



## PRESENTACIÓ DEL TESAUROS

1. Objectius d'un tesaurus
2. Camps temàtics
3. Estructura
  - a. Camps temàtics i microtesaurus
  - b. Relacions semàntiques
    - i. Notes d'aplicació i definició
    - ii. Relació de pertinença a un microtesaurus
    - iii. Relació d'equivalència
    - iv. Relació jeràrquica
    - v. Relació associativa

### 1. Objectius d'un tesaurus

#### **Què és un tesaurus?**

Tenint en compte que la nostra eina de treball és el tesaurus Eurovoc, elaborat pel Parlament Europeu i l'OPOCE (Oficina de Publicacions Oficials de les Comunitats Europees), ens remetem a la definició reflectida en aquest vocabulari: Un tesaurus de descriptors és una llista estructurada d'enunciats de conceptes. Aquests enunciats pretenen representar de manera unívoca el contingut conceptual dels documents i de les preguntes en un sistema documental.

#### **Per a què serveix?**

Per representar de manera unívoca el contingut dels documents.

La utilització d'aquesta mena de vocabulari es fa necessària atesa l'ambigüitat que representa el llenguatge emprat tant pels autors del material documental com pels usuaris que pretenen trobar els documents.

#### **Avantatges de la utilització d'un tesaurus**

Un tesaurus té uns avantatges clars davant qualsevol altre llenguatge d'indexació:

– Evita la *sinonímia*: el mateix concepte es pot expressar mitjançant diversos sinònims o termes molt pròxims (per exemple, "agricultura", "sector agrari", "sector agropecuari"). Com que seria impossible trobar un document indexat sota qualsevol d'aquests termes a partir d'una pregunta referent a un dels seus sinònims, una de les comeses fonamentals del tesaurus consisteix a evitar els inconvenients derivats de la sinonímia: s'escull, de manera més o menys arbitrària, un dels sinònims com a descriptor; i als altres sinònims, o termes que designen conceptes molt pròxims al que representa el descriptor, se'ls dona l'estatus de no-descriptors.

– Evita l'*homonímia*: un terme pot tenir significats diferents (per exemple, "premsa" pot significar màquina per comprimir o, en sentit figurat, impremta, o també conjunt de publicacions periòdiques, especialment, les diàries, o conjunt de persones dedicades al periodisme). La interrogació d'un terme homònim remetria a documents sense cap relació amb el tema que interessa l'usuari. Una altra de les comeses fonamentals del tesaurus consisteix, doncs, a evitar els inconvenients derivats de l'homonímia: tots els descriptors se situen en un context determinat, de manera que el significat és unívoc. Les resolucions adoptades per als altres casos poden passar per introduir qualificadors.



### Elements constitutius d'un tesaurus

1. *Descriptors*: paraules o expressions que designen sense ambigüitat els conceptes constitutius de l'àrea coberta pel tesaurus (per exemple: 'sector agrari').
2. *No-descriptors*: termes pertanyents a un mateix conjunt de sinònims o quasisinònims però que no han sortit elegits com a entrada d'autoritat o descriptor. Tenen una relació d'equivalència amb el terme escollit com a descriptor. No es poden utilitzar per indexar, però permeten millorar la coherència de la indexació.

## 2. Camps temàtics

El tesaurus Eurovoc amb el desenvolupament propi del Parlament pretén cobrir totes les àrees d'activitat de les institucions, tant de les europees com de les catalanes:

Vida política  
Relacions internacionals  
Comunitats europees  
Dret  
Vida econòmica  
Intercanvis econòmics i comercials  
Afers financers  
Afers socials  
Educació i comunicacions  
Ciència  
Empresa i competència  
Treball i ocupació  
Transports  
Medi ambient  
Agricultura, silvicultura i pesca  
Sector agroalimentari  
Producció, tecnologia i investigació  
Energia  
Indústria  
Geografia  
Organitzacions internacionals  
Toponímia de Catalunya

El desenvolupament de cada un d'aquests camps està en relació amb l'interès o la proximitat de la institució als temes a què fan referència.

D'altra banda, cal tenir en compte que la divisió dels descriptors en camps temàtics és, d'alguna manera, arbitrària. Molts descriptors podrien pertànyer a més d'un camp temàtic. Cal dir, però, que de manera general s'accepta que per poder gestionar correctament un tesaurus se n'ha de restringir la polijerarquia, és a dir, la inclusió d'un descriptor en més d'un camp. D'acord amb aquest principi, el tesaurus Eurovoc inclou els descriptors en un sol camp de tots aquells en què es podrien incloure. El camp elegit acostuma a ser el que la majoria d'usuaris pot considerar més natural.



### 3. Estructura

#### 3.a) Camps temàtics i microtesaurus

L'estructura del tesaurus consisteix en una classificació jeràrquica organitzada en dos nivells:

Camps temàtics: identificats per un número de dos dígit i un enunciat.

camp	enunciat
10	comunitats europees

Microtesaurus: identificats per un número de quatre dígit –els dos primers corresponen al camp temàtic i els altres dos a una seqüència dintre aquest camp temàtic– i un enunciat.

microtesaurus	enunciat
1016	construcció europea

#### 3.b) Relacions semàntiques

En l'àmbit específic dels descriptors i no-descriptors, el tesaurus es presenta amb una estructura formada per les relacions semàntiques següents:

- 3.b.1) Notes d'aplicació i definició
- 3.b.2) Relacions de pertinença a un microtesaurus
- 3.b.3) Relacions d'equivalència
- 3.b.4) Relacions jeràrquiques
- 3.b.5) Relacions associatives

##### 3.b.1) Notes d'aplicació i definició

Alguns descriptors en tenen. La introdueix el símbol **SN** (Scope Note). Tenen una doble funció:

- 3.b.1.a)       precisar el sentit del descriptor (nota de definició o explicativa)

evasió de capitals SN    Fugida de diners cap a l'exterior.
--

- 3.b.1.b)       determinar els supòsits d'utilització del descriptor (nota d'aplicació).

fauna SN    L'utilitzarem, com els termes específics, per als animals salvatges, en llibertat o en captivitat.
---

##### 3.b.2) Relacions de pertinença a un microtesaurus

Tots els descriptors van acompanyats d'una relació de pertinença, introduïda pel símbol **MT** (Microtesaurus) per indicar el microtesaurus o els microtesaurus a què pertany.

jove MT    2816 demografia i població
--



PARLAMENT DE CATALUNYA

3.b.3) *Relacions d'equivalència*

a) **UF** (Used for= Utilitzeu-lo per), situat entre el descriptor i el no-descriptor o no-descriptors que representa.

protecció civil UF protecció de la població
--

b) **USE** (Use= Utilitzeu), situat entre un no-descriptor i el descriptor que representa.

protecció de la població USE protecció civil (0431)
--

La relació d'equivalència abasta, de fet, diverses relacions, de les quals podem destacar:

b.1) la sinonímia total (la identitat de sentit)

Banc Europeu d'Inversions USE BEI (1006)
---

b.2) la quasisinonímia (la proximitat de sentit)

estimació de costos USE càlcul de costos (4026)
--

b.3) l'antonímia (l'oposició de sentit)

llengua morta USE llengua viva (3611)
--

b.4) la inclusió (un descriptor determinat engloba un concepte específic o més d'un i li dona categoria de no-descriptor perquè no té la freqüència d'ús del descriptor)

cinturó de seguretat USE dispositiu de seguretat (4811)
--

3.b.4) *Relacions jeràrquiques*

a) **BT** (Broader Term= Terme genèric), acompanyat d'una xifra que indica el nombre de nivells jeràrquics entre el descriptor específic i cada un dels genèrics.

educació d'estrangers

MT 3206 educació

BT1 educació

b) **NT** (Narrower Term= Terme específic): situat entre un descriptor genèric i un d'específic acompanyat d'una xifra que indica el nombre de nivells jeràrquics que hi ha entre el genèric i cada un dels específics.

educació

MT 3206 educació

BT1 educació

NT1 educació d'adults

NT1 educació per a estrangers



## PARLAMENT DE CATALUNYA

Cal tenir present que, els descriptors que no tenen termes genèrics, l'Eurovoc els anomena 'caps de jerarquia' o 'capçaleres de sèrie d'una jerarquia' ('top terms'). També els podem anomenar 'descriptors capçalera'.

D'altra banda, a l'Eurovoc hi han alguns descriptors dels camps temàtics 72 (Geografia) i 76 (Organitzacions internacionals) que són polijeràrquics pel que fa a grups temàtics, en el sentit que tenen diversos termes genèrics de primer nivell.

### 3.b.5) *Relacions associatives*

S'indica mitjançant el símbol **RT** (Related Term= Terme relacionat/associat) situat entre dos descriptors associats.

La relació associativa abasta, de fet, diverses relacions, entre les quals distingirem les següents:

#### 3.b.5.a) Concomitància

salari baix RT    renda baixa
----------------------------------

#### 3.b.5.b) Successió en el temps o en l'espai

invenció RT    patent
--------------------------

#### 3.b.5.c) Relació amb elements constitutius

carn equina RT    equí
---------------------------

#### 3.b.5.d) Relació de propietat

govern a l'exili RT    legitimitat
---------------------------------------

#### 3.b.5.e) Objecte d'una acció, procés o disciplina

soroll RT    aïllament acústic
-----------------------------------

#### 3.b.5.f) Localització

gestió comptable RT    societat fiduciària
---

#### 3.b.5.g) Similitud (quan s'ha conferit la categoria de descriptors a termes pràcticament sinònims)

finançament dels partits RT    finançament electoral
---



PARLAMENT DE CATALUNYA

3.b.5.h) Antinòmia

censura RT llibertat d'informació
--------------------------------------

A més, la relació associativa té un conjunt de característiques importants:

- Simetria: la relació dels termes sempre és recíproca

salari baix RT renda baixa	renda baixa RT salari baix
-------------------------------	-------------------------------

- Incompatibilitat amb la relació jeràrquica: *si dos descriptors estan vinculats per una relació jeràrquica, no hi pot haver cap relació associativa que els uneixi, ni a l'inrevés.*
- Els descriptors situats sota un mateix terme capçalera *no poden estar units per una relació associativa.*