

Universidade Federal de Minas Gerais
Escola de Ciência da Informação

Criação de linguagens de indexação



Graduação em Biblioteconomia
Disciplina: Tratamento da Informação V

Apostila estruturada e digitada pela bolsista Alessandra Rodrigues da Silva.
Projeto de Iniciação à Docência – PID: para a área de Tratamento e Serviços da Informação.
Orientação Profª. Gercina Lima.

SUMÁRIO

<i>1^a Parte - Estabelecimento de categorias</i>	3
<i>2^a Parte - Arranjo sistemático das categorias</i>	8
<i>3^a Parte - A notação de um sistema de classificação</i>	12
<i>4^a Parte - Proposta para a área de zoologia</i>	14
<i>5^a Parte - Estabelecimento de categorias em zoologia</i>	16
<i>6^a Parte - Arranjo sistemático em zoologia</i>	24
<i>7^a Parte - Ordem de citação X ordem de arquivamento</i>	32
<i>8^a Parte - Notação em um sistema de classificação em zoologia</i>	33
<i>9^a Parte - Criação de Índice Alfabético</i>	39

Criação de linguagens de indexação

1ª Parte - Estabelecimento de categorias

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

RAMSDEN, M. J. **An introduction to index language construction**, a programmed text. London, C. Bingley, 1974.

Todas as linguagens consistem de dois elementos básicos: um vocabulário e uma estrutura gramatical (sintaxe). As pessoas que escrevem sobre recuperação da informação usam freqüentemente, o termo **linguagens naturais** para denominar a linguagem falada e escrita. É possível em indexação empregar a linguagem natural simplesmente como é falada ou usada nos documentos sem tentar, por exemplo, controlar sinônimos ou indicar os relacionamentos entre os termos. Um índice feito desta maneira chama-se índice de **linguagem natural**. Como alternativa ao índice de linguagem natural pode-se usar uma linguagem artificial às necessidades do sistema, isto é, uma **linguagem de indexação**. Esta linguagem refletirá um vocabulário controlado para o qual foram tomadas decisões cuidadosas sobre os termos a serem usados, o significado de cada um e os relacionamentos que apresentam. Estes dois processos antagônicos são conhecidos como “**indexação por conceitos**”.

Uma linguagem de indexação deve indicar os relacionamentos entre os termos no seu vocabulário. Por exemplo, se um sistema inclui os termos EDUCAÇÃO SECUNDÁRIA E ESCOLAS PROFISSIONALIZANTES, deve-se indicar a relação entre esses dois termos, pois um usuário procurando informação sobre educação secundária, pode encontrar documentos relevantes sob o termo escolas profissionalizantes que, usualmente, se ocupam também de educação secundária. Uma linguagem de indexação que tem um vocabulário controlado e que procura indicar as relações entre os termos no seu vocabulário é chamada **estruturada**.

Há várias espécies de linguagens de indexação estruturadas, como por exemplo: sistemas de classificação, listas de cabeçalhos de assunto, tesauro, etc.

Todas elas consistem de: a) um vocabulário,
b) uma sintaxe.

O vocabulário é a lista de termos usados no sistema. É sempre menor do que o vocabulário dos usuários e do próprio indexador devido ao controle de termos. Por causa disto, é necessário fornecer, além do vocabulário de indexação (termos usados no índice), um **vocabulário de abordagem**, consistindo de termos que não são usados no sistema para representar os conceitos, mas que podem ser usados pelos autores, indexadores, e especialmente por usuários ao consultar o índice. Por exemplo, o vocabulário de um sistema usa o termo TESTES DE INTELIGÊNCIA. Porém, o assunto pode ser procurado por:

TESTES MENTAIS	-	Termo sinônimo
TESTES PSICOLÓGICOS	-	Termo mais abrangente
TESTES DE APTIDÃO	-	Termo mais específico
DESENVOLVIMENTO MENTAL	-	Termo relacionado

Estes termos constituem o **vocabulário de abordagem**.

A sintaxe é o padrão de relações reconhecidas entre os termos usados no sistema, isto é, entre os termos do vocabulário de indexação. Este reconhecimento é produto de uma análise de assunto cuidadosa que é básica para todas as linguagens de indexação estruturadas.

O vocabulário de uma linguagem de indexação pode ser **verbal** ou **codificado**. Um sistema de classificação emprega um vocabulário codificado na forma de sua notação já uma lista de cabeçalho de assunto que emprega um vocabulário verbal a representação dos assuntos ocorre por palavras. Assim o assunto História da Grã-Bretanha será representado na CDD por 942 e, na lista de cabeçalho de assunto como Grã-Bretanha – História.

Independente do vocabulário escolhido o primeiro passo é selecionar os conceitos existentes no assunto a ser indexado e os termos que o representarão.

Há diversas maneiras de se escolher os termos. Pode-se estudar o assunto e sugerir termos relevantes ou consultar especialistas. Alternativamente, pode-se consultar sistemas de classificação ou tesouros existentes. Contudo, a fonte mais útil e confiável será a literatura do assunto que o sistema irá indexar. Deve-se portanto, examinar a literatura para estabelecer os termos usados por aqueles que escrevem e trabalham na área. Uma linguagem de indexação cujo vocabulário se baseia no estudo da literatura tem “garantia literária”.

O exame da literatura revela a terminologia básica do assunto. Deve-se, então, decidir quais termos adotar no vocabulário de indexação e quais incluir no vocabulário de abordagem. Esta decisão se torna mais pertinente após a análise de assunto devido as relações entre os termos serem visualizadas mais facilmente.

A sintaxe em uma linguagem de indexação é o reconhecimento de certas relações entre os termos da linguagem ou, mais precisamente, entre os conceitos representados por esses termos. Análise de assunto é o processo pelo qual se estabelecem as relações a serem reconhecidas.

O processo de agrupamento dos termos que compartilham uma relação essencial é primordial para a análise de assunto. É um processo familiar existente na própria linguagem de comunicação. Por exemplo, quando se usa a palavra criminosos, entende-se que se refere àquelas pessoas que compartilham a característica essencial comum de ter “agido contra a lei criminal”. Assaltantes, falsários, chantagistas e seqüestreadores são espécies de criminosos (isto é, pertencem ao mesmo grupo ou classe) porque possuem a mesma característica essencial. O uso da palavra é importante e deve ser levado em consideração cuidadosamente. Qualquer pessoa, objeto ou fenômeno possui diversas características, mas aquela que se seleciona para governar um processo de agrupamento será a essencial para a finalidade de classificação do assunto. Assim, por exemplo, ao agrupar pessoas com a finalidade de criminologia, pode-se ignorar o fato de serem brancas ou negras, altas ou baixas. Considera-se somente se ofenderam ou não a lei criminal. Este é um processo classificatório (os grupos são, de fato, classes) e pode, eventualmente levar a produção de um esquema de classificação. Este processo é básico para qualquer linguagem de indexação estruturada.

O processo de agrupamento procura reunir os termos em grandes categorias. Examina-se cada termo indagando qual a característica essencial que reflete. Por exemplo, quando se trata do assunto VESTUÁRIO o termo **algodão** reflete a característica essencial de **material**, isto é, **algodão** pertence à classe **material**. Em linguagem classificatória isto se chama divisão, pois os termos dentro do assunto ou classe básica estão divididos em categorias. As características que governam a divisão são chamadas **características de divisão**. Quando se tem dúvida sobre que características um termo reflete pode-se resolver a questão procurando outros termos relacionados identificando a característica essencial que compartilham entre si.

Algumas categorias podem ser identificadas em qualquer área de assunto e são denominadas **categorias comuns**. São assim chamadas porque a característica que as determina ocorre em todos os assuntos. Por exemplo: BRASIL é um termo que poderá aparecer em todos os assuntos (Criminalidade no **Brasil**, Problemas de trâfego no **Brasil**, Relações familiares no **Brasil**). Entretanto, nem sempre a mesma característica, aplicada em diferentes áreas do assunto, resultará no mesmo grupo de termos. Assim, a característica de divisão de ATIVIDADES produzirá:

Em Agricultura: preparação do solo, semeadura, colheita, etc.

Em Engenharia: soldagem, perfuração, lixação, etc.

Examinado a literatura sobre Educação, podemos encontrar, entre outros, os seguintes exemplos de textos dissertativos na área:

QUADRO 1

- 1 - A situação das **escolas governamentais** no **Brasil** no **século XIX**.
- 2 - **O retroprojetor**: um novo **auxiliar didático**.
- 3 - A **ciência** na escola de **1º grau**.
- 4 - Guia das **Escolas Católicas** em **Belo Horizonte** em **1980**.
- 5 - Um levantamento da informação fornecida por **prospectos** de **universidades**.
- 6 - **Escolas profissionalizantes** no **nordeste**.
- 7 - **Padrões** de **aquisição** de **livros** em **bibliotecas** de **escolas** de **2º grau**.
- 8 - **Educação de adultos** e **antropologia**.
- 9 - **Televisão** na escola de **1º grau**.
- 10 - O uso de **slides** no **ensino de química elementar**.
- 11 - **Brinquedos pedagógicos** em **playgrounds** de **jardim de infância**.
- 12 - **Economia doméstica** para **meninas** nas **escolas primárias** da **Inglaterra**.
- 13 - A educação de **crianças** de **áreas de favelas**.
- 14 - Educação **superior** nos **Estados Unidos** na **década de 30**.
- 15 - **Esporte** nas **escolas públicas** da **Rússia**.
- 16 - **Ensino de leitura** para **crianças** com **dislexia**.
- 17 - **Recrutamento** de **professores** **pós-graduados** para **escola primária**.
- 18 - **InSTRUÇÃO** **programada** para **cegos**.
- 19 - **Pesquisa** sobre o uso de **vídeo – tape** na educação de **deficientes mentais**.
- 20 - **Ensino de inglês** para **imigrantes** na **Austrália**.
- 21 - **Educação física** para **deficientes físicos**.
- 22 - **História** em **nível avançado**: um **currículo** para **estudantes** de **tempo parcial**.
- 23 - **Ciência social** nas **politécnicas**.
- 24 - O problema do **vocabulário** na **população** da **ciência**.
- 25 - Uma comparação das atitudes com relação à educação entre **crianças da classe operária** e as de **classe média**.

O exame de literatura revela a terminologia básica de uma área. Nós precisamos decidir quais termos devemos adotar na nossa linguagem de indexação, ou seja, o vocabulário base, o vocabulário de abordagem e as relações entre os termos e conceitos. Essa decisão é facilitada após a análise de assunto quando os relacionamentos se tornam mais evidentes.

Deve-se observar que nesta linguagem foram identificados alguns termos que representam um elemento único, por exemplo, Austrália = área geo-política identificada pelo mesmo nome. Alguns termos podem expressar dois conceitos como **menina** no item 12 que corresponde aos conceitos **criança + sexo feminino**.

Como já visto a natureza dos conceitos pode determinar diferentes categorias de conceitos: categorias comuns e aquelas que incluem conceitos diferentes em diferentes áreas e são chamadas **categorias fundamentais**. Exemplo desta última é a categoria ATIVIDADE vista como exemplo nas áreas de agricultura e engenharia. Coates reconhece as seguintes categorias fundamentais aplicáveis a qualquer área do conhecimento: COISA, PARTE, ATIVIDADE, PROPRIEDADE. Ranganathan, responsável pela criação da abordagem facetada sugeriu as categorias conhecidas por PMEST (PERSONALIDADE, MATÉRIA, ENERGIA, ESPAÇO E TEMPO).

Examinando o quadro 1 veremos que alguns conceitos estarão numa mesma categoria, por exemplo: Brasil, Austrália, Rússia, Inglaterra que denotam o elemento “lugar”. A categoria lugar pode ocorrer em qualquer assunto, assim como século XIX, década de 30 e 1980 denotam “períodos históricos” e também se aplicariam a qualquer assunto. Outro tipo de categoria aplicável a qualquer assunto é usualmente chamada “subdivisão padrão” e denota a forma na qual a informação é apresentada (bibliografias, dicionários, periódicos, entre outras). Estas são as chamadas categorias comuns utilizadas em vários sistemas de classificação.

Em resumo, as categorias comuns incluem os mesmos termos e conceitos em qualquer área, enquanto que as categorias fundamentais podem incluir diferentes tipos de termos e conceitos, mesmo se considerando que a idéia básica do agrupamento efetuado (característica de divisão) seja a mesma (o caso de ATIVIDADES no contexto de agricultura e engenharia).

Se examinarmos o Quadro 1 poderemos então definir quais as categorias fundamentais seriam necessárias à área de educação. Note que, em alguns casos, novos termos foram acrescentados à lista original para melhor demonstrar a abrangência de cada categoria:

Equipamentos: retroprojetor, livros, televisão, slides, brinquedos pedagógicos, carteiras, cadeiras, vídeo-tape.

Agentes (pessoa que exerce a atividade educacional): professores, monitores, pais.

Atividades (operação): ensino, recrutamento, avaliação, aquisição.

Propriedade (qualificação): pós-graduados, tempo parcial, casados.

Instalações: biblioteca, playgrounds, sala de aula, laboratórios.

Métodos de ensino/aprendizagem: instrução programada, aulas expositivas, palestras, estudo de caso.

Métodos de ensino/aprendizagem: instrução programada, aulas expositivas, palestras, estudo de caso.

Curriculum: biologia, ciência, esporte, inglês, economia doméstica, educação física.

Escolas: governamentais, jardins de infância, católicas.

Educandos: adolescentes universitários, crianças, meninas, adultos, deficientes físicos, classe operária, classe média, imigrantes.

Vale notar que alguns conceitos não foram categorizados. Deve-se lembrar que o estabelecimento de uma classificação é arbitrário e subjetivo com relação a se idealizador.

Considere os conceitos:

- Nível avançado
- 1º grau/ 2º grau
- Superior
- Universidade
- Dislexia
- Profissionalizante

*Esses elementos poderiam estar relacionados com diversas categorias, como por exemplo: currículos, educandos, escolas. Concordam? Pense no assunto e dê sua sugestão justificada para o caso.

Criação de linguagens de indexação

2ª Parte - Arranjo sistemático das categorias

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

RAMSDEN, M.J. **An introduction to index language construction; a programmed text.** London: Clive Bingley, 1974.

TURNER, C. **Organizing information; principles and practice.** London: Clive Bingley, 1987.

VICKERY, B.C. **Classificação e indexação.** Rio de Janeiro: BNG/BRASILART, 1980.

BUCHANAN, B. **Theory of library classification.** London: Clive Bingley, 1979.

Ao se construir qualquer linguagem de indexação deve-se considerar a importância dos chamados assuntos compostos e complexos. O processo de criação de classes prevendo somente a existência de assuntos simples é relativamente descomplicado. No entanto, sabemos que cada vez mais o conhecimento evolui resultando na fragmentação e especialização das unidades conceituais, e que o fenômeno de interdisciplinaridade muitas vezes nos surpreende. A literatura principalmente na área científica e técnica tende a cada vez mais apresentar assuntos compostos e complexos. Puranil, discípulo de Ranganathan, analisou assuntos apresentados em periódicos por um período de 50 anos e mostrou que em 1900 mais de 50 % da amostra consistia em assuntos simples, e em 1950 mais de 85% eram compostos mostrando relacionamentos de assuntos. Imagine a situação contemporânea, considerando a evolução científica pós II Guerra Mundial em termos de complexidade e rapidez...

Esta evolução representa um desafio para as propostas de classificação do conhecimento, uma vez, que ela deve acompanhar de perto a mutante estrutura do conhecimento humano.

As se propor estruturar uma área do conhecimento ou mesmo o conhecimento como um todo podemos optar por uma abordagem hierárquica (enumerativa) ou uma abordagem analítico-sintética (facetada).

Abordagem hierárquica

As dificuldades da abordagem hierárquica priorizando as relações gênero x espécie tornaram-se evidentes na medida em que a ciência evoluiu apresentando assuntos cada vez mais complexos. Os esquemas enumerativos são o resultado das idéias e teorias tradicionais que emergiram dos criadores de esquemas de classificações filosóficas e científicas cujos primórdios datam da antiga Grécia.

Os esquemas enumerativos são produzidos por um processo regular de divisão que resulta na determinação de uma posição para todo e qualquer conceito nas tabelas do sistema em uma ou várias hierarquias.

Através de um exemplo simples da área de educação podemos visualizar a elaboração de um esquema enumerativo que possui como ponto de partida a escolha de uma característica

básica, por exemplo: “níveis de ensino”. Esta característica seria então usada para dividir toda a área de educação numa primeira ordem de assuntos de modo a produzir as classes:

Primária
Secundária
Profissionalizante
Superior

É importante que este primeiro agrupamento de assuntos seja exaustivo porque se omitirmos um elemento qualquer possivelmente surgirão problemas futuros na classificação de assuntos. Teremos de considerar uma segunda característica, talvez “disciplinas”, para a subdivisão da primeira ordem de elementos listando todas as disciplinas lecionadas nos diferentes níveis de ensino. Assim teremos:

Primária
Arte
Linguagem
História
Secundária
Arte
Linguagem
História
Profissionalizante
Arte
Linguagem
História
Superior
Arte
Linguagem
História

Uma terceira característica na área de educação poderia ser “Métodos de ensino”. Aplicando-se esta característica ao nosso esquema teríamos a seguinte configuração:

Primária
Arte
Aulas expositivas
Seminários
Laboratório
Linguagem
Aulas expositivas
Seminários
Laboratório
Etc...

Assim podemos ver que o esquema seria desenvolvido em etapas consistentes pela adição sucessiva de características que nos dariam assuntos cada vez mais específicos. É óbvio que este processo envolve um número considerável de decisões. Particularmente, é importante a escolha de qual característica deve ser usada para a divisão primária do assunto. Deve-se considerar também a ordem de aplicação das características e a ordem dos conceitos em cada agrupamento. Um sistema de classificação enumerativo agrupará assuntos do 1º elemento da divisão, enquanto dispersará os assuntos resultantes da aplicação das características subsequentes.

Este processo de produção de esquemas tradicionais de classificação enumerativos é extremamente trabalhoso e complexo. Para ser eficiente em termos de especificidade, deve prever “a priori” todas as relações possíveis entre os conceitos (o que talvez seja impossível) e representá-las nas tabelas.

A etapa seguinte seria o estabelecimento de uma notação que representasse os conceitos e indicasse por configurações lógicas a idéia de subordinação e coordenação entre eles.

Abordagem facetada

Ao analisarmos a literatura na área de educação definimos a existência de categorias comuns e daquelas chamadas fundamentais.

QUADRO 2

Lugar: Brasil, Belo Horizonte, Nordeste, Inglaterra, Favelas, Estados Unidos, Rússia, Austrália.

Tempo: Século XX, 1980, Década de 30.

Subdivisões padrão: prospectos, ensino, vocabulário, pesquisa, periódicos, padrões.

QUADRO 3 – Categorias fundamentais

Equipamentos: Auxiliares didáticos, retroprojetor, livros, cadeiras, carteiras, televisão, slides, quadro negro, brinquedos pedagógicos, vídeo-tape.

Agentes: supervisor, professores, psicólogo, monitores, pais.

Atividades: ensino compra, recrutamento, avaliação, aquisição.

Propriedade: tempo integral, pós-graduados, tempo parcial, decisão exclusiva, graduados, casados.

Instalações: biblioteca, auditórios, playgrounds, sala de aula, museu, laboratórios.

Métodos de ensino/aprendizagem: instrução programada, aulas expositivas, palestras, estudo de caso.

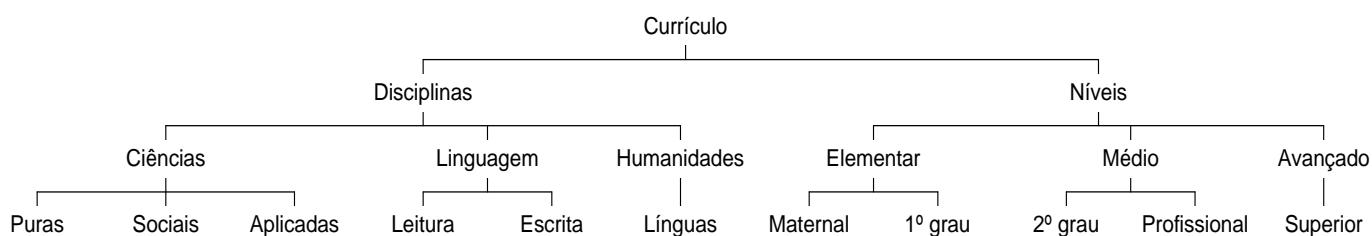
Curriculum: biologia, profissionalizante, jardim de infância, ciência, 1º grau, esporte, elementar, escola primária, inglês, economia doméstica, superior, antropologia, educação física, nível avançado, universidade.

Escolas: governamentais, escolas católicas.

Educandos: dislexia, imigrantes, adolescentes, universitários, crianças, deficientes físicos, meninas, classe operária, adultos, deficientes mentais, classe média, cegos, mulher.

Ao se examinar a lista de categorias encontradas, notar-se-á que algumas “pedem” uma subdivisão. Isto pode acontecer se os termos dentro das categorias são mutuamente exclusivos. Quando os termos são mutuamente exclusivos há uma superposição de definições, o que indica que podem ser combinados para representar assuntos compostos. Por exemplo, na categoria CURRÍCULO, os termos elementar e ciências não são exclusivos, enquanto que

elementar e avançado ou superior seriam. Isto indicaria que a categoria em questão comporta uma subdivisão, o que corresponderia no caso a:



Voltando aos quadros 2 e 3 estabeleça uma estrutura para a organização de conceitos dentro de cada categoria. Na terminologia de Ranganathan seria a ordenação dos “focos” dentro de cada “faceta”. Discuta a importância das facetas para a área de educação e as ordens segundo a definição feita por você. Apresente sua proposta.

Criação de linguagens de indexação

3ª Parte - A notação de um sistema de classificação

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

BUCHANAN, B. **Theory of library classification**. London: Clive Bingley, 1979.

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

Ao estabelecermos uma estrutura para a área de educação definirmos um instrumento que serviria de base para o agrupamento de documentos da área.

Nós vimos que nossa atividade profissional é responsável pela organização de coleções e pela criação de instrumentos de recuperação de informações. Logo, para que seja possível a organização da coleção em si, devemos estabelecer um sistema que permita ao usuário o acesso aos documentos ao mesmo tempo em que dê uma ordenação lógica à coleção. Isto é conseguido pela criação de um sistema de notação.

Ao contrário do arranjo alfabético a ordem sistemática não é evidente por si mesma. Ela é baseada em princípios lógicos, mas não é evidente. Após o estabelecimento da tabela para estruturação do assunto é acrescentada a notação a mesma. A notação deve ter uma ordem evidente por si mesma. As tabelas proporcionam o arranjo sistemático e somente depois de termos decidido sobre a ordem das facetas e focos é possível se pensar em uma notação. Deve-se chamar a atenção de que a notação não é o arranjo sistemático, mas deve espelhar este arranjo. A notação não pode transformar um esquema ruim em um esquema bom de classificação. Mas o sistema de notação deve evidenciar de maneira clara e lógica a estrutura proposta.

Existem dois sistemas de símbolos que têm uma ordem reconhecida amplamente: os números arábicos e o alfabeto latino. Uma notação que empregue um único conjunto de símbolos é chamada notação pura, enquanto que a que usa mais de um tipo é conhecida como notação mista.

Um bom sistema de notação deve incorporar os seguintes quesitos:

- ordenação evidente,
- memorabilidade,
- simplicidade,
- hospitabilidade,
- expressividade.

A notação num sistema de classificação deve ser capaz de fornecer para cada documento um endereço dentro da coleção, deve ser o elemento que permita a correspondência entre o índice alfabético e o arranjo sistemático proposto para a coleção, e deve ainda, possibilitar a expressão de combinações de classes elementares para expressão de assuntos compostos.

Já mencionamos a sintaxe como a extensão da capacidade descritiva do vocabulário. Assim, num sistema de notação para uma linguagem de indexação deve-se estabelecer as regras para a síntese de elementos na representação de assuntos compostos ou complexos. Se

estamos trabalhando com sistemas pré-coordenados (como sistemas enumerativos ou facetados) a previsão da possibilidade de combinação dos elementos e a determinação da ordem de citação dos mesmos é de suma importância.

Considerando o exposto, responda:

1. A notação num sistema de classificação hierárquico é mais complexa do que num sistema facetado?
2. Quais as soluções que você vê com relação à representação de assuntos compostos e complexos nos sistemas enumerativos e nos facetados?
3. Estude uma notação para a nossa proposta de estrutura para a área de educação.
4. Apresente uma proposta para estabelecimento das relações sintáticas.
5. Classifique os documentos da listagem do primeiro texto sobre Criação de Linguagem de Indexação de acordo com sua proposta.

Criação de linguagens de indexação

4^a Parte - Proposta para a área de zoologia

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

BUCHANAN. B. **Theory of library classification**. London: C. Bingley. 1979.

Como já foi visto a criação de um esquema facetado é muito mais fácil do que a de um esquema enumerativo, ainda que várias decisões sejam necessárias: a ordem de citação das facetas considerando a sua importância, o arranjo dos focos nas facetas e a possibilidade do uso de síntese na representação de assuntos compostos ou complexos.

Para se criar um esquema facetado em qualquer área deve-se listar os conceitos elementares que são chamados isolados. Nestes esquemas cada conceito passa a ser um “foco”. Os focos são reunidos e organizados para formar as facetas que representam uma categoria fundamental, e geralmente dão o nome à faceta. Por exemplo, em um esquema de classificação para à área de assistência social possivelmente teríamos uma faceta contendo os termos crianças, imigrantes, classe média, mulheres, operários, idosos, e que provavelmente seria denominada pessoas. teríamos também uma faceta chamada problemas, que conteria focos como sem-casa, doenças mentais, solidão, pobreza, crime, violência, e a faceta terapia, com focos como adoção, prisão, terapia ocupacional, psicanálise.

O termo “faceta” é usado no sentido “aspecto”. As pessoas, os problemas, as terapias são um aspecto da assistência social e poderíamos identificar vários outros aspectos da área.

Deve-se lembrar que muitas vezes as facetas comportam sub-agrupamentos (subfacetas) que seriam definidas com base no fato dos focos serem ou não mutuamente exclusivos. Uma subfaceta é definida como um conjunto de focos produzidos pela aplicação de uma característica de divisão. São meramente pontos de vista sobre o mesmo conjunto de elementos que compõem uma única faceta. É importante chamar a atenção que teoricamente todos os focos de uma faceta poderiam constar de todas as subfacetas criadas. Por exemplo, considerando a faceta pessoas, poderíamos definir as subfacetas faixa etária, sexo, condições sociais. Deve-se chamar a atenção que os focos que compõe a faceta pessoa representam um mesmo conjunto de elementos que podem ser produzidos pela aplicação das seguintes características de divisão: faixa etária, sexo e condições sociais.

Já temos uma definição precisa dos termos “foco”, “faceta” e “subfaceta”. Podemos então estabelecer os passos a serem seguidos na construção de um esquema facetado:

- 1) Examinar uma amostra significativa da leitura da área com o objetivo de identificar os conceitos trabalhados por autores especialistas;
- 2) Agrupar os isolados em facetas, quando se tornarão focos;
- 3) Se necessário, aplicar uma característica de divisão às facetas com propósito de criar subfacetas;
- 4) Organizar os focos nas suas subfacetas usando preferencialmente a ordenação geral/específico ou qualquer outra que no caso seja mais lógica;
- 5) Organizar as subfacetas nas suas respectivas facetas;
- 6) Escolher uma ordem de citação das facetas;

- 7) Estabelecer uma notação para cada classe que servirá como endereço dos documentos, e mostrar sua posição em uma estrutura lógica da área;
- 8) Criar um índice alfabético que remeta à estruturação lógica das classes usando sua notação como “chave”.

Seguindo a metodologia proposta na criação de uma linguagem de indexação para a área de Educação examine os títulos abaixo e proponha uma estrutura para a área de Zoologia.

1. Olhando de perto borboletas e mariposas.
2. Ensaio sobre a fisiologia da fauna marinha.
3. Animais de montanhas.
4. Experiências de laboratório sobre a respiração de vertebrados.
5. Comportamento social de animais.
6. Sistema visual dos anfíbios.
7. Sobre animais venenosos.
8. Fauna do litoral.
9. Como os mamíferos correm.
10. Hábitos alimentares do leão marinho.
11. Aracnídeos tropicais.
12. O papel do jogo no desenvolvimento dos primatas.
13. Pássaros da costa.
14. A tegumento dos insetos.
15. Animais de beira-mar.
16. Os sentidos dos mamíferos marinhos.
17. Animais peçonhentos.
18. Répteis do deserto.
19. Adaptações sensoriais da pele dos macacos Colobus.
20. Pássaros migratórios.
21. Os nomes genéricos de mariposas do mundo.
22. O vôo das abelhas.
23. Necessidades nutricionais dos Pandas.
24. Peixes de água salgada.
25. O cérebro das abelhas.
26. Invertebrados de rios e córregos.
27. Maravilhas da natureza das planícies.
28. Mariposas tropicais.
29. Pássaros de bosques.
30. A vida da águia dourada.
31. Taxonomia entomológica (período sobre classificação de insetos).
32. Hibernação.
33. Agressão ao homem.
34. Inteligência dos ratos.
35. Serpentes equatoriais.
36. Sobre o “rei dos animais”.
37. Como coletar e identificar aranhas.
38. A migração dos pássaros.
39. O leão e a leoa.
40. A reprodução dos mamíferos.

Criação de linguagens de indexação

5ª Parte - Estabelecimento de categorias em zoologia

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

BUCHANAN, B. **Theory of library classification**. London: Clive Bingley, 1979.

O objetivo deste exercício é mostrar uma metodologia para criação de linguagens de indexação em uma perspectiva facetada. Por isso trabalharemos de uma forma limitada quanto ao conceito da área “Zoologia”. Não teremos a intenção de ser exaustivos neste conteúdo, uma vez, que isto implicaria um trabalho extremamente intenso.

Ficaremos limitados aos termos/conceitos relativos aos 40 documentos listados e o acréscimo de alguns termos/conceitos servirá somente para melhor compreensão da estrutura proposta.

Podemos examinar cada um destes títulos e notar que à exceção do número 27 todos são representativos de conteúdo na área de zoologia: pode-se analisá-los em termos de seus componentes. Neste processo devemos lembrar de:

1. Reconhecer os sinônimos ou quase sinônimos (nº 07 e 17, 2 e 24)
2. Diferenciar os homógrafos ou quase homógrafos (nº 10 e 39)
3. Distinguir entre classes de conceitos únicos e classes de conceitos compostos (nº 20 e 38).

Ao identificar os conceitos para as facetas que definimos como necessárias deve-se considerar a existência de documentos que tratam de dois conceitos distintos relativos a uma única faceta (nº1 – Faceta Animais) e a existência de documentos que se referem a focos de diferentes facetas (nº 2 – Facetas Animais, Processo e Forma).

A análise do documento 3 e 4 daria mais elementos para a faceta Animais e também para a categoria Processo, ainda acrescentando um novo foco que poderia estar na faceta denominada Operação. No documento 5 a faceta Animal não está presente no documento já que não se trata de nenhum tipo específico de animal, no entanto “comportamento social” é um tipo de atividade, o que determina a necessidade de uma faceta Atividade. A análise dos enunciados é bastante complexa considerando que muitas vezes definem a necessidade de dois ou mais focos com relação a um conceito único, por exemplo: o documento 18 possivelmente criaria duas classes na faceta Animal “animais do deserto” e “répteis” e o 24 nos daria “animais de água salgada/marininhos/do mar” como também “peixes”. O resultado da análise de assunto dos documentos nos daria a seguinte listagem, considerando as facetas criadas:

Animal

Borboletas

Mariposas

Animais marinhos/do mar/água salgada

Animais da Montanha

Vertebrados

Animais/Fauna

Anfíbios

Animais Venenosos/Peçonhentos

Fauna do Litoral/Animais da Beira Mar

Mamíferos

Primates
Pássaros
Animais da Costa
Insetos
Animais do Deserto
Répteis
Macacos Colobus
Animais Migratórios
Abelhas
Pandas
Peixes
Invertebrados
Animais das Planícies
Animais de Rios/Córregos
Animais Tropicais
Animais dos Bosques
Águia Dourada
Homem
Ratos
Animais Equatoriais
Serpentes
Aranhas/Aracnídeos

Deve-se lembrar que fizemos uma distinção entre as coisas que os animais fazem = Atividades, e as coisas que acontecem aos animais = Processos.

Processos

(Esta categoria compreende elementos relativos aos processos fisiológicos como também as suas partes, pois não seria justificável separar, por exemplo “vista” de “olhos”)

Fisiologia
Respiração
Sistema visual
Correr (sistema locomotor)
Tegumento/pele
Sentidos (sistema sensorial)
Voar (sistema locomotor)
Nutrição/alimentação
Cérebro/sistema nervoso
Reprodução

Atividades

Atividades/comportamento
Comportamento social
Jogo
Hibernação
Migração

Operações

Experiência de laboratório
Terminologia/nomenclatura
Classificação
Coleção
Identificação

Propriedades da Faceta Animal

Agressão
Inteligência

Processos Gerais

Desenvolvimento/maturação
Adaptação

Formas

Ensaios
Periódicos

Vocês devem notar que, muitas vezes, houve necessidade de algumas justificativas e de extrapolarmos em algumas decisões, uma vez, que devemos lembrar que sempre existe uma interferência subjetiva no processo de classificação. Além disso, ficou evidente que quanto menos dominarmos a área, maior será a dificuldade na estruturação e classificação dos conceitos. A decisão de se criar uma faceta especificamente para Atividades foi feita de forma empírica. A justificativa formal para isto seria a possibilidade da existência de assuntos compostos envolvendo focos da faceta Processo e da Faceta Atividades, por exemplo: a respiração durante a hibernação ou comportamento social na alimentação. Quando isto acontece, deveríamos ter duas facetas separadas – isto prova que mais de uma espécie de conceito está envolvido. Por outro lado deve-se chamar a atenção que não é possível se ter documentos relativos a facetas PARTES e PROCESSOS, por exemplo: a vista e o olho. Isto justificaria a necessidade de uma única faceta que combinasse processos fisiológicos e partes.

A próxima etapa seria o exame das facetas com o objetivo de identificar a necessidade de serem subdivididas por mais de uma característica de divisão criando subfacetas. De fato, parece que somente a faceta Animal precisaria ser dividida em subfacetas, por exemplo: Animais migratórios é claramente uma classe definida pelo “comportamento”, enquanto que animais das planícies é uma classe definida pelo “habitat”. Podemos raciocinar da mesma forma que fizemos com relação à separação das facetas Processos e Atividades. Isto é, é possível imaginar a existência de documentos sobre “animais migratórios das planícies”. Isto indica que nós precisamos criar grupos que, no caso, serão subfacetas da faceta Animal. As características de divisão que podem ser usadas na organização da faceta Animal parecem ser:

- habitat: exemplo, “Animais das planícies”
- efeito sobre o homem: exemplo, “Animais peçonhentos”.
- comportamento ou hábitos: exemplo, “Animais migratórios”.
- e, uma quarta característica de divisão com uma dificuldade de ser denominada e cuja aplicação produz espécies familiares de animais como: vertebrados, mamíferos, répteis e insetos. As espécies de animais deste tipo são definidas pelos zoólogos usando critérios complexos. Para um leigo é fácil definir “animais das montanhas” como “animais que vivem nas montanhas”. No entanto, a definição da classe “inseto” exige um detalhamento e conhecimento mais profundo, tal como, “insetos são animais invertebrados cujo corpo é dividido em três seções distintas e possuem 6 pernas”. Nós podemos então, denominar esta subfaceta “Taxonomia zoológica”, uma vez, que usamos os procedimentos classificatórios dos zoólogos para estabelecer os focos.

Neste estágio da construção de um esquema devemos relembrar as regras aristotélicas para divisão lógica das coisas. Este procedimento vai nos assegurar a produção eficiente aplicada nesta operação e que somente uma característica de divisão deve ser aplicada de cada vez.

Isto quer dizer, na nossa situação, que todos os focos de uma sub-faceta devem ser produzidos pela aplicação de uma única característica.

Não há obviamente nenhuma razão que nos faça agir de outra forma, mas muitos esquemas parecem ignorar esta regra, e com isso, sua eficiência fica comprometida. Neste caso poderíamos ter os resultados adversos que seriam classificação cruzada (isto é, a existência de mais de uma classe válida para um único conceito) e o inverso (ou seja, a ausência de classe para o conceito).

Considere por exemplo, esta seqüência da faceta Educando:

Svj - Adolescente
Sw - Educandos fora da faixa etária natural
Swb- Adulto
Swd- Pais
Swf - Donas-de-casa
Swg- Idosos

Está bastante claro que mais de uma característica foi aplicada para produzir este conjunto de focos. Três destes se referem a pessoas pela idade, um representa pessoa pelo relacionamento e um pela ocupação. Como consequência às classes não são mutuamente exclusivas – uma dona-de-casa pode ser uma pessoa idosa e ser também mãe – e não há nenhuma indicação de qual seria a classificação mais adequada para um documento cujo conteúdo se referisse a “Pessoa Idosa” e também a dona-de-casa e mãe. E o que é mais complicado, a impossibilidade de se expressar à classe “Mães adolescentes”, porque a única classe que representa o conceito “pais” está enumerada na seqüência fora da “faixa etária natural”.

A observância da regra de divisão lógica mencionada vai produzir subfacetas cujos focos são mutuamente exclusivos e isto evitará este tipo de problema. Além disso, vai assegurar que a ordem de arquivamento entre as classes foi considerada com relação à ordem de citação, isto é: o princípio de inversão foi observado. Isto não é possível quando os focos produzidos pelo emprego de diferentes características foram reunidos em uma única seqüência como no caso mencionado.

Vamos agora aplicar as quatro características que definimos para a faceta Animal com o objetivo de estudar as possíveis subfacetas.

Faceta Animal

Animal/Fauna

(Subfaceta habitat)
Marinhos/ mar/ animais de água salgada
Animais das montanhas
Fauna litoral/ animais de praia
Anfíbios
Animais da Costa
Animais do deserto
Animais do rio
Animais das planícies
Animais tropicais
Animais da floresta

Animais equoriais

(Subfaceta animais por hábito)

Animais migratórios

(Subfaceta efeitos sobre o homem)

Animais venenosos

(Subfaceta taxonomia)

Borboletas

Mariposas

Vertebrados

Anfíbios

Mamíferos

Rapinas

Primates

Pássaros

Insetos

Répteis

Macacos Colobus

Abelhas

Pandas

Peixes

Invertebrados

Águia Dourada

Homem

Ratos

Cobras/serpentes

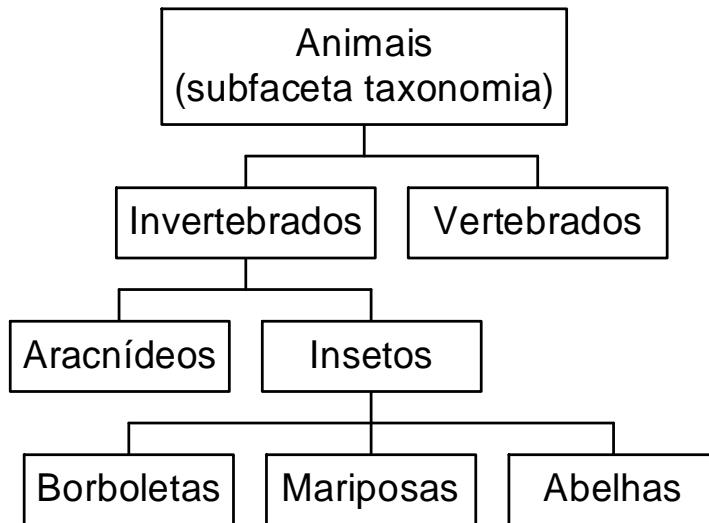
Aranhas

Rã

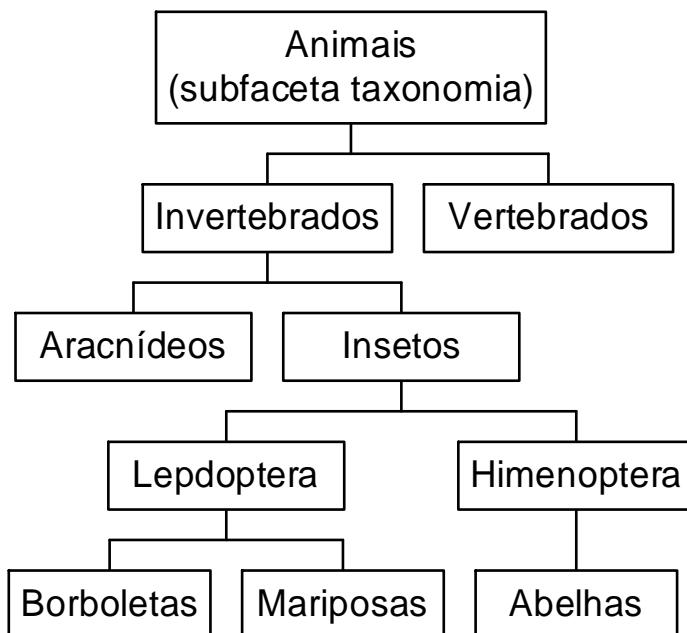
Aracnídeos

Agora devemos aplicar uma outra regra de divisão lógica. Esta é relativa a exaustividade na aplicação de uma característica de divisão, isto é, deve-se incluir todas as classes produzidas pela aplicação de cada característica. Já chamamos a atenção que o nosso esquema não será completo, mas vamos apresentar alguns focos com o objetivo de melhorar a compreensão. Se temos “animais maléficos” (agressivos ao homem) teríamos também “animais benéficos”. Se temos animais migratórios na sub-faceta Animais por Hábito, teríamos “animais hibernadores” (principalmente se temos uma classe para hibernação na faceta ATIVIDADES), considerando o habitat poderíamos incluir “animais de planaltos”, “animais de lagos”, “animais polares”. Para a subfaceta Taxonomia Animal podemos incluir alguns companheiros para a classe “cobras” que seriam as lagartixas e tartarugas.

A terceira regra da divisão lógica também se refere à não omissão de classes. Ela define que a divisão deve ser gradativa considerando o fator gênero e espécie, isto é, nenhuma classe super ordenada (superior) deve ser omitida. Se esta regra não for observada dizemos que a cadeia de classes gerais para específicas não foi modulada. Falhas na modulação resultam não só na impossibilidade de se expressar classes omitidas, como também, evidenciam dificuldades do compilador do esquema no reconhecimento de relações. Como exemplificado nesta hierarquia:



Neste caso a cadeia Animais – Invertebrados – Insetos – Borboletas não modula. Classes que são mais específicas que inseto e mais amplas que as classes referentes no próximo passo foram omitidas. Por exemplo, entre insetos e borboletas a classe Lepidoptera foi omitida, e entre insetos e abelhas se omitiu a classe Hymenoptera. Naturalmente estas classes esquecidas não poderiam ser expressas e ao mesmo tempo mostrou-se que o compilador do esquema não tinha uma idéia clara entre as borboletas, mariposas e abelhas. Uma hierarquia correta seria:



Vamos examinar nosso esquema para inserir classes onde necessárias lembrando sempre que estaremos sendo seletivos. Na subfaceta Efeitos Sobre o Homem a classe “animais maléficos” é superordenada a “animais venenosos”, simplesmente na subfaceta Habitat deixamos de incluir a classe “animais aquáticos” que seria superordenada a “animais marinhos”, e também considerando “animais de lagos” e “animais de rios” não incluímos uma classe que seria “animais de água doce”. Tendo incluído “animais aquáticos” parece que seria razoável também incluir águias, ursos. Naturalmente na criação de um esquema geral acharíamos muitos outros focos que deveriam ser incluídos, mas estamos trabalhando com o propósito de demonstração.

Agora já temos alguma clareza quanto à organização dos focos em subfacetas da faceta Animal obedecendo às regras da divisão lógica, de modo que este processo seja feito de forma eficiente. Podemos notar que algumas subfacetas ainda pedem um reagrupamento, como é o caso de “animais por habitat” que deveria ser dividida “por latitude” (animais tropicais, animais polares), “por elemento” (animais terrestres, animais aquáticos), “por tipo de relevo” (animais das montanhas, animais de planícies), “por tipo de solo” (animais de deserto, animais de florestas). Devemos lembrar que a necessidade dessas subfacetas se deve a possibilidade de assuntos compostos tomados das diferentes classes, por exemplo, poderíamos ter documentos sobre “animais aquáticos de florestas tropicais” ou “animais terrestres de montanhas equatoriais”.

Como vimos nenhuma das outras facetas necessitou deste tipo de tratamento, pois não refletiram mais do que uma característica de divisão. Isto significa que a primeira regra de divisão lógica não se aplica a elas, mas a outras duas referentes à omissão de classes e modulação. Assim, para uma melhor visualização, vamos acrescentar mais dois focos a faceta Operação. “Experiências” que é uma classe superordenada a “experiências de laboratório” e a faceta Processo incluiremos a classe “locomoção” que é superordenada a “correr e a voar”. Também incluiremos “sistema Nervoso”. O resultado dos agrupamentos dos focos na faceta animal seria o seguinte:

(Faceta animal) Animais / Fauna			
(Subfaceta efeitos sobre o homem)		(Subfaceta animais por hábito)	
Animais benéficos		Animais migratórios	
Animais maléficos		Animais hibernantes	
Animais venenosos			
(Subfaceta animais por habitat)			
(Por latitude)	(Por elemento)	(Por tipo de relevo)	(Por tipo de solo)
Animais equatoriais	Animais aquáticos	Animais das montanhas	Animais do deserto
Animais polares	Animais terrestres	Animais das planícies	Animais das florestas
Animais tropicais	Animais marinhos	Fauna Litorânea	Animais da selva
	Anfíbios	Animais da Costa	
	Animais de rios	Animais do planalto	
	Animais de lagos		
	Animais de água doce		
(Subfaceta taxonomia)			
Borboletas		Cobras	
Mariposas		Serpentes	
Vertebrados		Aves	
Anfíbios		Aranhas	
Mamíferos		Rãs	
Aves de rapina		Sapos	
Primates		Largatixas	
Pássaros		Tartaruga	
Insetos		Lepidoptera	
Répteis		Hymenoptera	
Macacos Colobus		Aracnídeos	
Abelhas		Macacos	
Ursos			

Peixes
Invertebrados
Águia
Águia Dourada
Homem
Ratos

Criação de linguagens de indexação

6ª Parte - Arranjo sistemático em zoologia

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto e Marysia Malheiros Fiúza.

BUCHANAN.B. **Theory of Library Classification**. London: Clive Bingley, 1979.

Depois de agruparmos os focos é necessário colocá-los em ordem dentro dos grupos formados. Tomando por exemplo, a sub-faceta “habitat”, teremos em primeiro lugar o subgrupo “por latitude”. Nela encontraremos uma relação gênero – espécie: os animais equatoriais são obviamente uma espécie de animal tropical, uma vez, que a zona equatorial é uma parte da zona tropical. Assim colocaremos as duas classes na ordem: assunto geral – assunto específico.

Animais tropicais
Animais equatoriais

Entretanto não há uma relação “assunto geral – assunto específico” entre as classes Animais de Zona Temperada, Animais Polares e Animais Tropicais; estas classes são coordenadas e deveremos escolher uma ordem lógica para organizá-las. Chamaremos esta ordem lógica de “array” ou “ordem em fileira”. Quando se trata de conceitos envolvendo geografia a base apropriada para ordenação entre os conceitos seria a proximidade espacial e poderíamos aplicá-la intuitivamente através da escolha do direcionamento do equador aos pólos:

Animais tropicais
Animais equatoriais
Animais da zona temperada
Animais polares

ou o reverso

Animais polares
Animais de zona temperada
Animais tropicais
Animais equatoriais

Escolhemos a segunda versão, pois nos parece mais apropriado movimentar-se dos climas mais frios para os mais quentes e, além disso, a primeira versão parece interromper a progressão partindo do trópico para o equador antes de passar nas zonas temperadas.

Na subfaceta “por elemento” aplicaremos os mesmos princípios, embora de uma forma um pouco mais complexa. Devemos considerar a ordenação em fileira e também encontrar as relações gênero-espécie:

Animais Aquáticos	Animais terrestres	Animais anfíbios
Animais de rio		
Animais marinhos		
Animais de lago		
Animais de água doce		

Considerando os “animais aquáticos” notar as relações gênero-espécie:

Animais aquáticos

Animais marinhos

Animais de água doce

Animais de lago

Animais de rios

Devemos então escolher as ordenações em fileira “arrays”. A primeira é entre Animais de Lagos e Animais de Rios. Como os rios tendem a nascer em lagos, teremos a ordem:

Animais de lagos

Animais de rios

Para escolher a ordenação entre Animais Marinhos e Animais de Água Doce (com suas classes subordinadas Animais de Rios, Animais de Lagos), poderemos usar o princípio de trabalho (mares maiores que rios e lagos).

Animais aquáticos

Animais marinhos

Animais de água doce

Animais de lagos

Animais de rios

Por fim teremos que ordenar as classes coordenadas Animais aquáticos, Animais terrestres, Animais anfíbios. Como a vida provavelmente começou na água, colocaremos os animais aquáticos antes dos terrestres e os anfíbios entre os dois. Assim nossa ordenação final da subfaceta será:

Animais aquáticos

Animais marinhos

Animais de água doce

Animais de lagos

Animais de rios

Animais anfíbios

Animais terrestres

Para as subfacetas “por relevo de solo” sugerimos as seguintes ordenações:

(Por relevo de solo)

Animais da costa

Fauna

Animais de planície

Animais de terra elevada

Animais de Colina

Animais da montanha

Animais de planalto

(Por tipo de solo)

Animais do deserto

Animais de pastagem

Animais de capoeira

Animais de florestas

Animais da selva

Para completar nosso trabalho com a subfaceta “animais por habitat” precisamos colocar os focos em ordem para podermos decidir como arquivar Animais de Planície (seria antes ou depois de Animais de Pastagens) e Animais de Água Doce (seria antes ou depois de Animais Tropicais?). Este é o problema da ordem de arquivamento entre as classes colaterais e é

decidido pelo princípio de inversão, isto é, devemos decidir a ordem de citação entre as quatro subfacetas e organizá-las na ordem inversa – se desejarmos que nossos documentos sejam colocados nas estantes da ordem geral para o específico.

Tomemos, por exemplo, um documento sobre Animais Terrestres na Fauna Litorânea Florestal Tropical. Este documento contém elementos referentes as quatro subfacetas. Como ainda não existe um sistema lógico que determine a ordem de citação, ou melhor, existem várias maneiras de se ver o problema, vamos definir uma ordem de citação arbitrária: “por elemento” – “por latitude” – “por tipo de solo” – “por tipo de relevo”. A ordem de citação para o nosso documento seria: Animais Terrestres – Animais Tropicais – Animais de Florestas – Animais da Costa.

Usando o princípio de inversão a organização final dos focos da subfaceta “animais por habitat” seria:

(Animais por habitat)

(Por tipo de relevo)

Animais da costa

 Fauna litorânea / animais de praias

Animais das planícies

Animais das terras elevadas

 Animais das colinas

 Animais de montanha

 Animais do planalto

(Por tipo de solo)

Animais do deserto

Animais das pastagens

Animais de capoeiras

Animais de florestas

Animais da selva

(Por latitude)

Animais polares

Animais de zonas temperadas

Animais tropicais

Animais equatoriais

(Por elemento)

Animais aquáticos

 Animais marinhos/animais de mar/ animais de água salgada

 Animais de água doce

 Animais de lagos

 Animais de rios

Animais anfíbios

Animais terrestres

Segundo esta ordenação o arquivamento de documentos seria feito do geral para o específico, ou seja, o nosso documento seria arquivado depois de documentos de classes mais gerais, tais como:

Doc.1. Animais Terrestres – Animais Tropicais – Animais de Florestas

Doc.2. Animais Tropicais – Animais de Florestas - Animais da Costa

Doc.3. Animais de Florestas – Animais da Costa

Usando o mesmo método nas outras subfacetas da faceta animal, teríamos:

(Subfaceta efeito sobre o homem)

Animais benéficos

Animais nocivos

Animais venenosos

(Subfaceta animais por hábito)

Animais hibernantes

Animais migratórios

Esta ordem em fileira foi escolhida assumindo-se a Hibernação é uma reação mais primitiva do que a migração.

(Subfaceta taxonomia)

Invertebrados

Aracnídeos

Aranha

Insetos

Lepidóptera

Mariposa

Leaf moths

Hymenoptera

Abelhas

Abelhas produtoras de mel

Vertebrados

Peixes

Anfíbios

Rãs

Sapos

Répteis

Cobras

Lagartos

Serpentes

Tartarugas

Aves

Rapaces / aves de rapina

Águias

Águias douradas

Mamíferos

Roedores

Ratos

Carnívoros

Leões

Leões marinhos

Ursos

Pandas

Primates

Macacos

Macacos Colubus

Homem

A ordenação em fileira desta subfaceta segue mais ou menos a usada pelos zoologistas, sendo baseada na ordem de desenvolvimento – ex. Invertebrados – Vertebrados peixes – Anfíbios – Répteis – Aves mamíferos.

Nossa última decisão na organização da faceta animal é a ordem de arquivamento entre as subfacetas “habitat”, “efeito sobre o homem”, “hábito” e “taxonomia”. Devemos ordená-las ao inverso da ordem de citação que escolhemos: Taxonomia – Habitat – Hábitos – Efeito sobre o homem.

Depois de organizarmos os focos na faceta Animal podemos tratar as outras facetas da classe zoologia da mesma maneira, embora nenhuma delas apresente as complicações de subfacetas. Teremos, então, somente uma decisão final: a escolha da ordem de arquivamento entre as facetas. Os focos nas diferentes facetas são classes colaterais do primeiro tipo – são espécies diferentes de coisas (animais, processos, operações, entre outros). A ordem de arquivamento das facetas deve ser o inverso da ordem de citação para classes compostas, se desejarmos conservar a ordem geral-específico entre os documentos ou registros de documentos. Os três princípios de escolha de ordem de citação: concretividade crescente, finalidade e, talvez consenso – indicam como faceta primária: Animais.

O consenso certamente indicaria como faceta secundária os Processos fisiológicos e partes. Como as Atividades (ex. Jogo) dependem dos atributos (ex. Inteligência) como também dos Processos Fisiológicos, citaríamos estas facetas na seguinte ordem:

Processos Fisiológicos – Atributos – Atividades

A faceta Processos Gerais cujos focos são Desenvolvimento/Maturação e Adaptação deve vir em seguida e depois a faceta Operações e em último lugar a faceta Forma.

Assim a ordem de citação ideal para a área seria:

Animais – Processos Fisiológicos – Atributos – Atividades
– Processos Gerais – Operações – Forma.

O que geraria a ordem de arquivamento:

Forma – Operações – Processos gerais – Atividades – Atributos
– Processos Fisiológicos – Animais.

Então o nosso esquema de classificação será:

Zoologia

(Faceta Forma de Apresentação)

Ensaios
Periódicos

(Faceta Operação)

Coleção
Experiências
Experiências de campo
Experiências de laboratório
Identificação
Classificação / Taxonomia
Nomenclatura

(Faceta Processos Gerais)

Desenvolvimento / Maturação
Adaptação

(Faceta Atividades)

Atividades / Comportamento
Comportamento social
Jogo
Hibernação
Migração

(Faceta Atributos)

Inteligência
Agressão

(Faceta Processos Fisiológicos e Partes)

Anatomia

(n.b. órgãos específicos são enumerados abaixo, sob sua função ou sistema)

Fisiologia

Pele
Sistema Nervoso
Cérebro
Sentidos
Vista / Sistema visual
Respiração
Nutrição / Alimentação
Locomoção
Correr
Voar
Reprodução

(Faceta Animal)

Animais / Fauna

(Subfaceta por efeito sobre o homem)

Animais benéficos
Animais nocivos
Animais venenosos / Animais peçonhentos

(Subfaceta hábito)

Animais hibernantes
Animais migratórios

(Subfaceta habitat)

(Por relevo do solo)
Animais da costa
Fauna litorânea / Animais de praia
Animais de planície
Animais de terras elevadas
Animais de colinas
Animais de montanhas
Animais de planaltos
(Por tipo de solo)
Animais do deserto
Animais de pastagens
Animais de capoeiras

Animais de florestas
Animais de selvas
(Por latitude)
Animais polares
Animais de zonas temperadas
Animais tropicais
Animais equatoriais
(Por elemento)
Animais aquáticos
Animais marinhos / Animais do mar / Animais de água salgada
Animais de água doce
Animais de lagos
Animais de rios
Animais anfíbios
Animais terrestres

(Subfaceta Taxonomia)

Invertebrados
Aracnídeos
Aranhas
Insetos
Lepidóptera
Mariposas
Leaf moths
Borboletas
Hymenoptera
Abelhas
Abelhas produtores de mel

Vertebrados
Peixes
Anfíbios
Rãs
Sapos
Répteis
Cobras
Lagartos
Serpentes
Tartarugas
Aves
Rapaces / Aves de rapina
Águias
Águias douradas

Mamíferos
Roedores
Ratos
Carnívoros
Leões
Leões marinhos
Ursos
Pandas

Primates
Macacos
Macacos Colubus

Homem

Criação de linguagens de indexação

7ª Parte - *Ordem de citação X ordem de arquivamento*

Como vimos, a ordem de citação se refere ao estabelecimento de prioridades na indexação de documentos de assuntos compostos ou complexos. Deve-se definir a importância das facetas que compõem este assunto. Ao se criar uma linguagem de indexação pré-coordenada é extremamente importante esta definição, pois dela dependerá todo o processo de recuperação de informações. Escolhendo uma das características de divisão como primária ou a mais importante estaremos estabelecendo inclusive a estrutura de organização da coleção propriamente dita. As facetas consideradas como secundárias, terciárias, ..., estarão dispersas sob o primeiro ponto de acesso escolhido.

Ao se organizar uma coleção de documentos é interessante fornecer-lhe uma ordenação lógica, ou seja, manter juntos assuntos relacionados e mostrar o desenvolvimento do assunto geral para o específico.

Foram estas as considerações que orientaram a estrutura proposta para a área de ZOOLOGIA, tanto relativo a ordem de citação, quanto, a ordem de arquivamento.

Ordem de citação

Animais – Processos fisiológicos – Atributos –
Atividades – Processos gerais – Operações - Forma

Ordem de arquivamento

Forma – Operações – Processos gerais – Atividades
– Atributos – Processos fisiológicos - Animais

1 - Examine os enunciados relativos a 5 documentos na área de Zoologia:

- DOC.1 Bibliografia de zoologia
DOC.2 Adaptação da pele dos ursos polares para hibernação
DOC.3 O processo de hibernação
DOC.4 Ensaios sobre animais hibernantes
DOC.5 A hibernação dos animais que vivem no pólo norte

2 - Estabeleça a ordem de citação para os elementos de cada documento.

3 - Mostre como ficarão colocados na estante.

4 – Examine o modo como os documentos foram arquivados e comente se esta organização estaria de acordo com o princípio geral – específico.

Criação de linguagens de indexação

8ª Parte - Notação em um sistema de classificação em zoologia

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

BUCHANAN, B. **Theory of library classification.** London: C. Bingley, 1979.

Depois de criarmos uma estrutura para organizar uma coleção de documentos em Zoologia agrupando-os por assunto (segundo a ordem de citação estabelecida) e arquivando-os numa ordem lógica (segundo a ordem de arquivamento) é possível classificar todos os 40 títulos nos quais nos baseamos.

A ordenação proposta pela nossa estrutura é lógica, porém não é evidente. Para que ela seja inteligível é necessário que usemos um sistema de notação como já vimos anteriormente. Um sistema de notação nos garante:

- a possibilidade de localizar cada um dos documentos numa coleção ao mesmo tempo que nos permite dar uma ordenação lógica à mesma;
- a possibilidade de expressar assuntos compostos e complexos pela síntese de seus elementos notacionais constitutivos;
- a correspondência entre a linguagem natural, alfabética usada no índice e a estrutura proposta para organização da notação da coleção.

A notação não é essencial a um sistema de classificação, mas obviamente representa uma série de vantagens a este. A notação é usada para mostrar a ordem preferencial, mas determina esta ordem. Como Bliss diz, ela é “subsidiária” e nós “só devemos” acrescentá-la ao sistema depois de tomadas as decisões sobre a ordem de citação e arquivamento.

Sem dúvida a aceitabilidade de um esquema depende bastante das qualidades da notação, especialmente, se esta apresenta facilidade de uso: a Biblioteca da Escola de Biblioteconomia de Wales deixou de usar um excelente esquema de classificação para a área, feito pelo CRG, em favor da Classificação de Dewey (bastante inferior em qualidade para a área) pela complexidade da notação daquela classificação facetada.

“Os bibliotecários estão tão acostumados a ver anotação em primeiro lugar nas telas dos sistemas de classificação e nas fichas dos catálogos, que são levados a pensar na notação como elemento mais importante, mas a verdade é que a classificação é a coisa mais importante, e a notação a despeito de sua utilidade não faz a classificação, podendo, no entanto, comprometê-la”. BLISS (1979).

Se não quisermos prejudicar nosso esquema de classificação devemos considerar as propriedades e qualidades da notação cuidadosamente. Estas caem em dois grupos:

- aquelas que afetam a facilidade de uso;
- aquelas que asseguram que cada classe tenha uma notação única.

As decisões a serem tomadas neste ponto seriam:

- que conjunto de símbolos usar;
- que tamanho (extensão) deveria ter;
- se a notação deve evidenciar relações hierárquicas

- se deveríamos permitir colocações alternativas para a mesma classe / conceitos.

Essas decisões resultarão em procedimentos complexos e altamente inter-relacionados.

As notações devem ser baseadas em numerais ou letras, sendo que, os numerais mostram de uma forma mais clara a ordenação proposta. Compare estas duas notações para a classe Futebol no Bliss Classification e na CDD representadas respectivamente por HGK e 796.344 (futebol). Qual serviria de forma mais clara para um endereço? Essa é uma das grandes vantagens da CDD: sua notação simples mostra claramente a dependência hierárquica entre os assuntos e dá um endereçamento fácil dos documentos da coleção.

Usando um sistema de notação mista o valor dos números (ou símbolos) formados fica bastante complicado. Usando-se letras e números o que arquivar primeiro? E se adotarmos outros símbolos como parênteses, dois pontos, vírgula, ponto e vírgula, entre outros, a situação fica ainda mais complexa. A notação mista, ao contrário da pura, possui como grande vantagem a maior base para notação. Se usarmos os numerais, por exemplo, teremos 10 símbolos (0 a 9) ou o alfabeto teremos 26 letras, enquanto que se combinarmos ambas as opções nossa base passará de 36 símbolos. O benefício de uma notação longa é questionável em termos do arranjo de coleções, uma vez, que Bliss sugere que o “limite econômico” de uma notação de classe seria de 3 ou 4 dígitos.

As notações que mostraram as relações de subordinação pela adição de um novo símbolo a cada subdivisão sucessiva são chamadas hierárquicas ou estruturais, por exemplo:

D5 Veículos
D51 Veículos terrestres
D515 Veículos ferroviários
D5153 Trens de passageiros
D51532 Trens de 1 classe

Outro tipo de notação, não hierárquica, é a notação expressiva que mostra somente uma estrutura sintética. Veja os exemplos:

CDD(dewey)		BC1(Bliss)
796.3	Jogos de bolas	HKE Jogos de bolas
796.33	Futebol	HKF Futebol
796.333	Rugby	HKG Futebol Inglês
796.334	Futebol inglês	HKH Hokey
796.34	Jogos de raquete	HKI Pólo
796.342	Tênis	HKJ Lacrosse

A notação expressiva tem a vantagem da brevidade, mas a notação hierarquia é mais “agradável” no momento que mostra de forma tão evidente a estrutura e relacionamentos entre os assuntos. As notações expressivas muitas vezes confundem os usuários, principalmente, quando o número de símbolos relativos de uma classe não mostra as relações de forma clara. Veja na Classificação de Bliss:

JHN Professoras mulheres
JHNP Professores homens

JKG Língua materna
JKGY Leitura e escrita
JKH Leitura

Uma outra vantagem da notação hierárquica é a possibilidade de ampliarmos ou especificarmos nossa estratégia de busca mais facilmente. Se procurarmos documentos sobre Futebol inglês 796.334 e não encontrarmos nada, seria lógico procurar em 796.333, pois saberíamos de antemão não se referir ao assunto, mas a outro tipo de futebol. Deveríamos procurar em 796.33, pois documentos sobre futebol em geral possivelmente tratariam do futebol inglês.

A tendência nos sistemas especializados modernos tem sido estabelecer uma notação expressiva pela sua brevidade e talvez pela dificuldade das notações hierárquicas quanto à sua hospitalidade.

A capacidade de uma notação ser mnemônica tem importância menor ao considerarmos as qualidades de um sistema, mas sem dúvida seria um ponto a favor no processo da escolha entre dois tipos de sistema. Num sistema facetado o uso do mesmo símbolo para representar o mesmo conceito, em qualquer contexto, facilita a memorização do mesmo. Por exemplo, os parênteses na CDU sempre representam a área. Mesmo num esquema hierárquico a repetição de um mesmo símbolo leva a sua memorização. Por exemplo, na CDD independente da literatura, 1 será sempre poesia, 2 teatro, 3 ficção...

O uso dos sistemas de classificação em bibliotecas pretende dar uma ordenação ideal tendo em vista o interesse dos usuários. Assim é importante que o sistema possibilite classificações alternativas para conceitos. Na medida em que o interesse dos usuários varia de biblioteca para biblioteca.

“As classificações são relativas e adaptáveis, e os sistemas devem, dentro do possível, servir a interesses diferentes, desde que não incompatíveis. Coleções alternativas devem então ser oferecidas”. Bliss (1939).

Os sistemas neste caso seriam “flexíveis” e a flexibilidade seria oferecida pela notação. Existem dois métodos para prover a flexibilidade: um seria dar alternativas de classificação para um mesmo conceito em duas classes diferentes, o outro seria permitir ao classificador mudar a ordem de citação, segundo seu interesse. Em qualquer um dos dois casos, uma vez que a biblioteca escolha uma determinada coleção nunca deverá usar a outra possibilidade.

Por exemplo, na classificação de Bliss para o assunto “Ciência e Tecnologia de Materiais” reservou-se duas notações: BR no contexto de Ciências e UG no contexto de Tecnologia. Adotando-se uma notação a outra nunca deve ser usada.

Se o indexador escolher uma outra ordem de citação deve assegurar que a ordem de arquivamento parte do geral para o específico. Isto é conseguido também reservando-se diferentes notações para a mesma faceta ou sub-faceta, ou de outro modo, usando-se símbolos introdutórios de facetas cujo valor para arquivamento poderia ser alterado.

No primeiro caso, veja o exemplo de Bliss:

JI	Faceta métodos de ensino
JK	Faceta currículos
JL/JV	Faceta educandos
JW	Faceta métodos de ensino
JY	Faceta currículos

Este esquema dá ao indexador a opção das seguintes ordens de citação:

1^a: Educando – Currículo – Métodos de Ensino
2^a: Currículo – Métodos de Ensino- Educando

Se o indexador escolher a 1^a terá como ordem de arquivamento:
Métodos de Ensino – Currículos – Educandos

Se escolher a 2^a terá como ordem de arquivamento:
Educandos – Métodos de Ensino – Currículos

No primeiro caso deixaria sem uso as classes JW e JY e no 2^a caso não se usariam as notações JI e Jk.

O problema de indexação de assuntos compostos ou complexos também pode ser solucionado via uso de símbolos introdutórios de facetas. Na literatura da área são conhecidos como “Indicadores da faceta”, em inglês “Intercalator ou Fence”.

Estes símbolos devem ter um valor para arquivamento ao mesmo tempo em que fornecem uma noção do tipo de conceito que precedem.

Se não usarmos estes intercaladores podemos ter sínteses que serão confundidas com números de classes, como por exemplo:

W Indexação do assunto
WOD Arranjo sistemático
O Faceta material
OD Ilustrações

Se não usássemos um indicador de facetas, teríamos a mesma notação WOD tanto para indexação de ilustrações quanto para o arranjo sistemático.

A solução seria, por exemplo, usar uma barra para separar as facetas ou mesmo um símbolo qualquer que indicasse inclusive o tipo de faceta que está sendo introduzida = “W/OD ou W,OD”. Deveria-se também estabelecer o valor dos símbolos para arquivamento, como por exemplo:

W Indexação do assunto
B Faceta tipo de bibliotecas (introduzida pelo símbolo dois pontos :)
OD Ilustrações (introduzidas pelo símbolo ponto e vírgula ;)

Sendo a ordem de citação: operação – tipo de material – tipo de biblioteca, um documento sobre indexação de ilustrações em bibliotecas escolares seria identificado em W;OD:BE, e isto definiria a seguinte ordem de arquivamento:Tipos de bibliotecas – Tipo de material – Operação.

O uso da notação mista com indicadores de faceta torna a notação mais complexa e difícil de usar. Mas como os símbolos não têm valor para arquivamento intrínseco é possível estabelecer a ordem de citação que melhor convier.

Um exemplo na CDU ilustra este fato. A classe 8 (Literatura) é facetada. Assumindo-se que, o que chamamos faceta “Trabalho Literário” incorpora sub-facetadas períodos, forma e língua, deve-se citar primeiramente a língua, juntando sua notação específica ao símbolo 8 de literatura.

- 820 Literatura Inglesa
830 Literatura Alemã
840 Literatura Francesa

As sub-facetas forma e período são introduzidas pelo indicadores – (hífen) e “ “ (aspas):

- | | |
|--------------|----------------------------------------------------|
| - 1 Poesia | “I7” século I8 |
| - 2 Teatro | “I8“ século I9 |
| - 3 Prosa | “I9 “ século 20 |
| - 3I Romance | “I939/45” Literatura 2 ^a Guerra Mundial |

Classificando 6 trabalhos segundo ordens de citação diferentes:

- 1) Literatura Inglesa do século I9
- 2) Romance Inglês
- 3) Teatro Inglês do Século I9
- 4) Literatura Inglesa
- 5) Romance Inglês do século I9
- 6) Teatro Inglês

L/P/F

- 1) 820”I8^a
- 2) 820-3I
- 3) 820”I8^a -21
- 4) 820
- 5) 820”I8^a -31
- 6) 820-2

L/F/P

- 1) 820 “ I8^a
- 2) 820-31
- 3) 820-2 “ I8^a
- 4) 820
- 5) 820-31 “ I8^a
- 6) 820-2

O que definiria a seguinte ordem de arquivamento:

L/P/T

- 1) 820
- 2) 820-2
- 3) 820-31
- 4) 820 “ I8^a
- 5) 820 “ I8^a -2
- 6) 820 “ I8^a -31

L/F/P

- 1) 820
- 2) 820 “ I8^a
- 3) 820 -2
- 4) 820 -2 “I8^a
- 5) 820 -3I
- 6) 820 31 “ I8^a

O último aspecto a ser considerado quanto à notação seria a “hospitalidade”, ou seja, a capacidade do sistema de aceitar novos conceitos/classes. Isto é conseguido se toda notação existente for divisível, E este é um dos pontos a serem observados na criação de um sistema de notação, inclusive deixando-se classes vagas para possíveis expansões.

A necessidade de especificar em um esquema de classificação conceitos individuais também está relacionada com a hospitalidade . A CDU permite que se acrescente aos números um recurso individualizante chamado A/Z, por exemplo: para se individualizar os sistemas de classificação bibliográfica poderíamos usar 025.400, 025.4CDD, 025.4CDU, de modo a usar uma ordenação mais lógica, subdividindo esta parte da coleção se o volume do documento o exigir.

Enfim, ao se escolher um sistema deve-se avaliar as propriedades da notação escolhida assegurando que apresente códigos únicos para as classes de conceitos com valores inequívocos de arquivamento e facilidade de uso.

Agora que já vimos às possibilidades de escolha na criação de um sistema de notação podemos aplicá-la ao nosso sistema de classificação em Zoologia. Podemos considerar as possibilidades de vários tipos de sistema criando:

- a) um sistema de notação pura, hierárquica, usando somente letras maiúsculas sem indicadores de faceta;
- b) um sistema de notação pura, hierárquica, usando somente letras com indicadores de faceta;
- c) Um sistema de notação mista, não hierárquica, com indicadores de faceta.

Criação de linguagens de indexação

9ª Parte - Criação de índice Alfabético

Tradução e adaptação de Maria Cristina Mello Ferreira Pinto.

BUCHANAN.B. **Theory of library classification**. London: Clive Bingley. 1979.

Um índice, no sentido restrito usado neste capítulo, é uma lista de termos numa ordem conhecida que mostrará a localização destes termos em outra seqüência cuja ordem não é tão óbvia. Nós necessitaremos criar este tipo de índice para o esquema feito para a área de zoologia, uma vez, que o usuário ou o indexador não dominarão necessariamente a simbologia usada para a representação de conceitos na estrutura proposta. Este índice poupará tempo do usuário e do indexador na localização de informações. A notação será a chave entre a linguagem usada por eles e aquela definida pelo esquema. Tudo o que temos a fazer será reorganizar os termos que representam as classes de conceitos no esquema de acordo com uma estrutura alfabética e mostrar a notação representativa de cada classe :

Adaptação	DF
Agressão	FH
Anatomia	GH
Anfíbios	HQK
Animais	H
Animais Venenosos	HJJH
Atividades	EF
... até	
Vertebrados	HQ
Zoologia	A

Naturalmente devemos incluir os sinônimos no lugar apropriado da seqüência. Por exemplo: na letra f entraria:

Fauna	H
-------	---

Deve-se considerar a ineficiência da referência de um sinônimo para outro (por exemplo: "fauna ver animais") indicando-se para cada termo sinônimo o seu correspondente único na estrutura sistemática. Outro aspecto a ser considerado seria com relação a homógrafos que devem ser distinguidos de acordo com algum artifício (por exemplo: o uso de termos compostos, "leão" e "leão marinho").

Este tipo de índice onde a cada termo de entrada corresponde uma notação é chamado índice específico. É eficiente quando um conceito aparece somente em um único lugar das tabelas, como no nosso sistema facetado, mas quando um conceito aparece em diferentes contextos, como num sistema em um sistema enumerativo é necessário que se mostre todas as notações correspondentes aquele conceito nos contextos específicos. Este tipo de índice é chamado de índice relativo. Um exemplo seria o da CDD que mostra para um único conceito diferentes abordagens, ou seja, a dispersão do conceito em várias partes do esquema.

Criança: assistência social	362.7
Criança: cuidados	649.1
Criança: medicina	618.92
Criança: psicologia	155.4

O índice relativo é construído dentro dos mesmos princípios que usamos para distinguir homógrafos, ou seja, a cada termo de entrada são acrescentados outros termos que delimitem ou definam melhor a sua acepção. Normalmente estes termos se referem a classes superordenadas que os contêm. Este procedimento é chamado “indexação em cadeia” porque é baseado na análise da cadeia da classe subordinada. Por exemplo, para a classe “Leitura na escola primária”, na CDD, teríamos:

Educação	370
Ensino	371
Ensino elementar	372
Currículo	372.19
Leitura	372.4

O que definiria as entradas:

Leitura: ensino elementar	372.4
Curriculum: ensino elementar	372.19
Disciplinas: ensino elementar	372.19
Ensino elementar	372
Ensino primário	372
Ensino	371
Educação	370

Note que entradas como Ensino Elementar: Leitura 372.4 não são feitas pois neste caso estaríamos meramente repetindo as relações anteriormente estabelecidas na própria tabela (tudo sobre ensino elementar já está em 372).

A criação de índices relativos é bem mais complexa do que de índices específicos. Isto porque haverá necessidade de se contextualizar cada conceito, e também de se elaborar toda uma estrutura sindética responsável pelo controle de termos sinônimos e relacionados.

Alguns conhecidos esquemas de classificação apresentam índices mal elaborados. Pode-se notar isto através da evolução dos índices da CDD. O índice da 17^a edição teve que ser totalmente refeito depois de já publicado o esquema. A 18^a apresenta referencias do tipo “ver” para sinônimos o que é extremamente irritante. O índice da 20^a edição já se apresenta bem melhor elaborado diminuindo bastante a complexidade da rede sindética e mostrando de forma muito mais evidente as subordinações dos conceitos.