Table of Contents

[Actividade 1 3](#_Toc316866380)

[A indexação no processo de descrição documental 3](#_Toc316866381)

[1. A importância da indexação face à sociedade da informação 3](#_Toc316866382)

[2. Objectivos da indexação 3](#_Toc316866383)

[3. Análise do conteúdo documental 3](#_Toc316866384)

[4. Fases do processo de indexação: 4](#_Toc316866385)

[5. Sistemas de indexação 5](#_Toc316866386)

[Actividade 2 7](#_Toc316866387)

[As linguagens de indexação 7](#_Toc316866388)

[I - As linguagens documentais 7](#_Toc316866389)

[Breve apontamento histórico 7](#_Toc316866390)

[Tipologia das linguagens documentais: 8](#_Toc316866391)

[Qual é a necessidade de utilizar e desenvolver linguagens documentais? 9](#_Toc316866392)

[O controlo do vocabulário 9](#_Toc316866393)

[II - Os tesauros 12](#_Toc316866394)

[Funções do tesauro 13](#_Toc316866395)

[Composição do tesauro 13](#_Toc316866396)

[Normas de apresentação dos descritores 14](#_Toc316866397)

[Formas de apresentação dos descritores 15](#_Toc316866398)

[III - Tesauro EUROVOC 17](#_Toc316866399)

[Historial 17](#_Toc316866400)

[Evolução do EUROVOC 17](#_Toc316866401)

[Conclusões 17](#_Toc316866402)

[Estrutura 17](#_Toc316866403)

[Conselhos de utilização: 18](#_Toc316866404)

[Na prática a indexação realiza-se da seguinte forma: 18](#_Toc316866405)

[A representação dos conceitos 19](#_Toc316866406)

[A selectividade da indexação: 19](#_Toc316866407)

[A exaustividade é condicionada: 19](#_Toc316866408)

[A especificidade da indexação depende, igualmente: 19](#_Toc316866409)

[Vantagens e limitações da utilização do EUROVOC na indexação por assuntos 20](#_Toc316866410)

[IV - A Classificação Decimal Universal 21](#_Toc316866411)

[Antecedentes históricos da classificação 21](#_Toc316866412)

[Características do sistema 21](#_Toc316866413)

[Princípios fundamentais 22](#_Toc316866414)

[Flexibilidade da CDU 23](#_Toc316866415)

[A ordem de citação 23](#_Toc316866416)

[Análise do tema 25](#_Toc316866417)

[Estrutura da CDU (ver anexo A) 26](#_Toc316866418)

[Tabelas de Auxiliares 26](#_Toc316866419)

[Secção I: Sinais e subdivisões de auxiliares comuns 27](#_Toc316866420)

[Auxiliares comuns independentes 27](#_Toc316866421)

[Auxiliares Comuns Dependentes 31](#_Toc316866422)

[Auxiliares especiais 32](#_Toc316866423)

[Secção II: Tabelas auxiliares especiais 32](#_Toc316866424)

[Tabelas Principais da CDU 32](#_Toc316866425)

[Vantagens e desvantagens da utilização da CDU na indexação por assuntos 33](#_Toc316866426)

[Anexo A – Estrutura da CDU 33](#_Toc316866427)

[Actividade Formativa 3 34](#_Toc316866428)

[A qualidade da indexação: factores condicionantes 34](#_Toc316866429)

[1. Qualidade da indexação 34](#_Toc316866430)

[2. Controlo do ficheiro de autoridades de assunto e a qualidade da indexação 35](#_Toc316866431)

[3. Política de indexação 35](#_Toc316866432)

[4. Manual de indexação 40](#_Toc316866433)

[Índice Remissivo 42](#_Toc316866434)

# Actividade 1

##  A indexação no processo de descrição documental

### 1. A importância da indexação face à sociedade da informação

É definida pela **NP 3715** como a acção que consiste em descrever ou caracterizar um documento relativamente ao seu conteúdo, representando esse conteúdo numa linguagem documental

### 2. Objectivos da indexação

1. Dar a conhecer o conteúdo dos documentos com o fim de informar os utilizadores;
2. Possibilitar a realização de selecções temáticas;
3. Arrumar os documentos por assuntos;
4. Recuperar os documentos por assuntos;
5. Fazer a difusão selectiva da informação.

### 3. Análise do conteúdo documental

* Nos **documentos impressos** devem ser analisados os seguintes elementos:

**T**ítulo; **R**esumo; **S**umário; **I**ntrodução; **C**onclusão; **I**lustrações, diagramas, quadros e respectivas legendas; **P**alavras ou grupos de palavras sublinhadas ou realçadas por um tipo de letra diferente; **I**nício dos capítulos e parágrafos.

* Os **documentos não escritos** (audiovisuais, visuais ou sonoros):

**E**quipamento para proceder à sua leitura; **T**reino para identificar o assunto.

**Na análise de conteúdo distinguem-se dois níveis:**

* A **condensação ou resumos** (tem carácter informativo)
* A **indexação** (indica-nos apenas o conteúdo temático)

Os **resumos analíticos**, segundo a **NP 418**, são a representação abreviada e precisa do conteúdo de um documento sem interpretação ou crítica. Têm o objectivo destacar, de forma concisa, os pontos essenciais do documento: a metodologia, os resultados e as conclusões do documento.

#### Informativo

* A informação quantitativa e/ou qualitativa contida no documento, numa sequência de frases interligadas; Útil em textos que descrevem um trabalho experimental e consagrados a um único tema.

#### Indicativo

* Apresenta a informação como guia indicativo ou descritivo do tipo de documentos, dos principais assuntos focados e da maneira como são tratados; É utilizado em textos discursivos ou extensos, como revisão de conjunto, críticas e monografias.

#### Informativo – indicativo

* Combina as duas formas, sendo os aspectos essenciais abordados de forma informativa e os restantes de forma indicativa.

### 4. Fases do processo de indexação:

A **NP 4285-4** define linguagem documental como uma linguagem formal utilizada para caracterizar os dados ou o conteúdo de documentos e permitir o seu armazenamento e recuperação. Intervêm na terceira fase do processo de indexação.

1. Análise e definição do conteúdo do documento;
2. Identificação e selecção dos conceitos representativos do conteúdo;
3. Tradução dos conceitos em termos de linguagem documental, através dos instrumentos de indexação.

#### 1ª Fase Analisada anteriormente, no [ponto 3](#_3._Análise_do).

#### 2ª Fase – Identificação e selecção dos conceitos compreende:

* Identificação sistemática dos conceitos representativos do conteúdo;
* Selecção dos conceitos.

##### A exaustividade

* Número de conceitos considerados e que caracterizam o conteúdo integral do documento;
* O critério principal na selecção dos conceitos deve ser sempre o seu valor potencial, na expressão do conteúdo de um documento e na sua recuperação.

##### A especificidade

* Exactidão com que um conceito é representado por um termo de indexação.

#### 3ª Fase – Representação dos conceitos por termos de indexação:

* O conceito já está representado por um termo de indexação que deve permanecer;
* O conceito é novo e por isso temos que seleccionar o termo de indexação que o deverá representar, respeitando as regras e princípios da linguagem documental seleccionada.

**Linguagem documental** é uma linguagem artificial, convencional e controlada, utilizada para descrever o conteúdo dos documentos, tendo em vista o armazenamento da informação.

### 5. Sistemas de indexação

Os serviços de informação podem optar por dois tipos de sistemas de indexação:

* Sistemas de indexação pós-coordenados e pré-coordenados.

Estes sistemas diferem em três aspectos:

* **Quem** faz a coordenação?
* **Quando** é que a coordenação é feita?
* **Como** é que os termos de indexação são coordenados?

**Quem faz a coordenação?**

* Num sistema pré-coordenado é o indexador que combina os termos de indexação e cria os cabeçalhos de assunto ou a notação.
* Num sistema pós-coordenado é o utilizador que constrói a coordenação, no momento em que combina os termos.

**Quando é que a coordenação é feita?**

* No sistema pré-coordenado a coordenação dos termos de indexação é feita antes de se iniciar a pesquisa.
* No sistema pós-coordenado os termos de indexação são coordenados no momento da pesquisa.

**Como é que os termos são coordenados?**

* No sistema pré-coordenado os termos de indexação são coordenados pelo indexador de acordo com um conjunto de regras de sintaxe definidas para a construção desses termos. A relação entre os termos de indexação é gramatical.
* No sistema pós-coordenado, do ponto de vista do utilizador, a pesquisa implica a coordenação dos termos utilizando uma “sintaxe booleana”. Do ponto de vista do indexador, tem pouca ou nenhuma sintaxe (o seu vocabulário consiste apenas em termos simples). A relação entre os termos é uma relação lógica.

**Características do sistema pré-coordenado:**

* Garante a integridade da informação;
* Representa os assuntos complexos como tal e não através de vários conceitos sem qualquer relação explícita (tratados como assuntos independentes);
* Representa o assunto em toda a sua especificidade;
* Assegura a recuperação de cada termo no contexto em que é pertinente;
* A linguagem utilizada na representação dos conceitos utiliza vários termos que é preciso combinar;
* A combinação é feita pelo bibliotecário no momento da indexação, auxiliado por normas;
* O vocabulário é construído *a posteriori*;
* Utiliza cabeçalhos de assunto ou notações como termos de indexação;
* Adequam-se a bibliotecas gerais e especializadas;
* Utiliza uma linguagem flexível;
* A linguagem utilizada é mais específica;
* A informatização potencia as vantagens da linguagem utilizada, pois permite flexibilizar a pesquisa;
* Tem custos maiores para os serviços e custos menores para os utilizadores.

**Características do sistema pós-coordenado:**

* A combinação dos termos faz-se pelo utilizador no momento da pesquisa da informação, utilizando os operadores booleanos;
* O vocabulário utilizado é construído *a priori*, isto é, antes de se proceder a qualquer ato de indexação;
* Utiliza descritores como termos de indexação;
* Adequam-se a serviços especializados;
* A informação recuperada não é específica, nem coextensiva com o conteúdo da obra;
* O indexador despende um tempo menor no processo de indexação;
* Tem custos menores para o serviço e maiores para o utilizador.

Seja qual for a opção do gestor do serviço de informação, o objectivo é conseguir as duas qualidades que a pesquisa de informação deve possuir: **a precisão e a exaustividade**, ou seja que abarque **todos** os documentos relevantes **apenas** para aquele assunto.

# Actividade 2

## As linguagens de indexação

Nas **linguagens vocabulares**, analisamos os tesauros, a sua composição, formas de apresentação e a importância do controlo do vocabulário.

Nas **linguagens classificatórias** destacamos a CDU, pela implantação que tem e por ser um elemento imprescindível para se fazer parte da PORBASE.

## I - As linguagens documentais

As três características destas linguagens:

* Artificiais
* Controladas
* Convencionais
1. **Artificial** porque utiliza signos alfabéticos ou numéricos para representarem conceitos segundo regras próprias e dentro de um contexto restrito.
2. **Controladas** porque foram desenvolvidas, para permitirem a representação de um conceito de forma unívoca, recorrendo sempre ao mesmo termo de indexação, de modo a que a ambiguidade seja minimizada ou mesmo anulada.
3. **Convencionais** porque utilizam signos ou códigos (alfabéticos, numéricos ou alfanuméricos) cuja utilidade reside na univocidade de conteúdo. O que só é possível, respeitando as normas e as regras que permitem a cooperação entre bibliotecas.

### Breve apontamento histórico

As primeiras linguagens documentais desenvolvidas foram as ***classificações universais*** nos finais do século XIX. Destaca-se a CDD e a CDU. São ambas classificações de **tipo hierárquico**, que têm uma organização sistemática do conhecimento. A CDD tem a característica **enumerativa** e a CDU evoluí como uma classificação **sintética**, com vantagens na cobertura de assuntos e na sua combinação. A cadeia notacional que construímos quando classificamos um assunto corresponde a um **termo de indexação pré-coordenado.** Cada elemento tem o mesmo significado.

São **tipificadas** quanto à **abrangência** temática (universais, gerais ou especializadas), à **organização e estrutura** (enumerativas, hierárquicas, de facetas). São **pré-coordenadas**.

As **linguagens vocabulares ou terminológicas**, também designadas por **combinatórias** surgem, também, nos finais do século XIX, com a utilização da primeira lista de cabeçalhos de assunto ou lista de encabeçamento de matérias. Seguem os princípios da **pré-coordenação**, da especificidade e da entrada directa.

Os **tesauros** (linguagens vocabulares) aparecem apenas no século XX e, uma vez que utilizavam uma linguagem controlada próxima da linguagem natural, mais próximos de satisfazer as necessidades de informação por permitirem um tratamento dos conteúdos temáticos de um modo mais específico, ao nível do conceito, sem que se perdesse a rede de relações semânticas. Adequava-se a áreas especializadas, eram mais fáceis de utilizar, exigiam menos tempo ao indexador e eram mais rápidos na recuperação da informação. Seguem os princípios da **pós-coordenação**.

Os cabeçalhos de assunto são **linguagens terminológicas** dispendiosas. Exigem muito tempo por parte do indexador (mas garantem a integridade e a especificidade da informação), possibilitam uma maior precisão na descrição dos assuntos complexos e permitem uma percepção mais imediata, por parte do utilizador. A informática, nos meados do século XX, não tinha capacidade para potencializar as pesquisas por assunto, utilizando as linguagens de cabeçalhos de assunto. Esta incompatibilidade, associada aos custos do tratamento da informação, fizeram cair em desuso estas linguagens. Surge a era do tesauro, enquanto linguagem documental.

Na última década do século XX, com o aumento de recursos digitais disponíveis e com os desenvolvimentos da Web semântica surgiu “A nova geração de sistemas de organização do conhecimento”:

* As **ontologias** – Formas de representação de um domínio do conhecimento.
* **Mapas conceptuais** – A informação, os conceitos e as suas relações são representados na forma de diagramas ou mapas. Proporciona a representação estruturada do conhecimento em forma de rede – redes semânticas.
* **Taxonomias** – constituem um esquema ordenado do conhecimento em várias áreas, segundo determinadas regras ou normas.

### Tipologia das linguagens documentais:

Critérios de tipificação das linguagens documentais mais conhecidos:

* Os do **controlo**;
* Os da **coordenação dos termos**;
* O **estrutural**.

Do controlo exercido sobre o vocabulário, as linguagens podem organizar-se em duas categorias:

* **Livres** - Por exemplo: Listas de termos, Terminologias, etc.;
* **Controladas** – Classificações e tesauros. São vocabulários construídos *a priori*.

O critério **de coordenação dos termos** está dependente do momento em que se realiza a combinação dos vários elementos que compõem uma linguagem documental.

* Na elaboração da linguagem ou no momento da descrição, é **pré-coordenada**;
* No momento da recuperação, a linguagem é **pós-coordenada**.

Exemplificando: classificações e cabeçalhos de assunto são linguagens pré-coordenadas e os tesauros são linguagens pós-coordenadas.

Nos sistemas **pré-coordenados**, os termos de indexação (cabeçalhos de assunto) combinam-se de acordo com regras sintácticas no momento da indexação.

Nas linguagens **pós-coordenados** a sintaxe é booleana e os termos simples ligam-se no momento da pesquisa.

Quanto ao critério da estrutura, identificam-se dois tipos de linguagens:

* **Hierárquica** ou categorial cuja estrutura estabelece que no conceito ou categoria de assunto quanto maior for a especificidade do conceito maior é o número de dígitos que compõem o código representado por um código numérico. A hierarquia conceptual reflecte a hierarquia numérica;
* **Combinatória** onde podemos encontrar os tesauros e os cabeçalhos de assunto

### Qual é a necessidade de utilizar e desenvolver linguagens documentais?

Para obviar as dificuldades no uso da linguagem natural e da inerente ambiguidade.

### O controlo do vocabulário

A **NP 4036** define linguagem de indexação como o conjunto controlado de termos escolhidos numa linguagem natural e utilizados para representar sob forma coordenada o conteúdo dos documentos.

Obstáculos linguísticos comuns da linguagem natural:

#### A sinonímia

#### A polissemia

A fronteira entre a polissemia e a homonímia nem sempre é clara.

#### A homonímia

##### Homógrafas e Homófonas

#### Singular e plural

O vocabulário controlado deve estabelecer os termos aceites e contemplar a forma que foi rejeitada:

Aluno (termo não aceite) < - > USE Alunos

Termos que representam entidades enumerativas, como nomes ou objectos aos quais se pode aplicar a pergunta **QUANTOS**? Devem ser expressos no **plural:**

Quantos documentos? Quantos alunos? Quantos ovos?

Termos que representam entidades não contáveis como nomes de materiais ou de substâncias às quais se pode aplicar a pergunta **QUANTO**? Devem ser expressos no **singular**:

Quanto café? Quanto ferro?

**Devem ser expressas no singular os nomes de entidades abstractas como:**

* Fenómenos (personalidade, inverno);
* Propriedades (fragilidade, opacidade);
* Religiões (Catolicismo, Hinduísmo);
* Actividades (corte; emigração, respiração);
* Disciplinas (Física; Psicologia);

O indexador estabelece o **controlo do vocabulário** nas linguagens documentais a dois níveis:

1. Controlo formal que inclui o controlo morfológico e o controlo sintáctico;
2. Controlo semântico que abarca as relações semânticas; as notas explicativas e as definições.

#### 1. Controlo formal

Tem a ver com a normalização das unidades lexicais relativamente a três aspectos: *escolha da língua*, *o género* e o *número* em que se devem expressar”.

“No que concerne ao *aspecto sintáctico,* este prende-se com a normalização da forma dos termos, em particular com a estrutura dos termos compostos que se usam no tesauro.

#### 2. Controlo semântico

O objectivo é a normalização conceptual de modo a evitar a ambiguidade semântica que gera ruído e silêncio na recuperação da informação. Analisa-se o controlo semântico seguindo duas vias:

* ***Via significante*** *-* devemos atender às questões da polissemia, da homonímia, dos qualificadores e definições.
* ***Via significado*** *-* Aspectos da sinonímia e relações de equivalência

##### As relações semânticas são de tês tipos:

**1. Relações de equivalência** têm a função de controlarem a sinonímia:

* Termos de origem linguística diferente (lanchar e merendar);
* Termos antigos e actuais (aeroplano e avião);
* Termos com grafias diferentes (ouro e oiro)
* Termos equivalentes em línguas diferentes (abstract e resumos)
* Termos populares e termos científicos (aspirina e ácido acetilsalicílico);
* Termos de uso geral e de uso local (conchas e caços)
* Nomes comuns e nomes de marcas (garrafa isoladora e garrafa de termos)
* Entre siglas, acrónimos e o seu desenvolvimento (UE e União Europeia)

***Estas relações são assimétricas e irreversíveis***, na medida em que um termo preferencial ou aceite está ligado a um termo não preferencial e um termo não preferencial deverá estar ligado, obrigatoriamente, ao termo preferencial:

Doentes mentais (termo aceite)

UP Doentes psiquiátricos (termo não aceite)

Doentes psiquiátricos

USE Doentes mentais

**2. Relações hierárquicas** baseiam-se em graus ou níveis de superioridade ou de subordinação. ***Estas relações são recíprocas e assimétricas***. Existem três tipos de relações:

***Relação genérica***identifica a posição que os membros de uma classe ou categoria têm dentro dela. O descritor de nível inferior é específico do descritor de nível superior:

TG Pássaros

TE Papagaios

Papagaios

TG Pássaros

***Relação partitiva ou todo/parte*** - cobre uma gama limitada de situações nas quais a parte está implícita no todo. Os termos podem ser organizados hierarquicamente, servindo o todo de termo superior e a parte de termo subordinado:

TG Sistema cardiovascular

TE Artérias

Artérias

TG Sistema cardiovascular

***Relação de instância***identifica a relação que existe entre uma categoria geral de coisas ou acontecimentos, expressas por um nome comum e um espécime individual dessa categoria, que forma uma classe de um só elemento e é representado por um nome próprio:

TG Regiões montanhosas

TE Himalaias

Himalaias

TG Regiões montanhosas

**3. Relações associativas** incluem relações entre pares de termos que não fazem parte de um mesmo conjunto de equivalências. Os termos associados não fazem parte da mesma cadeia hierárquica, mas estão mentalmente associados e a relação entre os termos deve ser especificada. ***Estas relações são recíprocas e simétricas****,* uma vez que quando se associa um termo a outro, devemos obrigatoriamente fazer o inverso:

Doentes mentais

TR Hospitais psiquiátricos

Hospitais psiquiátricos

TR Doentes mentais

## II - Os tesauros

A sua expansão foi rápida, pois contribuíram de um modo muito eficaz para o controlo das terminologias em áreas disciplinares específicas, para o processamento mais rápido da informação e para a cooperação e a interoperabilidade dos serviços.

A **NP 4285-4** define tesauro como o “vocabulário controlado de termos com relações semânticas abrangendo um ou vários domínios particulares do conhecimento”.

A **NP 4036** define “tesauro: vocabulário de uma linguagem de indexação controlada, organizado formalmente de maneira a explicitar as relações estabelecidas *a priori* entre os conceitos.”

Os tesauros são construídos segundo o princípio combinatório, baseado na combinação de termos, utilizando os operadores booleanos quando se realiza a pesquisa.

O tesauro é uma lista normalizada e estruturada de termos aceites para a indexação de documentos. Os termos são combinados entre si, de modo a descreverem os assuntos dos documentos e a responderem às questões colocadas. Os tesauros são instrumentos muito úteis para a indexação pois fornecem uma visão estruturada do vocabulário de uma disciplina. Esta é uma das suas potencialidades e que seria totalmente anulada se pretendêssemos criar tesauros generalistas.

Existem tesauros de âmbito alargado, de tipo enciclopédico, como o **Macrothesaurus** da OCDE.

O EUROVOC, criado por e para os serviços de informação do Parlamento Europeu, abarca todos os assuntos ligados aos domínios políticos e jurídicos na óptica dos parlamentares.

Os tesauros são instrumentos de indexação e de pesquisa, comuns ao indexador e ao utilizador, permitindo estabelecer a relação entre a pergunta e a informação sobre o assunto X.

Um tesauro caracteriza-se pela natureza e o número de relações existentes entre os termos. Quando se pretende estabelecer um novo tesauro, começamos por:

* Recolher vocabulário relativo ao assunto em causa;
* Reunir uma lista de termos, apelidados de candidatos a descritores;
* Com base nessa lista, seleccionar os descritores e não descritores. Os termos são examinados partindo do princípio que os tesauros só devem autorizar os descritores necessários e suficientes para cobrir aquele domínio concreto. É o ponto de partida para a normalização e o controlo do vocabulário coligido;
* A cada conceito corresponde um único descritor

Na elaboração dos thesauri ou tesauros, o [controlo semântico](#_2._Controlo_semântico) evita a ambiguidade semântica que pode provocar ruído ou silêncio no acto da pesquisa de documentos. O controlo semântico, através de várias estratégias, como a utilização de **qualificadores**, de notas explicativas e o recurso às **relações semânticas** (relações de equivalência, relações hierárquicas e relações associativas), faz o **controlo da** [**polissemia**](#_O_controlo_do) e o **controlo da sinonímia**, com o objectivo de assegurar que cada termo tem um único significado.

### Funções do tesauro

1. Normalização do vocabulário;

2. Indução;

3. Representação

A **normalização do vocabulário** tem como objectivo unificar todo o léxico do tesauro de modo a controlar todas as possíveis entradas. Controlo formal e semântico dos termos.

A **indução** serve para que o tesauro indique sistematicamente todas as alternativas possíveis de entradas que permitam recuperar informação pertinente, para uma dada pesquisa.

A **representação** estabelece a correspondência entre os descritores (termos de indexação de um tesauro) e os conceitos presentes nos documentos.

### Composição do tesauro

* **Unidades lexicais:**
	+ Descritores;
	+ Não-descritores ou termos equivalentes;
* [**Relações semânticas**](#_2._Controlo_semântico)**:**
	+ Relações de equivalência;
	+ Relações hierárquicas;
	+ Relações associativas

**Os descritores podem ser tipificados do seguinte modo**:

* Quanto à sua **carga informativa** podem ser:
	+ *Primários* representam um conceito de maneira unívoca, por um único termo;
	+ *Secundários com* a representação feita por termos compostos
* Quanto ao seu **conteúdo** podem ser:
	+ *Onomásticos*;
	+ *Geográficos* (conceitos de âmbito geográfico);
	+ *Materiais ou temáticos* (representam conceitos muito diferentes e por isso são os mais difíceis de controlar. Cobrem entidades concretas e abstractas);
	+ *Cronológicos*
* Quanto à **composição** podem ser:
	+ *Descritores simples* representam conceitos com uma única palavra);
	+ *Descritores compostos* (representam o conceito utilizando:
		- Um sintagma nominal [nome + adjectivo] ou preposicional [conjunto de termos unidos por nexos gramaticais].

### Normas de apresentação dos descritores

A **NP 4036** estabelece algumas considerações essenciais sobre a morfologia dos descritores:

* Os **descritores** devem ser impressos em maiúsculas;
* Os **não descritores** devem ser impressos em minúsculas;
* **Nomes e expressões nominais**Os termos de indexação devem ser expressos na forma de nome (substantivo) ou expressão nominal. Estas apresentam-se sob duas formas:
	+ **Expressões adjectivas** [nome + adjectivo]
	+ **Expressões prepositivas** [nome+ preposição + nome]
* **Adjectivos** utilizados isoladamente podem ocorrer em situações excepcionais;
* **Advérbios** como “mais” ou “muito” nunca devem ser utilizados;
* **Verbos** no infinito ou no particípio não devem ser utilizados isoladamente como termos de indexação. As actividades devem ser representadas por nomes ou expressões verbais;
* **Abreviaturas ou acrónimos** não devem ser utilizadas;
* **Escolha da forma singular ou plural** quando existe distinção a decisão pode ser determinada pelos seguintes factores:
1. *Indexação pré-coordenada ou pós-coordenada*;
2. *Factores culturais*

Nos organismos onde se pode adoptar indiferentemente, a escolha entre as duas formas depende, em geral, do tipo de noção que ele exprime. A distinção deve ser indicada através de um **qualificador**:

MEMÓRIA (processo mental) / MEMÓRIAS (género literário)

* **Nomes de lugar** Variam de uma para outra língua e, por vezes existem variantes na mesma comunidade linguística:
1. Um nome oficial e uma variante corrente são empregues indiferentemente:

HOLANDA <-> PAÍSES BAIXOS

1. Coexistem a forma original e tradução

LEGHORNO <-> LIVORNO

Opta-se pelo nome mais familiar. Estabelecem-se remissivas recíprocas:

PAÍSES BAIXOS

UP Holanda

Holanda

USE PAÍSES BAIXOS

* **Nomes próprios de instituições ou de pessoas** são excluídos.
* **Notas explicativas e definições** Podem acompanhar o termo para limitar e precisar o sentido em que este deve ser utilizado:
1. *Data de adopção de um termo*;
2. *Fonte do termo*;
3. *Instruções de aplicação do termo* para os indexadores: não faz parte do termo e é registada num campo específico do registo de autoridade;
* **Apresentação dos termos e suas relações.** Existem várias formas de os apresentar. Esta Norma limita-se à descrição de três métodos de uso corrente:
	1. ***Apresentação alfabética***; contendo as notas explicativas e indicação das relações entre os termos;
	2. ***Apresentação sistemática***, acompanhada de um índice alfabético;
	3. ***Apresentação gráfica*,** acompanhada de um índice alfabético.

### Formas de apresentação dos descritores

Formas de apresentação dos termos e suas relações, segundo as orientações da **NP 4036:**

#### I) Apresentação alfabética

Todos os termos, descritores e não-descritores, estão organizados numa sequência alfabética única. Os não descritores são essencialmente acompanhados pela referência **USE**. As informações relativas a um descritor deverão ser enunciadas pela seguinte ordem:

1) **NE** ou (SN) notas explicativas ou definições

2) **UP** ou (UF) indicação de não-descritor

3) **TT** indicação do termo de topo, se necessário

4) **TG** ou (BT) indicação dos termos genéricos

5) **TE** ou (NT) indicação dos termos específicos

6) **TR** ou (RT) indicação dos termos relacionados

#### II) Apresentação sistemática

Um tesauro com os termos organizados sistematicamente deve conter duas partes:

1. Categorias ou hierarquias de termos dispostos segundo o seu significado e as relações lógicas;
2. Um índice alfabético que remete o utilizador para o local apropriado da apresentação sistemática.

A ligação entre as duas partes é feita por um sistema de endereços. É atribuído, a cada um dos descritores da parte sistemática, um código que funciona como uma referência no índice alfabético.

A parte sistemática é muitas vezes considerada como a parte principal do tesauro, a que contém a maior quantidade de informação definidora e relacional. O índice alfabético tem o papel de um componente complementar mas secundário.

**Organização primária de um tesauro sistemático**

Atender não apenas às relações entre os termos, mas também às relações entre as hierarquias ou categorias. É necessário uma estrutura “chapéu” ou uma macroclassificação à parte sistemática para assegurar que as noções similares estão reunidas e as noções sem relação estão separadas, em benefício do utilizador. Pode obedecer a duas formas:

* 1. Organização em domínios ou disciplinas;
	2. Organização em facetas;
	3. Combinação dos dois tipos de organização

##### a) Organização em domínios

Organiza-se o universo dos conhecimentos em classes principais ou disciplinas.

##### b) Organização por facetas

A noção é mais abstracta e a sua organização requer uma disciplina mental mais rigorosa.

Os termos são organizados em classes ou conjuntos sem ter em conta o ou os domínios.

##### c) Combinação dos dois tipos de organização

Na prática, os dois tipos de organização são combinados.

#### III) Apresentação gráfica

Os termos de indexação e as suas relações são dispostos numa figura a duas dimensões que permite ao indexador ou utilizador dispor de toda uma gama de termos e suas relações. Podem identificar-se dois tipos principais:

1. Estrutura arborescente (NP 4036);
2. Esquema em flecha (NP 4036)

Um tesauro que inclua uma **apresentação gráfica** deve comportar duas partes:

##### 1) A apresentação gráfica

Limitada aos descritores, pelo formato que não se adapta facilmente às notas explicativas, sinónimos, etc. É identificada por um símbolo, um número ou um elemento de notação hierarquicamente significativo.

##### 2) Índice alfabético

Contém as notas explicativas e as relações de equivalência e pode também incluir as relações hierárquicas e associativas.

O índice da estrutura arborescente não indica relações hierárquicas. O índice do esquema em flecha inclui não só notas explicativas, mas também as relações hierárquicas.

Quando a parte alfabética contém, em proporção, maior número de informações (definidoras e relacionais), funciona como a parte principal do tesauro.

## III - Tesauro EUROVOC

O Serviço das Publicações Oficiais das Comunidades Europeias e o Parlamento Europeu optaram pela elaboração de um tesauro multilingue que abrangesse todos os domínios de actividade das Comunidades Europeias.

### Historial

O trabalho de selecção dos descritores e estruturação do tesauro foi iniciado em 1982.

A primeira edição foi publicada em 1984, em dois volumes (apresentação alfabética e temática) e em sete línguas, tendo sido utilizada pelo Parlamento Europeu e pelo EUR-OP. A experiência dessa utilização permitiu uma revisão do tesauro, do qual foi publicada, em 1987, uma segunda edição, disponível em nove línguas. Esta edição foi alvo de interesse fora das instituições comunitárias.

### Evolução do EUROVOC

Estes centros de documentação reconheceram as qualidades do EUROVOC para o tratamento da documentação comunitária e parlamentar. Além disso, ao adoptarem o EUROVOC, podiam dispor desta última já indexada.

O Centro Europeu de Investigação e Documentação Parlamentares (CERDP) estabelece o contacto entre os serviços de documentação de todos os parlamentos e membros do Conselho da Europa. Um dos seus objectivos consiste em facilitar o intercâmbio de informação entre os parlamentos. Foi neste contexto que a Assembleia da República portuguesa decidiu utilizar o EUROVOC. O primeiro parlamento membro do conselho da Europa a traduzir o EUROVOC para a sua língua foi o da Turquia.

### Conclusões

Um tesauro é matéria viva, que evolui em função das necessidades dos indexadores e utilizadores de fundos documentais e de bases de dados cada vez mais numerosas.

### Estrutura

Encontra-se estruturado de modo genérico, segundo uma classificação em dois níveis:

* **Domínios**, número de dois algarismos e impressos em maiúsculas.

Abrange todas as actividades das instituições europeias;

* **Microtesauros**, número de quatro algarismos, dos quais os dois primeiros são os do domínio a que pertence o microtesauro, e por um enunciado; em minúsculas.

Encontra-se estruturado, ao nível específico dos descritores e não descritores, por relações semânticas:

* Relações de pertença ao microtesauro **(MT**);
* Relação de equivalência **(UF)** ou usado por **(UP);**
* Relação hierárquica: **(BT)** ou termo genérico **(TG) e (NT)** ou termo específico **(TE);**
* Relação associativa **(RT)**
* Notas de aplicação ou de definição **(SN)**

**Os descritores são impressos a negrito.**

### Conselhos de utilização:

A indexação é a operação que consiste em reconhecer e seleccionar os conceitos de que trata um documento e em representá-los por descritores do tesauro.

A sua finalidade é possibilitar a extracção, de entre uma colecção de referências bibliográficas, as que responderão às questões colocadas; esta extracção só é possível se os conceitos da questão estiverem na mesma linguagem: a dos descritores do tesauro.

### Na prática a indexação realiza-se da seguinte forma:

* Rápida tomada de conhecimento do conteúdo do documento a indexar;
* Selecção dos conceitos a utilizar para a indexação do documento:
	+ **A selectividade**: só escolher os conceitos em relação aos quais o documento fornece uma informação de interesse aos utilizadores;
	+ **A exaustividade**: devem ser escolhidos todos os conceitos úteis que figuram no texto, quer apareçam de forma explícita, quer implícita;
* Representação dos conceitos seleccionados por descritores do tesauro:
	+ Se em língua estrangeira, tradução dos conceitos;
	+ Procura na apresentação alfabética permutada dos termos.

Existem três possibilidades:

1. A expressão do conceito corresponde a um descritor que o designa;
2. Corresponde a um não descritor que o designa: este último remete para o descritor a utilizar (“subsídio escolar”, que nos envia para “subsídios de estudo”),
3. Não existe uma entrada correspondente à expressão do conceito. Neste caso, podem ser usados quatro métodos:
* Consultar a apresentação alfabética permutada sobre cada uma das palavras do conceito;
* Conceber uma outra formulação do conceito em linguagem natural e verificar se a mesma corresponde a uma entrada da apresentação alfabética permutada;
* Definir a ou as classes gerais (microtesauros) que incluem o conceito, consultar a apresentação temática e procurar os que melhor representam o conceito;
* Combinar dois ou mais descritores da apresentação alfabética permutada para expressar, por uma perífrase, o conceito a representar.

Deve-se ter em conta o respectivo sentido real, e não o sentido literal.

### A representação dos conceitos

* **A especificidade vertical**: o descritor deve situar-se no mesmo nível de especificidade do conceito ou, no caso de aquele não existir, no nível imediatamente superior no tesauro.
* **A especificidade horizontal**: se para um conceito composto existe no tesauro um descritor composto.

A correcta aplicação das regras de selectividade, exaustividade e especificidade é essencial na qualidade de um sistema de informação documental.

### A selectividade da indexação:

* É independente da qualidade do tesauro,
* Depende da política de indexação aplicada por cada sistema documental e, nomeadamente, da profundidade da indexação,
* Depende do conhecimento que o indexador possui dos utilizadores do sistema documental, bem como daquilo que lhes interessa.

### A exaustividade é condicionada:

* Pela riqueza do tesauro: só podem indexar-se, obviamente, os conceitos presentes na linguagem documental,
* Pela política de indexação, que fixa a profundidade da mesma,
* Pelo comportamento do documentalista e respectiva meticulosidade.

**Um aumento da exaustividade da indexação** origina, na fase da pesquisa documental:

* Uma melhoria da taxa de recuperação;
* Mas uma diminuição da precisão.

### A especificidade da indexação depende, igualmente:

* Do tesauro, que inclui ou não um número importante de descritores específicos
* Da política de indexação,
* Dos documentalistas.

Ao contrário do que acontece com a **exaustividade**, o aumento da especificidade provoca:

* Uma diminuição da taxa de recuperação
* Mas uma melhoria da precisão.

Observe-se, no entanto, que os documentos indexados com um nível mais genérico podem ser encontrados, precisamente através de uma pesquisa baseada nesses descritores. Infelizmente, o inverso não se verifica, ou seja, se a indexação for pouco específica, obtém-se:

* Melhor taxa de recuperação,
* Mas menor precisão.

É por este motivo que convém ser específico quando se procede à indexação.

### Vantagens e limitações da utilização do EUROVOC na indexação por assuntos

**Vantagens:**

* Normalização terminológica dos vocabulários de indexação dos documentos e de formulação das questões aquando da interrogação dessa mesma base, permitindo uma grande eficácia da pesquisa documental;
* Multilinguismo, permitindo indexar documentos na língua do documentalista e procurá-los na língua do utilizador;
* Possibilidade de estabelecer redes de cooperação entre serviços documentais.

**Limitações:**

* Foi construído para satisfazer as necessidades de sistemas documentais gerais no que se refere às actividades comunitárias, não sendo adequado para documentos especializados [nem para bibliotecas generalistas];
* Por outro lado, embora se desenvolvam esforços no sentido de ter em conta as necessidades dos utilizadores não pertencentes às instituições comunitárias, não pode ter a pretensão de abranger as diferentes realidades nacionais com suficiente especificidade.

## IV - A Classificação Decimal Universal

### Antecedentes históricos da classificação

Em 1895 Paul Otlet e Henry La Fontaine começaram a compilar este no Repertório Bibliográfico Universal e decidem organizá-lo de uma forma sistemática. Tal tarefa exigia uma classificação bibliográfica adequada e seleccionaram a *CDD*. A classificação possuía apenas alguns milhares de subdivisões, mas a sua notação apresentava a possibilidade de utilização universal já que abarcava todo o universo do conhecimento e empregava os números arábicos.

Acrescentaram uns quantos instrumentos de síntese e tabelas auxiliares que a tornaram mais flexível e detalhada. Estes elementos vieram a alterar a estrutura inicial da CCD, totalmente enumerativa, para uma estrutura híbrida. A estrutura da CDU possui características enumerativas, sintéticas e facetadas.

Em 1992 a responsabilidade de gerir e actualizar a CDU passa para o *Consórcio da CDU.* Uma das suas primeiras actividades, foi a criação de uma base de dados que constituiu a versão autorizada do sistema, ou seja, o **Master Reference File** (MRF). Esta base é actualizada todos os anos, após a publicação da lista anual de correcções autorizadas que se designa por *Extensions and Corrections to the UDC*.

### Características do sistema

**Classificação Decimal Universal**

***Classificação***

Desde sempre o Homem teve necessidade de classificar e ordenar tudo o que o rodeia. Classificar é um método de organização em categorias ou classes e pode ser empregue em múltiplas tarefas quotidianas.

A classificação tem por base o princípio da ordenação de conceitos de acordo com o seu grau de semelhança.

Esta perspectiva analítica da classificação é essencial para que o sistema seja coerente. A soma total dos termos derivados da aplicação de uma única característica de divisão designa-se como **faceta**.

As classificações sistematizaram-se tendo por base as classes. Podemos definir **classes** como um *conjunto cujos membros têm algo em comum.* Pode ser simples ou composta. Uma classe composta é constituída pela intersecção de dois ou mais tipos diferentes de conceito (ou faceta) dentro da mesma classe, tais como fisiologia vegetal, patologia vegetal, ecologia vegetal, etc.

***Decimal***

O número da classe é o código utilizado para representar a classe/conceito e determina a disposição dessa classe no sistema. Pode ser formado por um único algarismo, por uma sequência de algarismos árabes ou ainda por uma sequência que combine algarismos árabes e símbolos autorizados.

A numeração CDU é baseada em algarismos árabes ordenados de acordo com o sistema decimal. Isto significa que o universo do conhecimento é visto como a unidade e cada área temática/disciplina/classe é uma fracção desse todo.

A numeração aumenta por fracções decimais e a extensão reflecte o nível hierárquico.

***Universal***

A universalidade da CDU verifica-se duplamente:

* Abarca todos os domínios do conhecimento num único sistema classificatório.
* Pode ser utilizada em qualquer parte do mundo, independentemente da língua. O recurso aos algarismos árabes é um dos pontos fortes do sistema.

### Princípios fundamentais

A CDU é uma *classificação por aspecto*, como todos os outros sistemas gerais. Significa que um fenómeno/assunto pode ocorrer em mais do que uma classe, como por exemplo: ovos. Podem ser encontrados em ornitologia (classe 5), em culinária (classe 6), etc.

**A CDU é uma classificação que recorre às *facetas***(características comuns a várias categorias) mas não é uma classificação verdadeiramente facetada, pois estas enumeram apenas termos simples nas suas listas. Embora não sendo totalmente facetada nesse sentido, os princípios estão inerentes na estrutura da CDU. Existem muitas situações em que um termo composto é mais adequado do que a combinação de termos simples;

**A CDU é uma *classificação hierárquica***, significa que cada classe principal pode ser subdividida e cada subdivisão pode ser de novo subdividida. A hierarquia numérica reflecte a hierarquia conceptual.

**A CDU é uma *classificação analítico-sintética***, isto implica a ocorrência de um processo de análise e de outro de síntese, por outras palavras, implica que o classificador decomponha os conceitos em elementos simples (analisados), que são depois reunidos (sintetizados) na notação que melhor os representa.

**Termos simples, compostos e complexos**

Os termos simples são representados por números simples, normalmente das tabelas principais.

Os termos compostos são representados no esquema pela utilização de subdivisões auxiliares especiais.

Um termo complexo é habitualmente representado, sem o ser exclusivamente, por “:”.

### Flexibilidade da CDU

A flexibilidade da CDU é uma das características mais notáveis do sistema. Permite uma classificação dos assuntos de uma forma genérica ou mais específica de acordo com as necessidades dos serviços. O detalhe e o potencial para combinações de conceitos dão esse espaço. As várias partes do sistema podem ser citadas de múltiplas maneiras de acordo com as especificidades de cada serviço.

O técnico deve clarificar as suas necessidades, antes de iniciar o processo de classificação e deve manter um registo das decisões locais tomadas para a sua aplicação, de maneira a que seja seguido consistentemente o mesmo método. A consistência é essencial. O estabelecimento e a adesão a uma determinada ordem de citação ajuda tanto na disposição em prateleiras como na criação de uma sequência classificada, quer seja on-line ou impressa.

### A ordem de citação

Deve reflectir a forma como a maioria do material se encontra ordenado. A combinação dos termos isolados resulta da reunião de toda a informação sobre o assunto, tida como de primordial interesse e na dispersão de outros conceitos. Esta decisão depende do fim para o qual se destina a classificação. Para a disposição em prateleiras e para arquivo, são problemas de grande importância, e por isso a ordem escolhida deve ser consistente. Isto continua a ser válido mesmo ao utilizar um catálogo on-line dado ser extremamente útil a possibilidade de pesquisar uma sequência classificada para baixo e para cima. A seguir são indicadas as **categorias** que ocorrem mais frequentemente em documentos:

**Subdivisões comuns**

Forma do documento

Época. Aspectos históricos

Espaço. Subdivisões de local

Pesquisa, educação, comunicação e informação no tema; aspectos administrativos e legais

Influência de/Relações com outras áreas

Princípios e teoria

**Agentes**

**P**essoas e organizações; **I**nstalações e equipamento

**Acções**

**P**rocessos; **O**perações

**Características**

**M**ateriais; **P**artes; **E**ntidades completas; **S**istemas de entidades

É a ordem pela qual se necessita mais frequentemente do material quer seja na prateleira, quer seja em pesquisa. É disposto de uma forma que avança dos elementos mais abstractos para os mais concretos. A disposição mais utilizada é designada por ***ordem padrão de citação:***:

**Objecto – Tipo – Parte – Material – Propriedade – Processo – Operação – Agente – Espaço – Tempo**

Uma das vantagens da CDU é a facilidade de adaptar a ordem de citação de forma a ajustar-se a necessidades locais. Deve-se referir, contudo que o sistema é para ser usado para troca de informações.

 A ordem de citação utilizada pela Biblioteca Nacional de Portugal e recomendada para uso pelas bibliotecas cooperantes é a seguinte:

**Classificação Decimal Universal**

**Ordem de citação**:

Quando um elemento da notação é seleccionado como representativo de um aspecto do assunto de um documento e incorporado ao número de classificação desse documento, diz-se que ele é citado. Para a correcta representação de uma sequência de símbolos na formação de uma notação complexa ou composta estabeleceu-se a seguinte ordem de prioridades:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0/9** | **.01/.09** | **-1/-9** | **‘0/’9** | **A/Z** | **+** | **/** | **:** | **(=...)** | **(1/9)** | **«…»** | **(0...)** | **=...** |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** |

**1** CDU 0/9 das tabelas principais; **2, 3 e 4** Auxiliares especiais; **5** Extensão alfabética

**6** Reúne assuntos não consecutivos; **7** Reúne assuntos consecutivos

**8** Relaciona assuntos; **9** Auxiliar de nacionalidade e de raça

**10** Auxiliar de lugar; **11** de tempo; **12** de forma; **13** de língua

Por vezes surgem dificuldades em citar os diversos elementos. Algumas orientações tipo:

1. Índice principal + Auxiliar especial + Auxiliar comum
2. Índice principal + Auxiliar dependente + Auxiliar independente
3. Índice principal + Auxiliar especial ou dependente + Auxiliar de raça/nacionalidade + Auxiliar de lugar + Auxiliar de tempo + Auxiliar de forma + Auxiliar de língua

### Análise do tema

De forma a aplicar com sucesso uma classificação analítico-sintética como a CDU, é necessário tomarem-se certas decisões. Entre estas inclui-se a ordem de citação mais apropriada à situação.

O **tema** deve ser determinado e tomados os seguintes passos, pela ordem:

* Identificar o principal tema em discussão
* Seleccionar a classe principal apropriada
* Separar os diversos conceitos contidos no documento e atribuí-los ao grupo específico tal como descrito acima na fórmula da ***ordem padrão de citação***
* Seguir a ordem padrão de citação para combinar e reunir os vários elementos,

Em segundo lugar, a **forma** pela qual o material é apresentado deverá ser tomada em consideração. Pode estar relacionada com:

* A apresentação em relação aos assuntos que sejam significativos para o tema;
* A forma física

É apenas nesta fase que se devem verificar as exigências locais. A CDU assim o permite, mas quebrar a ordem de citação desta forma deverá ser o resultado de uma decisão política e deverá ser seguida de forma consistente, devendo por isso ser registada e justificada no **manual de procedimentos** ou no **ficheiro de autoridade**.

O procedimento normal é o seguinte:

1. Seleccionar o número das tabelas principais da CDU, juntamente com os auxiliares especiais apropriados que sejam necessários,
2. Seleccionar as numerações combinadas por meio de um dos instrumentos de ligação, tais como os dois pontos,
3. São aplicados os números das tabelas auxiliares comuns. A ordem de aplicação para estes últimos é normalmente:
4. Auxiliares para **pessoas, materiais** e **propriedades**,
5. Auxiliares comuns de **lugar, tempo**,
6. **Forma e de língua.**

Os elementos que se seguem sugerem combinações úteis e demonstram a importância de primeiro analisar cuidadosamente, depois aplicar a classificação e por último registar a decisão tomada.

### Estrutura da CDU (ver anexo A)

Todas as edições da CDU apresentam a mesma estrutura independentemente do tamanho (integral, média, abreviada) e da língua.

As tabelas médias impressas apresentam um **Índice Alfabético de Assuntos** que indica todas as classes possíveis para classificar determinado conceito.

As diferentes tabelas CDU possuem sempre uma **Introdução** que funciona como o corpo doutrinário do sistema, faz-se um breve apresentação da teoria e da história da CDU e definem-se as normas de uso das diferentes partes.

Seguem-se as **Tabelas de Auxiliares Comuns Gerais** que se subdividem em *Auxiliares independentes e dependentes*;

As **Tabelas de Auxiliares Especiais**;

As **Tabelas Principais** de 0/9

### Tabelas de Auxiliares

Cada edição indica, no início da tabela CDU, a **Secção I** que contem os **sinais e as subdivisões auxiliares comuns,** juntamente com notas sobre a sua aplicação e na **Secção II**, indica os sinais e as técnicas dos **auxiliares especiais** que indicam características recorrentes ao nível local de cada tabela principal.

* **Auxiliares comuns independentes**

Tabela 1c. Auxiliares de língua = …;
Tabela 1d. Auxiliares de forma (0…);
Tabela 1e. Auxiliares de lugar (1/9);
Tabela 1f. Auxiliares de raça, grupo étnico e nacionalidade (=…);
Tabela 1g. Auxiliar de tempo “…”.

* **Auxiliares comuns dependentes**

Tabelas1Ik. Auxiliares Comuns de Características Gerais:

-02 Propriedades;
-03 Materiais;
-05 Pessoas e características pessoais).

As primeiras, são designadas de **independentes** porque podem ser utilizadas independentemente de um número das tabelas principais e podem ser citadas no início, no meio ou no fim da notação. As designadas de **dependentes** só podem ser usadas como sufixos de um número das tabelas principais.

### Secção I: Sinais e subdivisões de auxiliares comuns

#### Tabela Ia. Adição. Extensão

**Secção 1. Coordenação. Adição**

**Símbolo: +** liga dois ou mais números CDU separados (não consecutivos):

*51+53 Matemática e Física*

**Secção 2. Extensão consecutiva**

**Símbolo: /** faz a ligação entre o primeiro número e o último de um grupo de números consecutivos da CDU para denotar um tema extenso ou uma série de conceitos.

*643/645 A casa, o equipamento e mobiliário doméstico*

Se o número depois da barra tem mais de três algarismos e começa com um grupo de algarismos, pode ser abreviado, desde que o primeiro item da barra seja um ponto.

*646.73/.74 Significa 646.73 e 646.74*

#### Tabela 1b. Relação. Subagrupamento. Fixação da ordem

**Secção 1. Relação simples (ver pág. 36)**

**Símbolo:** : (dois pontos) O sinal de relação é utilizado para ligar dois ou mais números da CDU. Ao contrário do sinal se somar e da barra, os dois pontos restringem em vez de alargarem os temas que ligam.

*631.11:502.17 Localização de quintas em reservas naturais*

Os números de ambos os lados dos dois pontos podem ser invertidos, dependendo da ênfase que se quer dar.

**Secção 2. Subgrupos (ver pág. 36)**

**Símbolo: […]** Os conceitos subordinados podem ser indicados até uma extensão limitada através da utilização de sinais algébricos de subagrupamento e de fixação de ordem.

*[622+669](485) Minas e metalurgia na Suécia*

**Secção 3. Ordenação (ver pág. 36)**

**Símbolo: : : (dois pontos duplos)** Os dois pontos em duplicado podem ser utilizados para estabelecer a ordem. Isto significa que o conceito que se segue a **::** tem uma relação de subordinação: *575::576.3 Citogenética*

### Auxiliares comuns independentes

#### Tabela Ic. Auxiliares Comuns de Língua

**Símbolo: = (igual)**

**=…‟0** Origens e períodos das línguas. Fases de desenvolvimento

**=00/03** Multilingue. Documentos traduzidos

**=1/=2** Línguas indo-europeias

**=3…** Línguas caucasianas, outras, Basco (Euskera**)**

**=4…** Línguas afro-asiáticas, nilo-saarianas, congo-kordofanas,

**=5…** Língua**s** uralo-altaicas, japonês, coreano, aino, paleo-siberianas

**=6…** Línguas austro-asiáticas, austronésias

**=7…** Línguas indo-pacíficas, australianas

**=8…** Línguas ameríndias

**=9…** Línguas artificiais

Os auxiliares comuns de língua expressam a língua ou a forma linguística de um documento cujo tema é expresso por um número principal da CDU. A Tabela Ic é o local principal para a enumeração de línguas e serve com fonte para a subdivisão da classe *811 Línguas* (enquanto assunto; como objecto de estudo), da classe *821 Literaturas de línguas particulares e da Tabela If. Auxiliares comuns de grupos étnicos.*

Na ordem de citação, o auxiliar de língua vem sempre em último lugar e serve para designar a língua em que o documento foi escrito, no entanto, pode ser citado no meio ou até em primeiro lugar num número composto, caso haja necessidade de classificar documentos por língua e não por tema.

#### Tabela Id. Auxiliares Comuns de Forma

**Símbolo: (0….) (parênteses 0)**

 **(0.0…)** Características físicas

**(01)** Bibliografia

**(02)** Livros em geral

**(03)** Obras de referência

**(04)** Separatas

**(05)** Publicações em série

**(06)** Publicações de organizações

**(07)** Documentos para o ensino, estudo, formação

**(08)** Poligrafias

**(09)** Apresentação em forma histórica. Fontes legais e históricas

Os auxiliares comuns de forma indicam a forma documental com que o tema se apresenta, representado por um número principal da CDU.

*572(035) Manuais de Antropologia Física*

Na ordem de citação, os auxiliares utilizam-se normalmente a seguir à numeração do tema, mas caso seja preferível, todos os documentos com a mesma forma de apresentação podem ser agrupados utilizando o auxiliar de forma adequada, como por exemplo:

*(05)… Todos os periódicos são arrumados juntos*

#### Tabela Ie. Auxiliares Comuns de Lugar

**Símbolo: (1/9) (parênteses, um a nove)**

 **(1)** Lugar e espaço em geral. Localização em geral. Orientação

**(2)** Designação fisiográfica

**(3)** Lugares do mundo antigo

**(4)** Europa

**(5)** Ásia

**(6)** África

**(7)** América do Norte e Central

**(8)** América do Sul

**(9)** Estados e regiões do Pacífico Sul e Austrália. Árctico. Antárctico

Na ordem de citação um auxiliar comum de lugar vem normalmente a seguir ao assunto/número principal da CDU.

#### Tabela If. Auxiliares Comuns de Raça, Grupo Étnico e Nacionalidade

**Símbolo: (=…) (parênteses sinal de igual)**

 **(=01/088)** Raças

**(=1:1/9)** Pessoas associadas a determinados lugares

**(=1:4/9)** Nacionalidades particulares do mundo moderno

**(=11/=8)** Povos, grupos étnicos, grupos linguístico culturais

Os auxiliares de grupo étnico derivam principalmente dos **Auxiliares Comuns de Língua (**Tabela Ic**)** e são utilizados para distinguir grupos linguístico-culturais.

*78(=411.16) Música judaica*

A nacionalidade política (cidadania de Estados-Nação) é expressa principalmente por (=1:4/9) derivados dos **Auxiliares Comuns de Lugar** (Tabela Ie)

*94(=1:469)(6) História dos portugueses em África*

#### Tabela Ig. Auxiliares Comuns de Tempo

**Símbolo: “…” (Aspas)**

 “0/2” Datas e períodos de tempo (d.C) no calendário convencional cristão

“3”Divisões e subdivisões cronológicas convencionais: numeradas, denominadas, etc.

“4” Duração. Período de tempo. Idades e grupos etários

“5” Periodicidade. Frequência. Repetição a intervalos determinados

“6” Divisões do tempo geológicas, arqueológicas e culturais

“7” Fenómenos no tempo. Fenomenologia do tempo

Os auxiliares comuns de tempo indicam a data, o momento ou o âmbito temporal de um tema representado por um número principal da CDU.

Na numeração utiliza-se o ponto para separar elementos temporais de magnitudes diferentes e não como em qualquer outra parte da CDU, após cada grupo de três algarismos. Os elementos temporais são citados por ordem decrescente de magnitude.

Na ordem de citação o auxiliar de tempo é normalmente colocado após o auxiliar geográfico.

Os séculos e as décadas são expressos por 2 ou 3 algarismos respectivamente.

*“03” Século IV d.*C ; *“20” Século XXI ; “196” Década de 60*

Extensões de vários séculos, décadas ou anos podem ser expressas pelos algarismos iniciais e finais ligados pela barra oblíqua / ([tabela Ia](#_Tabela_Ia._Adição.), Secção 2)

*“04/14” Século V a XV (Idade Média) ; “1939/1945” Período da IIª Guerra Mundial ;*

Os auxiliares de tempo não devem ser confundidos com os auxiliares comuns de forma para apresentação histórica: *53(091) História da Física ; 53”18” Física no séc. XIX*

#### Tabela Ih. Notações que não pertencem à CDU

**Símbolo:** \* (asterisco), **A/Z** (especificação alfabética)

1. \* Asterisco

* O asterisco pode seguir-se a um número da CDU para introduzir uma palavra, um símbolo, ou um número de uma fonte exterior à CDU que é acrescentado para especificar um assunto.
Quando é recomendado o uso de um código não pertencente à CDU a seguir ao \*, a fonte do código é indicada em nota. Ver exemplo da *Silvicultura 630* (pág. 611).
* O asterisco também pode ser utilizado para indicar uma numeração não autorizada, localmente atribuída, normalmente semelhante a um número CDU, para um conceito não existente na CDU.

2. A/Z Especificação alfabética directa

As especificações alfabéticas de nomes próprios, acrónimos e abreviaturas podem vir directamente a seguir ao número da CDU. As especificações alfabéticas dos auxiliares de lugar ([Tabela Ie](#_Tabela_Ie._Auxiliares)) devem ser colocadas entre parênteses, por exemplo:

*(469.521Beja) Concelho de Beja*

*821.134.3Saramago Literatura portuguesa: obra literária de Saramago*

*929Saramago Biografia de Saramago*

Pode ser conveniente reduzir todos os nomes a uma abreviatura de comprimento padrão, tal como as três primeiras letras, desde que o resultado não seja ambíguo: *929SAR*

Pode-se optar por expressar o nome na ordem invertida como a Biblioteca Nacional de Portugal faz, exemplo: *929Saramago, José*

***A regra é a consistência dos procedimentos, seja qual for a opção.***

### Auxiliares Comuns Dependentes

Existem três subdivisões dependentes que apenas podem ser utilizadas como sufixos. Não podem ser usados independentemente, nem ser citados em primeiro lugar numa notação composta. São sempre acrescentados a um número principal.

#### Tabela Ik. Auxiliares Comuns de Características Gerais

**Símbolos:**

**-02 (hífen zero dois);**

**-03 (hífen zero três);**

**-05 (hífen zero cinco)**

##### -02 Auxiliares Comuns de Propriedade

**-021** Propriedades de existência. Relação. Extensão. Valor. Origem. Ordem

**-022** Propriedades de magnitude. Quantidade, número. Valores temporais. Dimensão. Tamanho

**-023** Propriedades de forma

**-024** Propriedades de estrutura. Posição

**-025** Propriedades de disposição

**-026** Propriedades de acção e movimento. Direcção. Propriedades físicas. Ópticas. Radioactivas. Térmicas. Electromagnéticas. Propriedades da matéria

**-027** Propriedades operacionais. Desenvolvimento. Função. Produção. Propriedades organizativas. Membros

**-028** Propriedades de estilo e apresentação

**-029** Propriedades derivadas de outras classes principais

Se a forma é o assunto, utiliza-se o número principal: *050.9 Almanaques gerais*

Nos casos em que uma forma é parte de um tema, esta representa-se por meio do -02:

*050.9-028 Almanaques gerais ilustrados*

O que não se deve confundir com: *050.98 (084.1) Ilustrações de almanaques*

Nos casos em que se utilizem, ao mesmo tempo, números de mais de uma das [tabelas enumeradas em Ik](#_Tabela_Ik._Auxiliares).-0…, os auxiliares de *-02* são citados depois dos *-03* e *-05*, já que contém termos que matizam os enumerados nas outras duas tabelas.

##### -03 Auxiliares Comuns de Materiais

**-032** Materiais minerais de origem natural

**-033** Materiais manufaturados a partir de minerais

**-034** Metais

**-035** Materiais de origem principalmente orgânica

**-036** Materiais macromoleculares. Borrachas. Plásticos

**-037** Têxteis. Fibras. Tecidos

**-039** Outros materiais

##### -05 Auxiliares Comuns de Pessoas e Características Pessoais

**-051** As pessoas como agentes, como sujeitos activos

**-052** As pessoas como objeto, sujeitos passivos: clientes, utilizadores

**-053** Pessoas segundo a idade, ou grupo de idade

**-054** Pessoas segundo as características étnicas, nacionalidade, cidadania

**-055** Pessoas segundo o sexo e parentesco

**-057** Pessoas segundo a sua ocupação, trabalho, tipo de vida, educação

**-058** Pessoas segundo a sua classe social e estado civil

Se as tabelas principais já contemplam o aspecto pessoal, então as subdivisões de **-053/-058** podem acrescentar-se directamente: *347.96-055.2 Mulheres advogadas*

Tal como acontece com a maioria dos auxiliares comuns, as subdivisões de *-05* podem ser combinadas entre si ou com outros auxiliares:

*647-053.6-055.2 Mulheres adolescentes como empregadas domésticas*

*78.071-056.45(=411.16) Músicos prodígio hebreus*

### Auxiliares especiais

### Secção II: Tabelas auxiliares especiais

**Símbolos:**

**-1/-9 (hífen um barra oblíqua hífen nove);**

**.01/.09 (ponto zero um barra oblíqua ponto zero nove);**

**‘0/’9 (apóstrofe zero barra oblíqua apóstrofe nove)**

Deve-se ter em atenção estas subdivisões especiais e o classificador deve sempre comprovar, no início da classe ou da subclasse adequada, a tabela dos auxiliares especiais.

### Tabelas Principais da CDU

0 Generalidades. Ciência e conhecimento. Gestão. Informação. Documentação. Biblioteconomia. Organizações. Documentos e publicações

1 Filosofia. Psicologia

2 Religião. Teologia

3 Ciências Sociais. Estatística. Demografia. Sociologia. Política. Economia. Comércio. Direito. Administração Pública. Assuntos militares. Bem-estar social. Seguros. Educação. Folclore. Etnologia

5 Matemática. Ciências naturais

6 Ciências aplicadas. Medicina. Tecnologia

7 Arte. Recreação. Entretenimento. Desporto

8 Línguas. Linguística. Literatura

9 Geografia. Biografias. História

### Vantagens e desvantagens da utilização da CDU na indexação por assuntos

#### As vantagens de utilização da CDU são várias:

* Permite utilizar **vários níveis de especificação**
* Permite **organizar e recuperar a informação por assuntos** e simultaneamente permite **arrumar os documentos por áreas temáticas na estante em livre acesso.**

É importante estabelecer em simultâneo o seu Plano de Classificação e o Plano de Cotas.

* Permite realizar a **difusão selectiva da informação**

O serviço de informação deverá conhecer o perfil do utilizador.

* É uma **linguagem universal**, os técnicos de qualquer lado do mundo entendem o sistema de códigos numéricos e podem cooperar si na troca de informação.
* **Compatibiliza as diferentes linguagens de indexação terminológicas independentemente do idioma**
* Permite a **constituição de listas bibliográficas ordenadas por assunto**
* A estrutura hierárquica permite **alargar ou restringir a pesquisa.**

#### As desvantagens da utilização da CDU são várias:

* A **demora na atualização do sistema** em relação aos avanços da ciência;
* **Índice alfabético** insuficiente pois não estabelece as relações de equivalência nem as associativas e é exclusivo das tabelas médias;
* A organização do saber numa estrutura de classes já desactualizadas e que não pode ser alterada, ao comprometer toda a estrutura do sistema;
* A dificuldade de os utilizadores compreenderem as notações complexas.

# Anexo A – Estrutura da CDU



# Actividade Formativa 3

## A qualidade da indexação: factores condicionantes

* **Qualidade da indexação;**
* **Controlo do ficheiro de autoridade;**
* **Política de indexação;**
* **Manual de procedimentos de indexação.**

**A qualidade da indexação** é assunto estudado ao nível da biblioteconomia e os diversos autores têm apontado múltiplos elementos condicionadores dessa propriedade no processo de indexação. O que dificilmente encontramos é bibliografia e estudos que liguem a qualidade da indexação, às políticas de indexação dos diferentes serviços e que assinalem a importância das segundas sobre a primeira.

### 1. Qualidade da indexação

A qualidade da indexação é definida em função da eficácia da recuperação.

**A qualidade da indexação é condicionada por múltiplos factores:**

* Ligados ao indexador;
* Ligados ao documento;
* Ligados aos instrumentos de indexação;
* Ligados aos instrumentos auxiliares;
* Ambientais;
* Ligados ao processo de indexação

**Factores ligados ao indexador:**

* Formação do indexador (conhecimento das áreas de assunto a indexar);
* Conhecimento das necessidades dos utilizadores;
* Domínio dos instrumentos de indexação e aplicação da política de indexação;
* Características pessoais

**Factores ligados ao documento:**

* Conteúdo temático;
* Complexidade da abordagem;
* Língua e linguagem;
* Apresentação e sumarização

**Factores ligados aos instrumentos de indexação:**

* Qualidade do vocabulário (mede-se pelo nº de relações semânticas estabelecidas e pela sua actualização);
* Nível de especificidade e sintaxe;
* Ambiguidade e precisão

**Factores ligados aos instrumentos auxiliares:**

* Sistema de gestão da base de dados (capacidade, tempo de resposta);
* Interface do sistema de pesquisa

**Factores ambientais:**

* Iluminação;
* Ruído;
* Temperatura;
* Condições de trabalho em geral

### 2. Controlo do ficheiro de autoridades de assunto e a qualidade da indexação

Um registo de autoridade tem por função fixar, de forma normalizada, critérios uniformes de controlo de autoridade sobre a base bibliográfica, nomeadamente de Autor Pessoa Física; Colectividade, Família e Assunto.

A interacção do ficheiro de autoridade com o ficheiro bibliográfico garante a aplicação uniforme da linguagem documental, quer na indexação quer na pesquisa.

É no ficheiro de autoridade que o indexador pode estabelecer o controlo do vocabulário documental, face à linguagem natural, em dois aspectos essenciais: a forma dos termos e o seu significado. Este controlo realiza-se de acordo com as normas e directrizes definidas pelo serviço para o estabelecimento de referências de equivalência de modo a evitar a sinonímia, pelo estabelecimento de notas explicativas, destinadas a limitar o âmbito de aplicação de um termo e o estabelecimento das relações hierárquicas e associativas para cada termo de indexação. Este processo de normalização e controlo do vocabulário realiza-se no ficheiro de autoridade de assuntos utilizando as regras para o seu preenchimento prescritas no Manual UNIMARC/Formato Autoridades. Lisboa, BNP. 2008.

### 3. Política de indexação

Na implantação de um serviço de indexação há que se levar em conta um grande número de variáveis que irão afectar consideravelmente o desempenho do serviço. Definir essas variáveis, estabelecer princípios e critérios que servirão de **guia na tomada de decisões para optimização do serviço**, **racionalização dos processos e consistência das operações nele envolvidas**, são objectivos de uma política de indexação.

#### Requisitos para o estabelecimento de uma política de indexação

O principal propósito de um serviço de indexação é assegurar da forma mais eficiente e económica possível, que qualquer documento ou informação seja fornecido ao usuário no momento preciso.

Ao se estabelecer uma política de indexação tem que se levar em conta alguns factores que são imprescindíveis ao planeamento de qualquer sistema de recuperação de informação:

* 1. Identificação da organização à qual estará vinculado o sistema de indexação;
	2. Identificação da clientela a que se destina o sistema;
	3. Recursos humanos, materiais e financeiros.

##### A identificação da organização

O conhecimento dos objectivos e actividades da organização é fundamental para a determinação do tipo de serviço a ser implementado.

O tipo de organização determinará também que sistema de indexação usar, bem como os níveis de exaustividade e especificidade exigidos.

##### Identificação da clientela

A principal finalidade de um sistema de recuperação de informação é fornecer aos utilizadores a informação na forma e no momento exigidos por eles, a identificação desses utilizadores torna-se um pré-requisito para o planeamento de qualquer sistema de informação. Os estudos do perfil do utilizador são importantes para definição de uma política de indexação.

Através dos estudos de perfil do utilizador, obteremos:

* Conhecimento do alcance exigido pelo sistema quanto aos assuntos centrais e periféricos e quanto aos níveis de tratamento exigidos;
* O núcleo de um vocabulário que reflectirá os interesses do trabalho e necessidades de informação da clientela do sistema;
* Conhecimento do tipo de resposta exigido do sistema, isto é, se é exigido maior **revocação**, maior **precisão** ou ambas simultaneamente;
* Conhecimento do **nível de exaustividade** exigida na indexação; do grau de especificidade necessária à linguagem de indexação; o nível de sofisticação desejável no sistema,
* Conhecimento das exigências dos utilizadores quanto à forma de apresentação dos resultados de busca.

##### Determinação dos recursos financeiros, materiais e humanos

A limitação quanto às despesas constitui um factor importante no planeamento de um sistema de recuperação de informação.

Quando há limitações de recursos financeiros, devem-se tomar medidas tais como: adoptar uma linguagem já existente; quando se tornar necessária uma nova linguagem de indexação, usar, na sua elaboração termos simples, especificidade moderada, estrutura mínima e evitar artifícios complexos.

Um sistema deverá ser desenvolvido com um grau de sofisticação que seja necessário ao atendimento das exigências dos utilizadores. O uso de mecanismos para aumentar **a revocação** ou **a precisão** aumentará, consequentemente, o custo do sistema.

O **nível de exaustividade** na indexação é também um factor de aumento de custo. Na indexação dos documentos deve ser estabelecido um nível óptimo de exaustividade.

O **nível de especificidade** deverá ser mantido de acordo com as exigências feitas ao sistema. Será antieconómico usar um vocabulário muito específico quando o tipo de documentos e as solicitações feitas ao sistema são mais gerais.

A disponibilidade de equipamentos influenciará o tipo de organização de sistema bem como a escolha da linguagem de indexação.

A existência de recursos humanos adequados é outro factor a ser considerado no planeamento de um sistema de indexação. Quando há carência de pessoal deve-se optar por um sistema que exija uma manutenção mínima e um menor esforço na indexação, principalmente quando se tem um volume grande de documentos a serem indexados.

#### Elementos da política de indexação

Um sistema de recuperação de informação envolve uma série de decisões relativas a cada fase dos processos nele envolvidos e que irão afectar o desempenho do sistema.

Elementos que devem ser considerados na elaboração de uma política de indexação:

##### Cobertura de assuntos

Pelo estudo do perfil de utilizador, o técnico que planeia um sistema de recuperação já terá uma ideia dos assuntos a serem cobertos pelo sistema. Poderão estar identificadas as áreas onde se torna necessário um tratamento em profundidade e aquelas a serem tratadas superficialmente. Entretanto algumas áreas de interesse podem já estar cobertas por outros sistemas. Uma medida mais económica será a de se utilizarem os produtos desses sistemas (índices, resumos, bibliografias correntes, etc.) ao invés de indexar os documentos já cobertos por eles.

##### Selecção e aquisição dos documentos-fonte

Um importante aspecto a ser considerado no planeamento de um serviço de informação é o estabelecimento de uma adequada política de selecção e aquisição dos documentos a incluir no sistema. A adequação do nível intelectual e técnico dos documentos ao nível dos utilizadores; domínio, por parte dos utilizadores da língua em que os documentos estão escritos; limitações de recursos financeiros, o que implicará a determinação de prioridades na aquisição.

##### O processo de indexação

Cada etapa do processo de indexação é afectada por variáveis que vão influenciar praticamente todo o processo de recuperação da informação. Tais variáveis referem-se aos **níveis de exaustividade** e **especificidade** requeridos pelo sistema, **linguagem de indexação**, **capacidade de revocação e precisão** do sistema.

1. **Nível de exaustividade**

O nível de exaustividade da indexação é uma decisão política estabelecida pela administração do sistema de recuperação, de acordo com o propósito do mesmo.

O tipo de documento poderá determinar o nível de exaustividade de indexação.

1. **Nível de especificidade**

Há uma relação entre especificidade e a capacidade de precisão de um sistema. Maior grau de especificidade aumenta a taxa de precisão e diminui a revocação.

Uma importante decisão política na indexação refere-se ao nível da especificidade. Um sistema poderá decidir entre uma indexação em maior ou menor profundidade dependendo da procura ou das pesquisas que lhe são feitas.

1. **Escolha da linguagem**

A linguagem de indexação afecta o desempenho de um sistema de recuperação de informação em dois pontos; na estratégia de busca, estabelecendo a precisão com que o técnico de busca pode descrever os interesses do utilizador e na indexação, estabelecendo a precisão com que o indexador pode descrever o assunto dos documentos.

Se o sistema é automatizado e a sua área de assunto é bastante específica, poderá ser escolhida uma linguagem livre cuja maior vantagem é a rapidez na operação de indexação e a possibilidade de se utilizar pessoal menos qualificado. Porque as palavras usadas como termos de indexação são as próprias palavras dos autores, para assegurar recuperação adequada, o pesquisador precisa de pensar em todas as alternativas de grafia possíveis, singulares e plurais, sinónimos.

Com a linguagem controlada, a operação de indexação é mais lenta, mas o esforço despendido na busca é reduzido. A linguagem controlada permite uma maior consistência na indexação, o que a torna mais indicada a um sistema de recuperação que actue cooperativamente.

A linguagem pré-coordenada tem a vantagem de ser mais precisa e de facilitar a estratégia de busca, evitando falsas associações e relações incorrectas. A sua desvantagem é o aumento do custo da indexação devido ao crescimento do arquivo pela repetição de termos na entrada e o uso de referências.

A linguagem pós-coordenada permite uma maior revocação do que a linguagem pré-coordenada, apesar de não oferecer uma alta precisão. Uma outra desvantagem é que, com a combinação dos termos apenas na saída, podem-se obter associações falsas e relações incorrectas. Tais problemas podem ser evitados pelo uso de artifícios de precisão como elos e indicadores de função, mas tais artifícios acarretam um aumento de custo tanto no estágio de indexação quanto no de busca.

1. **Capacidade de revocação e precisão do sistema**

A **revocação** relaciona-se com a capacidade do sistema em assegurar a recuperação de um número desejável de documentos relevantes e a **precisão** relaciona-se com a capacidade do sistema em impedir a recuperação de documentos não relevantes.

Um serviço deverá decidir quanto aos seus níveis de revocação e precisão tendo por base o interesse dos seus utilizadores.

1. **Estratégias de busca/pesquisa**

Uma importante decisão a ser tomada no estágio de busca é se ela é delegada ou não. Na pesquisa delegada o utilizador transfere a responsabilidade da pesquisa a um especialista da informação.

1. **Tempo de resposta do sistema**

O tempo de resposta de um sistema de recuperação será determinado em função das necessidades dos utilizadores a que o sistema se destina. Sendo crucial em determinadas situações, o factor tempo de resposta é secundário em relação ao factor precisão.

1. **Forma de saída**

Um sistema pode apresentar como resposta a uma pesquisa: números de acesso referentes aos documentos, referências bibliográficas, resumos, ou o texto completo.

O formato tem grande influência sobre a tolerância do utilizador quanto à precisão dos resultados. Um utilizador que recebe resumos como resultado da sua pesquisa pode tolerar uma precisão mais baixa do que o utilizador que recebe os resultados em forma de referências bibliográficas ou apenas números de acesso aos documentos. É mais fácil fazer a selecção a partir dos resumos do que ter de consultar um grande número de documentos para chegar aos que realmente interessam.

1. **Avaliação do sistema**

A avaliação do sistema determinará até que ponto se está a satisfazer as necessidades dos seus utilizadores, que falhas estão a ocorrer e de que forma poderão ser corrigidas.

Embora um programa de avaliação possa concentrar-se sobre um determinado subsistema do sistema de indexação, este subsistema não pode ser avaliado isoladamente. Isto porque os vários subsistemas (indexação, pesquisa, linguagem de indexação, interface) estão intimamente ligados, pelo que, a mudança num deles repercutir-se-á nos outros. Assim, não será possível avaliar significativamente um sistema como um todo, sem se considerar o efeito da política e procedimentos de indexação sobre o subsistema de pesquisa. Da mesma forma, uma mudança de vocabulário afectará tanto a operação de indexação como a de pesquisa.

### 4. Manual de indexação

Com base no que foi exposto, no tópico relativo à política de indexação, vamos identificar quais os elementos essenciais a constar no **Manual de Procedimentos de Indexação**.

Entende-se por política de indexação o conjunto de directrizes que explicam e justificam as opções técnicas que a biblioteca/serviço de informação faz e às quais os profissionais deverão responder tecnicamente, ou seja, as suas rotinas de trabalho submetem-se a essas orientações que são estabelecidas considerando: o acervo e o utilizador.

O Manual de Procedimentos de Indexação consiste na formalização documental da política de indexação. Deve começar por **enunciar a política de indexação e os elementos que a constituem**.

Em segundo lugar, deve **dispor acerca do tratamento que deverá ser dado às diferentes áreas temáticas que o acervo contempla** – os serviços não têm que realizar um tratamento igualitário a todo o fundo documental. Em terceiro lugar, **deverá seleccionar a ou as linguagens documentais** que irá utilizar no processo de indexação.

O passo seguinte é **definir como se procederá à actualização da linguagem de indexação adoptada**.

No decorrer do processo de indexação **as decisões que são tomadas quotidianamente devem ser registadas** de forma a dar consistência ao processo.

Deverá ficar definida qual a periodicidade para se **realizar a avaliação crítica das decisões tomadas**.

**Ao elaborar os manuais de procedimentos de indexação devemos incluir**:

1. Introdução (onde se esclarece o propósito do manual)
2. Índice (para facilitar a consulta do documento)
3. Bibliografia
4. Deve conter a política de indexação e os seus elementos constituintes:
	1. Missão do serviço de informação;
	2. Perfil do utilizador;
	3. Características do acervo;
	4. Estrutura organizacional do sistema de informação;
	5. Qualidade e quantidade de recursos humanos, materiais e financeiros afectos;
	6. Tipo de sistema de indexação seleccionado: pré ou pós coordenado;
	7. Linguagem documental seleccionada;
	8. Processo de indexação (níveis de exaustividade e de especificidade requeridos pelo sistema, linguagem de indexação, capacidade de recuperação/revocação e precisão do sistema);
	9. Decisão acerca da estratégia de busca;
	10. Decisão sobre o sistema de gestão de dados (capacidade, interface amigável, tempo de resposta, formatos de extracção das informações recuperadas, interoperabilidade, etc.);
	11. Indicação dos procedimentos de avaliação da própria política de indexação
5. Deve dar múltiplos exemplos para os diversos procedimentos adoptados
6. Deve ser didáctico
7. Deve ser aberto e permitir actualizações constantes que resultam das dificuldades e das soluções práticas encontradas pelos indexadores no decorrer da sua experiência servindo como subsídio para as actualizações do manual de indexação.

O Manual SIPORbase é um bom exemplo de um manual de indexação onde se estabelecem os princípios orientadores de todo o processo de indexação para a Biblioteca Nacional de Portugal cujas orientações podem ser aplicadas a qualquer serviço de informação, independentemente da opção por um sistema pré ou pós-coordenado:

1. Organização e directrizes de utilização do manual

2. Princípios básicos e política de indexação

3. Estrutura e caracterização da linguagem documental

4. Terminologia

# Índice Remissivo

A

análise de conteúdo 3

condensação 3

indexação 3

resumo 3

Análise do tema 25

Apresentação sistemática

Combinação dos dois tipos de organização 16

Organização em domínios 16

Organização por facetas 16

Auxiliares Comuns Dependentes 31

Auxiliares especiais 32

C

cabeçalhos de assunto 8

cadeia notacional 7

CDU

Estrutura 26

**Master Reference File** 21

Classificação

faceta 21

***classificação analítico-sintética*** 22, 25

Classificação Decimal de Dewey 7

Classificação Decimal Universal 7

*classificação hierárquica* 22

***classificações universais*** 7

Composição do tesauro 13

Relações semânticas 13

Unidades lexicais 13

controlo do vocabulário

homonímia 9

Singular e plural 9

controlo do vocabulário 9

NP 4036 9

polissemia 9

sinonímia 9

**controlo do vocabulário** 10

**controlo do vocabulário**

Controlo formal 10

**controlo do vocabulário**

Controlo semântico 10

Controlo formal 10

Controlo semântico 10

via significado 10

via significante 10

**coordenação dos termos** 9

pós-coordenada 9

pré-coordenada 9

critérios de tipificação 8

controlo 8

coordenação dos termos 8

estrutural 8

D

**descritores** 13

Apresentação alfabética 15

Apresentação gráfica 16

Apresentação sistemática 15

carga informativa 13

composição 14

conteúdo 13

cronológicos 13

Descritores compostos 14

Descritores simples 14

Formas de apresentação 15

geográficos 13

materiais 13

Normas apresentação 14

onomásticos 13

primários 13

secundários 13

sintagma nominal 14

E

estrutura combinatória

cabeçalhos de assunto 9

Estrutura da CDU 33

estrutura hierárquica

categorial 9

EUROVOC 12

Evolução do EUROVOC 17

Historial 17

Tesauro 17

Vantagens e limitações 20

F

faceta *See* Classificação

***facetas*** 22

Fases do processo de indexação 4

1ª fase 4

2ª fase 4

3ª fase 4

especificidade 4

exaustividade 4

Identificação sistemática dos conceitos 4

Selecção dos conceitos 4

Funções do tesauro 13

indução 13

normalização do vocabulário 13

representação 13

H

homonímia

Homófonas 9

Homógrafas 9

I

**Índice Alfabético de Assuntos** 26

L

linguagem artificial 4

**Linguagem combinatória** 7

linguagem controlada 8

linguagem documental

Artificial 7

**controladas** 8

Controladas 7

Convencionais 7

**livres** 8

**Linguagem documental** 4

linguagem natural 8

**linguagens classificatórias** 7

**linguagens vocabulares** 7, 8

M

**Mapas conceptuais** 8

N

Norma ISO 2788 10

NP 3715 3

NP 4036 9, 12, 14, 15

NP 418 3

**NP 4285-4** 4, 12

O

**ontologias** 8

ordem de citação 23, 25, 28, 29, 30

**Ordem de citação** 24

**ordem padrão de citação**

estrutura 24

P

pré-coordenado 7

Q

qualidade da indexação 34

Qualidade da indexação 34

R

**Relações hierárquicas** 11

**Relação de instância** 11

**Relação genérica** 11

**Relação partitiva ou todo/parte** 11

resumos analíticos

Indicativo 4

Informativo 3

Informativo – indicativo 4

S

Sinais e subdivisões de auxiliares comuns 27

Símbolo (=…) 29

Símbolo (1/9) 29

Símbolo \* 30

Símbolo / 27

Símbolo : 27

Símbolo […] 27

Símbolo “…” 29

Símbolo + 27

Símbolo = 27

Símbolo dois pontos duplos 27

Sistemas de indexação 5

coordenação 5

exaustividade 6

pós-coordenados 5

precisão 6

pré-coordenados 5

sistema pós-coordenado 6

sistema pré-coordenado 5

T

Tabela 1b. Relação. Subagrupamento. Fixação da ordem 27

tabela Ia 30

Tabela Ia. Adição. Extensão 27

Tabela Ic 29

Tabela Ic. Auxiliares Comuns de Língua 27

Tabela Id. Auxiliares Comuns de Forma 28

Tabela Ie 29, 30

Tabela Ie. Auxiliares Comuns de Lugar 29

*Tabela If. Auxiliares comuns de grupos étnicos* 28

Tabela If. Auxiliares Comuns de Raça, Grupo Étnico e Nacionalidade 29

Tabela Ig. Auxiliares Comuns de Tempo 29

Tabela Ih. Notações que não pertencem à CDU 30

Tabela Ik. Auxiliares Comuns de Características Gerais 31

Tabelas Auxiliares 26

Tabelas auxiliares especiais 32

**Tabelas de Auxiliares Comuns Gerais** 26

**Tabelas de Auxiliares Especiais** 26

**Tabelas Principais** 26

Tabelas Principais da CDU 32

**Taxonomias** 8

tesauros 7, 8, 12

Thesaurus 12

Tipologia das linguagens documentais 8

tipos de linguagem 9

**estrutura combinatória** 9

**estrutura hierárquica** 9

V

Vantagens e desvantagens da utilização da CDU 33