUISA		
	-- nível 2	TE2	PESQUISA APLICADA		
	-- nível 3	TE3	PESQUISA DE PROCESSOS		
			PESQUISA DE PRODUTOS		
		PESQUISA FUNDAMENTAL			
		tr	CRIATIVIDADE		N. Termos relacionados
		tr	ENGENHARIA		
- J. Nota Explicativa
- K. Referências USE (ver abaixo)
- L. Referências Ver/Ou (abaixo)
- M. Termos específicos

Entrada do descritor intermediário

435 PESQUISA APLICADA

Termos genéricos

- G.

-- nível 1	tg1	PESQUISA	}	O. Termos genéricos
-- nível 2	tg2	P & D		
	TE1	PESQUISA DE PROCESSOS		
	TE1	PESQUISA DE PRODUTOS		
	tr	DESENVOLVIMENTO EXPERIMENTAL		
	tr	PESQUISA FUNDAMENTAL		

Entradas de não-descritores

435 PESQUISA APLICADA

- H. Referências VER/OU (não descritor) _____ Processos inovadores ver P&D (5757) _____
- ou TRANSFERÊNCIA DE TECNOLOGIA (6978) _____
- I. Referências USE (não-descritores)

}	pd&e	USE	ENGENHARIA (2190)	}	R. Todos os descritores para serem usados no lugar de "pd&e".
	E	P & D (5757)			
	USE	Pesquisa e desenvolvimento P & D (5757)			
- P. A etiqueta numérica de "P & D"
- Q. Só um descritor deve ser usado no lugar de "processos inovadores"
- S. Descritor único para ser usado no lugar de "pesquisa e desenvolvimento".

Ilustração 53

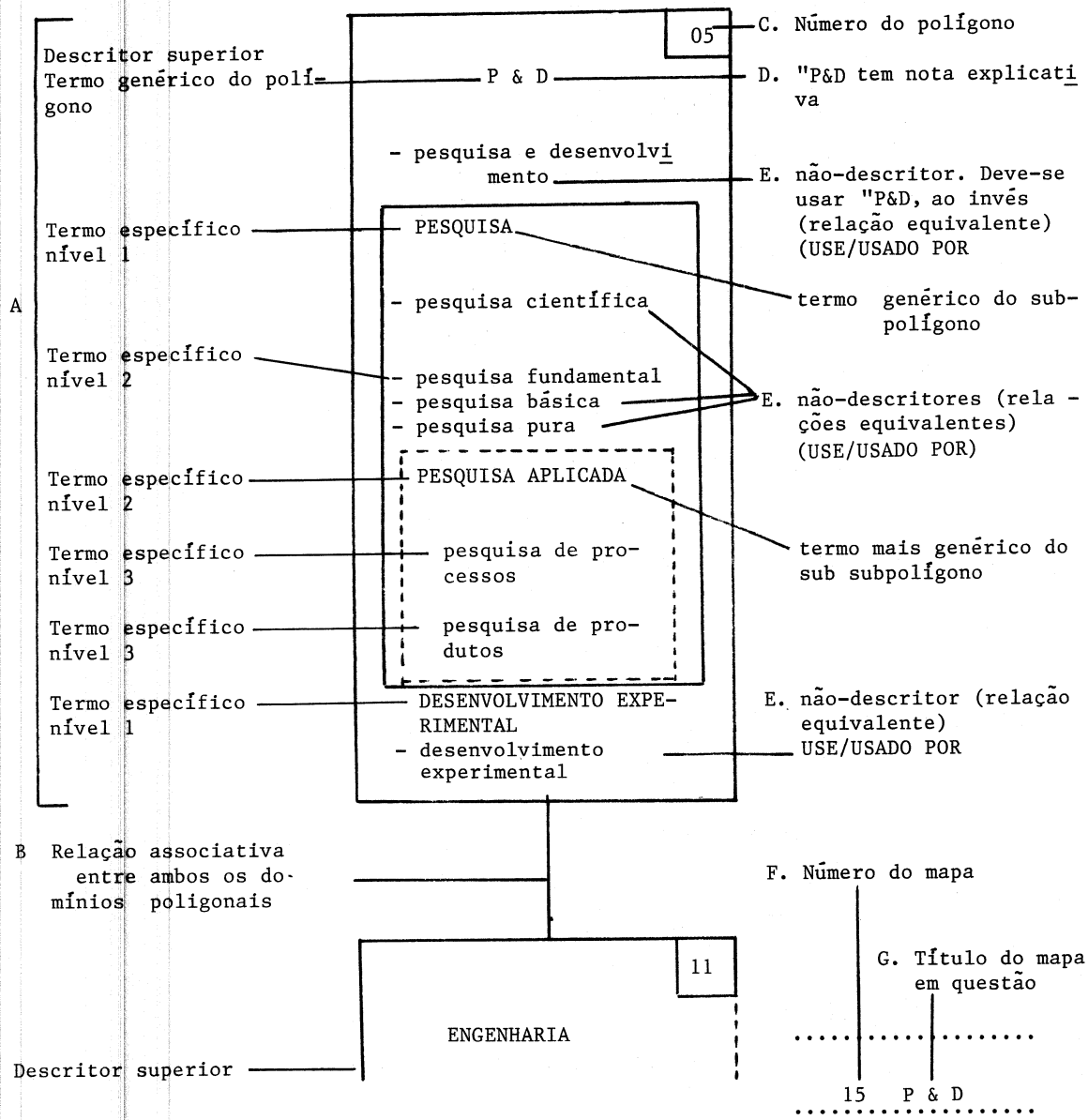


Ilustração 54

Apresentação facetada

La	Ácidos
Lac	Ácidos inorgânicos
Lacd	Ácido fosfórico
Lace	Ácido hidrolórico
Lach	Ácido hidroluórico
Laci	Ácido nítrico
Lack	Ácido nitroso
Lacl	Ácido perclórico
Lacq	Ácido fosfórico
Lacr	Ácido sulfúrico
Lad	Ácidos orgânicos

Apresentação do tesouro

Ácido sulfúrico		Lacr
TG(A)	Poluentes atmosféricos	Ppb
TR	Chuva ácida	Pps

Ácidos		La
TR	Substâncias corrosivas	Ps

Ilustração 55

R

Sociology

R70/84
R75/78

Social Problems (cont.)

Crime

Administration of justice M33/39*Criminal law M52***Law enforcement L56/59*

R75.10

Criminology

R76

Offences

UF Crimes

Addiction R73.50/90*Civil disturbances L73***Subversive activities L73.10***Terrorism L73.65***War crimes M98.50**(By crimes against property)*

R76.10

Theft

UF Burglary

Larceny

Robbery

R76.15

Arson

R76.20

Vandalism

UF Malicious damage

R76.25

Fraud

UF Embezzlement

Forgery

R76.30

Blackmail

R76.35

Hi-jacking

(By crimes against the person)

R76.40

Homicide

UF Murder

R76.40.10

Assassination

R76.45

Violence

UF Assault

Mugging

**Cruelty V15.40.20*

R76.45.10

Domestic violence

UF Violence in the home

Battering

**Ill-treated children R84.60.10*

R76.48

Kidnapping

UF Abduction

R76.50

Hostage-holding

R76.55

Sexual offences

UF Prostitution

White slave trade

Rape

R76.55.40

Pornography

(By crimes against public order)

R76.60

Disorderly conduct

UF Hooliganism

R76.68

Lynching

(By crimes against the state)

R76.70

Political offences

Subversive activities L73.10*Terrorism L73.65*

R76.70.10

Espionage

(By offence against justice)

R76.75

Perjury

R76.78

Corruption

UF Bribery

**Political corruption L76.50*

R76.90

(By organization)
Transnational crime

R76.90.40

Extradition

UF (Used For) = UP (Usado por)

Ilustração 56

Craftsmen (cont.)

RT Artists
Handicrafts

Creative role R59.80

BT Role
TT Social structure
RT Creativity

Creative thinking P54.20

BT Thinking
TT Higher mental processes
RT Creativity

Creative writing J46.03 *J60.25 *X60 *Y45.30

SN Students literary composition
UF Student composition
NT Oral composition
BT Literary composition
Study
Writing
TT Communication process
Study
RT Language education
Mother tongue instruction
Study

Creativity P52.10

UF Originality
RT Artistic creation
Creative role
Creative thinking
Cultural creation
Geniuses
Imagination
Visualization

Creches

USE Day Nurseries

Credit N88.15

NT Loans
BT Financing
TT Economics
RT Debts
Loans
Mortgages
Subsidies

Credit courses J40.19.13

SN Course work in higher education measured by a credit system whereby credit hours are assigned to a course defined by the number of hours per week in class and the number of weeks in the session.
BT Educational courses
TT Educational courses

Creep C27.30.40C

BT Plastic deformation
TT Damage
Faults
Physics
Properties
RT Fracture
Plasticity

Cremation

USE Crematoria

Crematoria E62.30

UF Cremation
RT Cemeteries
Disposal of the dead

Cretaceous period

USE Mesozoic period

Crime R75/78 *L57 *M52

NT Offences
BT Social problems
TT Social problems
RT Administration of justice
Crime victims
Criminal law
Criminals
Criminology
Delinquency
Law enforcement
Police
Punishment

Crime victims R84.30

UF Ill-treated persons
Victims of crime
Violence victims
NT Ill-treated children
BT Socially disadvantaged
TT People
Social problems
RT Crime
Cruelty
Violence

Crimes

USE Offences

Criminal courts M34.50

BT Courts
TT Courts
RT Criminal law

Criminal investigation L59.50/80

NT Evidence gathering
BT Police activities
TT Police activities
RT Detectives
Forensic medicine

Criminal law M52 *R75

BT Public law
TT Public law
RT Crime
Criminal courts
Criminology
Police

Criminals R77.30

UF Delinquents
Offenders
NT Gangs
War criminals
RT Crime
Criminology
Delinquency
Juvenile delinquency
Maladjusted
Offences
Prisoners

Criminology R75.10

RT Crime

BT(Broader Term) = TG(Termo Genérico)
NT(Narrower Term) = TE(Termo Específico)
RT(Related Term) = TR(Termo Relacionado)
SN(Scope Note) = NE(Nota Explicativa)
TT(Top Term) = TS(Termo Superior)
UF(Used For) = UP(Usado por)

N Mechanical engineering

NI/NL Fluid engineering (continued)

NJ/NL Fluid equipment (continued)

NJQ/NJT Valves (continued)

(By component)

NJS.S Valve components

NJS.SB Valve bodies

(By part)

NJS.SBJ Outlet connections
= Gas cylinder connections
* - Gas cylinders NIO.G

[For others synthesize, for example]

NJS.SBS **Valve bosses
- Valve bodies
+ Bosses

NJS.SJ Valve yokes

NJS.SO Valve plugs
= Plugs (valves)

NJS.SS Tap washers
* - Draw-off taps NJR.D

NJV/NJX Pipes

- = Tubes (pipes)
- * > Delivery pipes (firefighting) GON.D
- * > Drainpipes RMF.N
- * > Rainwater pipes RJJ.J
- * - Bending springs PSS
- * - Boiler tubes NEN
- * - Bursting discs GVQ
- * - Conduits (hydraulic) RCM.C
- * - Flues ROC.E
- * - Heat engineering NB/NE
- * - Hot-water supply systems RMB.M
- * - Pressure vessels NIN/NIP
- * - Rising mains RMB.FF
- * - Small-bore central-heating systems RKK.KR
- * - Sewerage GKG
- * - Space-heating systems RKJ/RKO
- * - Threads NWC/NWD
- * - Traps (drainage) RMF.P
- * - Ventilation equipment RKR.T

(By property)

NJV.B Rigid pipes

NJV.C Flexible pipes

= Hosepipes

= Hoses

* > Fire hoses GON.B

* - Hose connectors NKB.G

NJV.CC Flexible tubing

= Tubing (flexible)

(By application)

NJV.CP Suction hoses

(By storage device)

NJV.CV Hose reels

* > Fire hose reels GON.BN

Ilustração 58

- > Pipe tees
- * - Manifolds NLL.F
- Pipe plugs** NKD.N
 - Pipe couplings
 - * - Stoppers NSX.F
- Pipe reducers**
 - Reducing couplings NKD.K
- Pipe supports** NKH
 - = Pipe anchors
 - = Pipe brackets
 - = Pipe hangers
 - = Supports (pipes)
 - < Pipe fittings
 - Pipe clips
- Pipe systems**
 - Pipework systems NLF
- Pipe tees** NKF.F
 - = Tees (pipes)
 - < Pipe junctions
- Pipe threads**
 - Threads NWC/NWD
- Pipe tobacco** IYB
 - < Tobacco
 - * - Pipes (smokers) XUD
- Pipe wrenches** PTB.P
 - < Wrenches
- Piped services (laboratory)**
 - Laboratory piped services BPE
- Pipelaying equipment** RUJ
 - < Construction equipment
- Pipeline transportation** QBC.Q
 - < Land transport
 - * - Pipelines NLC
- Pipelines** NLC
 - [Systems of pipes]
 - = Fluid pipelines
 - < Fluid equipment
 - > Asbestos cement pipelines
 - > Cast-iron pipelines
 - > Gas pipelines
 - > Oil pipelines
 - > Plastic pipelines
 - Spigot-and-socket joints
 - * > Water mains GHS.S
 - * - District heating RKO.K
 - * - Drilling rigs STG
 - * - Flame traps GOJ.J
 - * - Pipeline transportation QBC.Q
- Pipelines**
 - + Water
 - =** Water pipelines NLC.R

- Piperaceae** EWK.W
 - < Dicotyledones
- Piperazine** DRR.R
 - = Diethylenediamine
 - < Heterocyclic compounds
- Piperidine** DRR.N
 - < Heterocyclic compounds
- Pipes** NJV/NJX
 - = Tubes (pipes)
 - < Fluid equipment
 - > Casing pipes
 - > Flexible pipes
 - > Gas pipes
 - > Goosenecks
 - > Insulation pipes
 - > Overflow pipes
 - > Pressure pipes
 - > Rigid pipes
 - > Seamed pipes
 - > Seamless pipes
 - > Taper pipes
 - > Vent pipes
 - Ducting
 - Pipe coatings
 - * > Delivery pipes (firefighting) GON.D
 - * > Drainpipes RMF.N
 - * > Rainwater pipes RJJ.J
 - * - Bending springs PSS
 - * - Boiler tubes NEN
 - * - Bursting discs GVQ
 - * - Conduits (hydraulic) RCM.C
 - * - Flues ROC.E
 - * - Heat engineering NB/NE
 - * - Hot-water supply systems RMB.M
 - * - Pressure vessels NIN/NIP
 - * - Rising mains RMB.FF
 - * - Sewerage GKG
 - * - Small-bore central-heating systems RKK.KR
 - * - Space-heating systems RKJ/RKO
 - * - Threads NWC/NWD
 - * - Traps (drainage) RMF.P
 - * - Ventilation equipment RKR.T
- Pipes**
 - + Asbestos cement
 - =** Asbestos cement pipes NJW.S
- Pipes**
 - + Ceramics
 - =** Ceramic pipes VVI
- Pipes**
 - + Plastics
 - =** Plastic pipes NJW.R
- Pipes**
 - + Water
 - =** Water pipes NJV.L

Ilustração 59

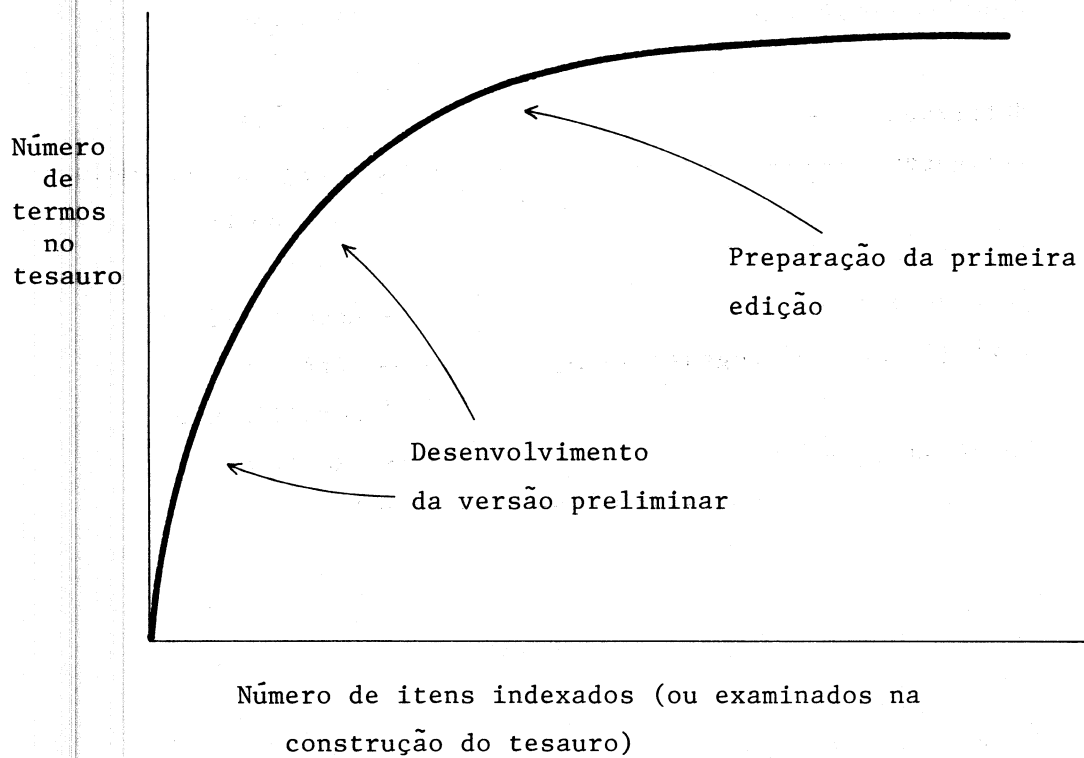


Ilustração 60

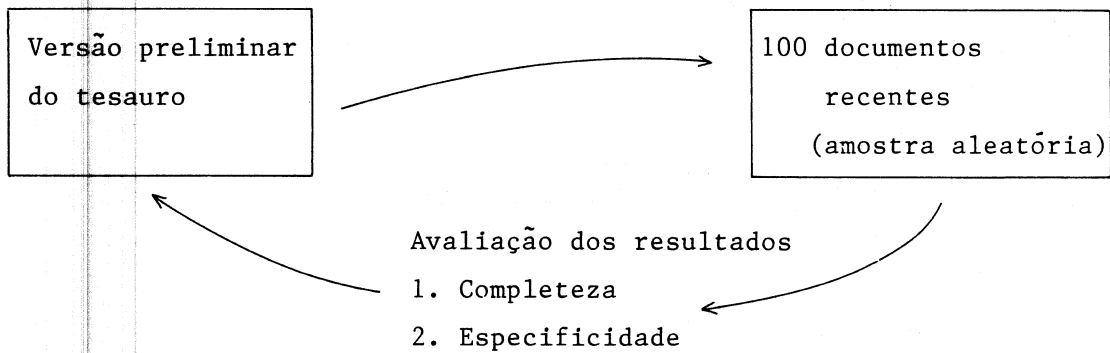


Ilustração 61

Análise conceitual	Tradução
Beterraba	Beterraba
Irrigação	Irrigação por meio de aspersores
Aspersores	
Água de esgoto	Água residual
Fertilizantes nitrogenados	Nitrogênio
	Fertilizantes
Sacarina	Sacarina
Israel	Israel

Ilustração 62

Número de conceitos	872
Número de termos traduzíveis	860
Percentagem de termos traduzidos	98%
Número de termos traduzíveis no nível de especificidade desejado	835
Percentagem de termos traduzíveis no nível de especificidade desejado	96%

Ilustração 63

Atualização

1. Termos coletados na literatura
2. Termos coletados nas consultas

Ilustração 64

Função do computador

1. Coletar os termos
2. Expandir a estrutura do tesauro
3. Gerar tesauros impressos
4. Facilitar o trabalho de atualização e manutenção
5. Fazer a interface entre a indexação e as operações de busca
6. Fazer a interface com a produção de índices impressos
7. Proporcionar apresentação dos dados em terminal em linha

Ilustração 65

Arquivo-Mestre do Vocabulário

1. Verificar a consistência e aceitabilidade dos termos
2. Proporcionar mapeamento automático
3. Manter informações estatísticas sobre o uso dos termos
4. Proporcionar buscas em hierarquias completas
5. Otimizar estratégias de busca
6. Gerar referências cruzadas para índices impressos
7. Proporcionar apresentação dos dados em terminal em linha

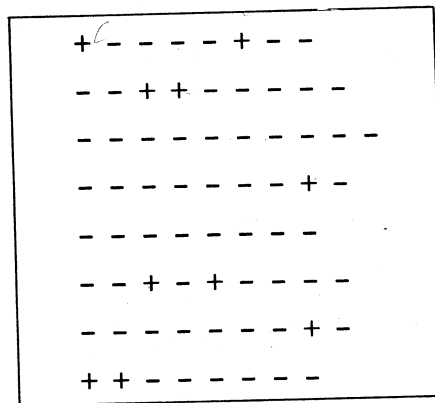
Ilustração 66

1. Liberdade de informação	8
2. Língua alemã	207
3. Língua árabe	82
4. Língua estrangeira	24
5. LINGUAGEM	
6. Linguagem de computação	43
7. Linguagem de indexação	58
8. Lingüística aplicada	109
9. Lingüística estrutural	124

Ilustração 67

1		Frutas cítricas
2	TG	Frutas
3	TE	Toranjias
4		Laranjas
5		Limas
6		Limões
7		Tangerinas
8	TR	Frutas cristalizadas
9		Sucos de frutas

Ilustração 68



Busca A: + + +

Busca B: + + + + + + + +

+ - - - - - - -

Ilustração 69

1. Consultas que representam de forma imperfeita as necessidades de informação
2. Fatores de indexação
3. Fatores da estratégia de busca
4. Fatores do vocabulário

Ilustração 70

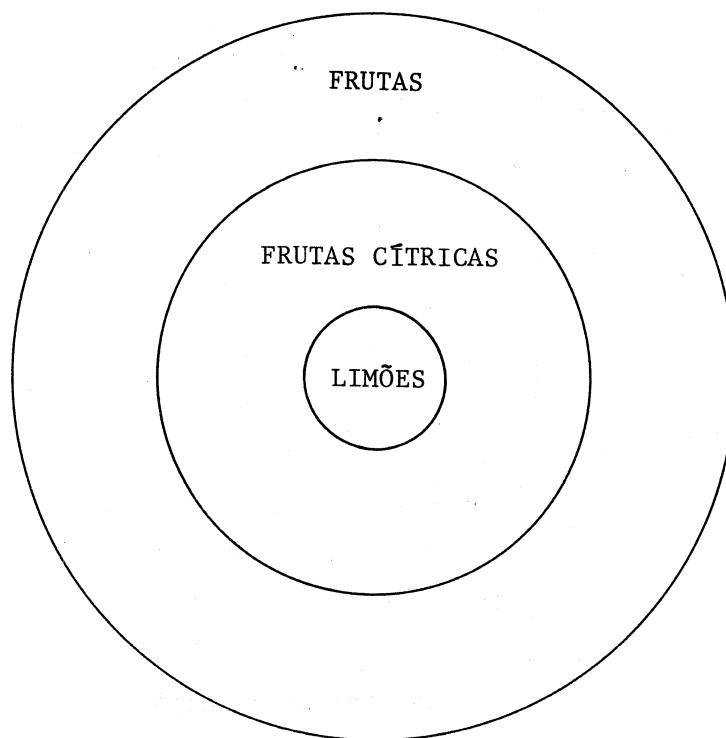


Ilustração 71

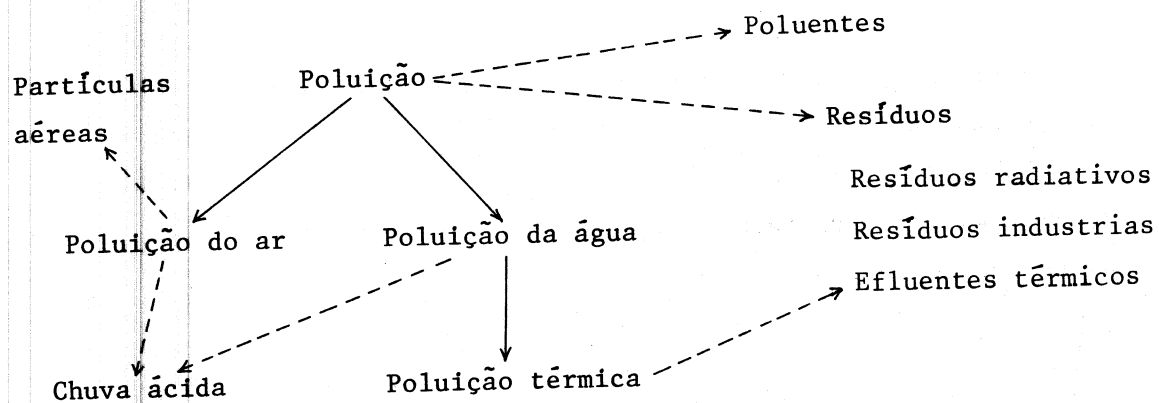


Ilustração 72

Brasil
 Exportações
 Importações
 Balança comercial
 Argentina
 México

Controle monetário
 Licença de exportação

Ilustração 73

Controles monetários - Brasil
 Brasil - Balança comercial
 Brasil - Exportações - Argentina
 Brasil - Exportações - México

Ilustração 74

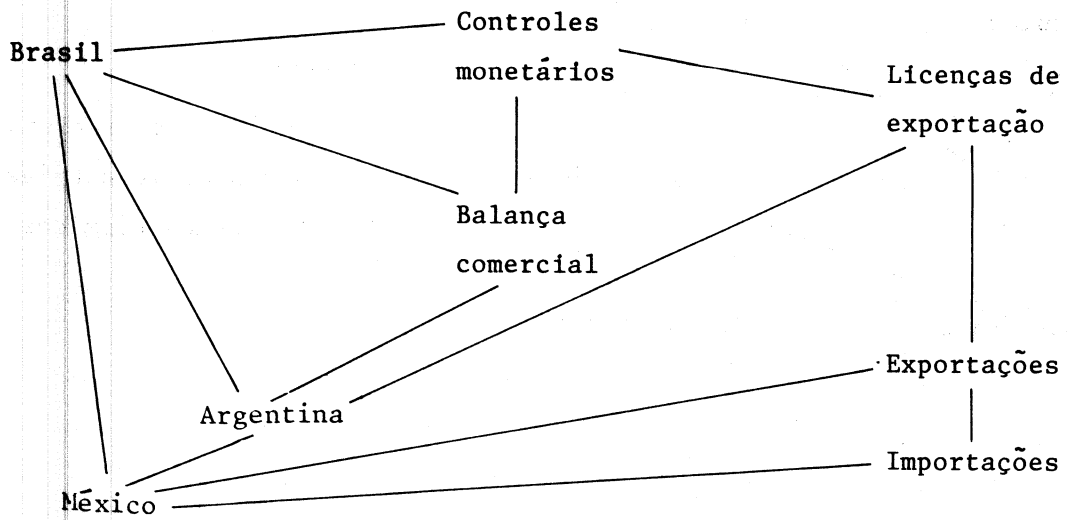


Ilustração 75

ARGENTINA e EXPORTAÇÕES e MÉXICO

BRASIL)
 EXPORTAÇÕES) 1
 ARGENTINA)
 IMPORTAÇÕES)

BRASIL)
 BALANÇA COMERCIAL) 3
 CONTROLES MONETÁRIOS)
 LICENÇAS DE EXPORTAÇÃO)

BRASIL)
 EXPORTAÇÕES) 2
 MÉXICO)
 IMPORTAÇÕES)

Ilustração 76

A = fonte, doador
B = alvo, recipiente

Brasil (A)
Exportação
Argentina (B)
Importação

Ilustração 77

Mulheres
Poesia

Mulheres na poesia

Computadores
Desenho

Desenho de computadores
Desenho de aeronaves
Desenho com ajuda de
Computador

Computadores/Desenho
Aeronaves/Desenho

Ilustração 78

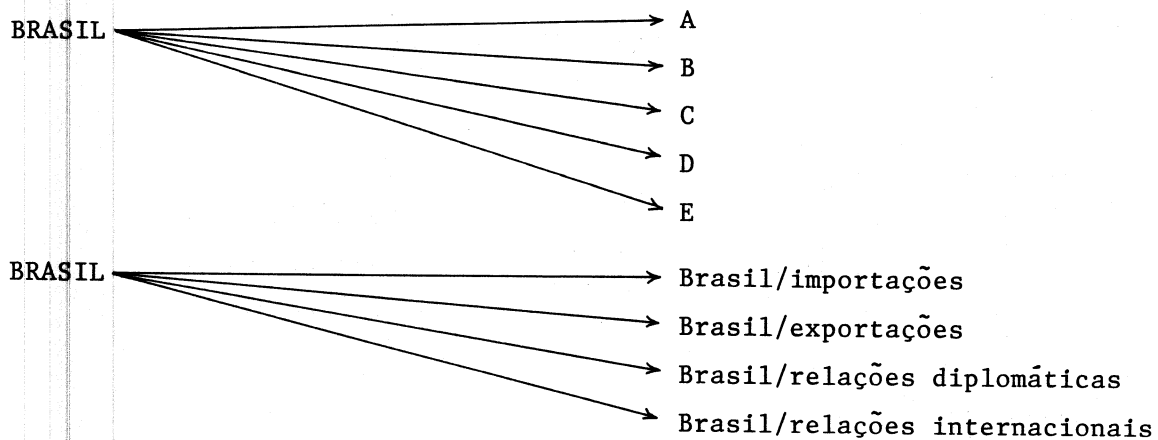


Ilustração 79

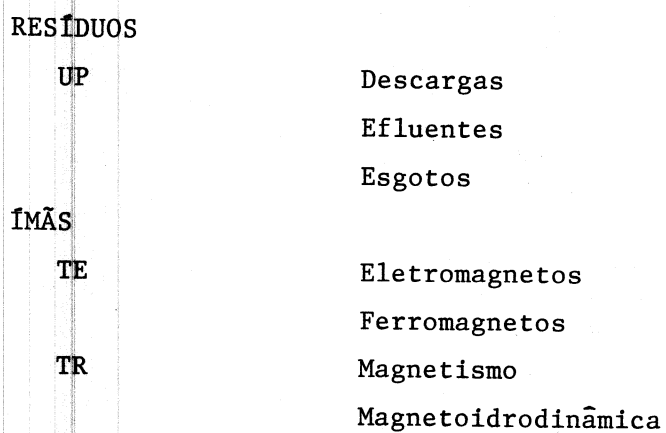


Ilustração 80

Fragmentos de palavras

Magnet...

Magnet

Magnético

Magnetos

Magneticamente

Magnetismo

Magnetoidrodinâmica

...magneto

Eletromagneto

Ferromagneto

...magnet...

Eletromagnetismo

Ferromagnetismo

Eletromagneticamente

Ilustração 81

Corrosão.....cobre

(10 palavras)

"Ligas de cobre"

Ilustração 82

TRABALHO
GUERRA

Piquetes
Piquetes

Ilustração 83

RESÍDUOS

RESÍDUOS*
EFLUENTES*
DESCARGAS*
BORRA*

ANTIBIÓTICOS

ANTIBIÓTICO*
*MICINA
*ILINA
*BIÓTICO
*MIXINA
*CICLINA

Ilustração 84

BIBLIOGRAFIA

Normas

UNISIST guidelines for the establishment and development of monolingual thesauri. 2. ed. Paris, UNESCO, 1981. 64p. (PGI-81/WS/15).

UNISIST guidelines for the establishment and development of multilingual thesauri. Revised ed. Paris, UNESCO, 1980. 85p. (PGI-80/WS/12).

Textos gerais

AITCHISON, Jean. *Manual para construção de tesouros*. Trad. Helena Medeiros Pereira Braga, Rio de Janeiro, BNG; Brasilart, 1979. 142p.

LANCASTER, F. W. *Vocabulary control for information retrieval*. 2. ed. Arlington, VA., Information Resources Press, 1985.

SOERGEL, D. *Indexing languages and thesauri: construction and maintenance*. Los Angeles, Melville, 1974. 632p. (Mais uma útil obra de consulta do que um texto para ser lido sem interrupção).

Artigos selecionados de

(Documentos que tratam de vários aspectos da compilação e estrutura dos tesouros, que o autor julgou particularmente ricos em informação, ou úteis).

DYM, E. D. A new approach to the development of a technical thesaurus. *Proceedings of the American Documentation Institute* 4:126-31, 1967.

JONES, K. P. Problems associated with the use of compound words in thesauri with special reference to BS 5723: 1979. *Journal of Documentation*, 37(2):53-68, 1981.

MANDERSLLOT, W. G. B.; DOUGLAS, E. M. B.; SPICER, N.

Thesaurus control - the selection, grouping and cross-referencing of terms for inclusion in a coordinate index word list. *Journal of the American Society for Information Science*, 21(1):49-57, 1970.

PICFORD, A. G. A. An objective method for the generation of an information retrieval language. *Information Scientist*, 2(1):17-37, 1968.

WILLETTS, M. An investigation of the nature of the relation between terms in thesauri. *Journal of Documentation*, 31(3):158-84, 1975.