

MINI-MANUAL PARA EL USO DEL
TESAURO MULTILINGUE POPINS
POR DOCPAL

Preparado por Sandra Acuña
septiembre de 1978.

INTRODUCCION

Un tesauo puede definirse en términos de su función o de su estructura. De acuerdo a su función, un tesauo es un instrumento de control terminológico que se utiliza para traducir a un lenguaje documentario el lenguaje natural usado tanto por los individuos y los usuarios como por los propios documentos. Desde el punto de vista de su estructura, es un vocabulario controlado y dinámico de términos que guardan relación entre sí, estas relaciones pueden ser de tipo semántico, es decir, referentes al significado de las palabras, o bien, relaciones de tipo jerárquico, que expresan superioridad o subordinación 1/.

A diferencia de un diccionario que entrega definiciones de palabras o de términos, un tesauo entrega las palabras o descriptores, expresando relaciones entre ellas y sin incluir, obligatoriamente, explicaciones o definiciones.

La presentación gráfica de los tesauos puede variar, pero habitualmente son dos las ordenaciones que se prefieren:

- a) Presentación sistemática de términos. Se agrupan los términos por su afinidad y cada grupo afin recibe el nombre de faceta la cual se identifica con una anotación numérica o alfa-numérica.
- b) Presentación alfabética de términos. En estricto orden alfabético se reordenan los términos de las facetas. El término, por tanto, irá siempre precedido o seguido por una anotación que equivale a la de su ubicación dentro de una faceta.

1/ UNESCO, UNISIST. guidelines for the establishment and development of monolingual thesauri. Paris, UNESCO, 1973. (SC/WS/JJJ).

1. Relaciones

Las relaciones entre términos (descriptores cuando se trata de un tesoro) son básicamente tres y en todos los casos de carácter recíproco, es decir, la relación opera en ambos sentidos.

1. 1. Relaciones de equivalencias (UF; USE)

Esta relación controla los sinónimos o cuasi-sinónimos, esto es, términos que tienen el mismo sentido en una disciplina específica. Se entiende por cuasi-sinónimos aquellos términos cuyo significado puede ser diferente en otro vocabulario especializado, pero que se consideran como sinónimos para las necesidades de un sistema específico de información.

La relación se expresa en la siguiente forma:

- a) Utilizando el símbolo USE para remitir desde un término ilegal (no usado) en el tesoro hacia aquel considerado correcto.

Ej.: SOBREPoblACION USE SUPERPOBlACION

- b) Utilizando el símbolo UF (used for) en español "usado por".

Ej.: SUPERPOBlACION

UF: SOBREPoblACION

SUPERPOBlAMIENTO

Aquí, como la relación es de carácter recíproco, bajo el término legal, esto es, "SUPERPOBlACION" aparecerán precedidos por el símbolo UF aquellos descriptores considerados como sinónimos del primero.

Ej.: SUPERPOBlACION

UF: SOBREPoblACION

SUPERPOBlAMIENTO

1. 2. Relaciones jerárquicas (BT; NT)

Son las relaciones de superioridad y subordinación entre los descriptores, utilizándose:

1. 2. 1. BT ("broader term") en español "término genérico" ^{para} indicar el descriptor que representa un concepto del cual forma parte el descriptor subordinado, y,

1.2.2. NT ('Narrower term') en español "término específico" para indicar el término que designa los términos subordinados.

Ej.: MIGRACION SELECTIVA

BT: MIGRACION

NT: INMIGRACION SELECTIVA

Como estas relaciones son de carácter recíproco, bajo MIGRACION se encuentra:

MIGRACION

NT: MIGRACION SELECTIVA

1.3. Relaciones de asociación (RT)

Estas se utilizan para indicar relaciones entre conceptos que están estrechamente relacionados, sin ser estas relaciones de carácter jerárquico ni de equivalencia. La relación de asociación se representa por medio del símbolo RT ("related term") en español, "términos relacionados".

Ej.: PARIENTE

RT: AFINIDAD

GRADO DE PARENTESCO

Los tesauros incluyen también otras aclaraciones como son las Notas de Alcance (SN) que tienen por objeto informar acerca del sentido en que se utiliza el concepto en el tesauro, la que va precedida del símbolo SN ("Scope Note"). Estas no son definiciones propiamente tales sino que indican el uso del término en el Tesauro. Es decir, éstas son sólo notas aclaratorias que no forman parte del descriptor.

Ej.: MORTALIDAD INFANTIL

SN: Mortalidad de los niños nacidos vivos que no han alcanzado su primer aniversario

Descripción del "TESAURO MULTILINGUE POPINS"

En una reunión acerca de información sobre población celebrada en México en 1975, se acordó confeccionar un tesauro multilingue sobre población y temas afines que se publicaría en 3 idiomas: español, francés e inglés. Es lo que se conoce con el nombre de tesauro POPINS. En esa misma fecha, CELADE fue designado, junto con el Colegio de México, representantes de las instituciones de habla hispana en el Grupo de Trabajo coordinado por Jean Viet, experto en la construcción de tesauros.

Se han celebrado 3 reuniones de trabajo que han permitido estructurar el actual documento de trabajo que es un borrador final del Tesoro POPINS que sería publicado en diciembre de este año.

Físicamente, este vocabulario consta de dos partes que se interrelacionan y en que cada descriptor se encuentra identificado por un número de 6 dígitos que representa la faceta a la que pertenece.

- 2.1. Índice alfabético. En él los términos aparecen ordenados de acuerdo a los elementos que lo forman, apareciendo estos tantas veces como elementos tengan. Por ejemplo, el descriptor 'AGLOMERACION URBANA' aparecerá alfabetizado en la A, bajo 'Aglomeración' y en la U, bajo 'Urbana'. En ambos casos, habrá indicación de la faceta a la que pertenece el descriptor, la que remite al Índice por facetas.
- 2.2. Índice por facetas. En este índice los descriptores se presentan agrupados de acuerdo a su afinidad temática. Los temas se clasificaron en 16 campos de interés (campos semánticos) para el área de población (fecundidad, migración, mortalidad, etc.) numerados del 01 al 18, los cuales se han dividido en "facetas," identificadas por los cuatro primeros dígitos del número de identificación. Dentro de las facetas pueden también distinguirse sub-facetas que se diferencian de las anteriores a través de los últimos dos dígitos del número de identificación.

Es así como bajo cada faceta y subfaceta se presentan los términos agrupados por su afinidad, permitiendo seleccionar el término más adecuado para describir el concepto por el cual se desea indizar, es decir, el o los descriptores por los cuales será posible recuperar el contenido intelectual del documento. En el índice se presentan bajo cada término, todas las relaciones que tiene el descriptor con otros términos en el tesoro. Es decir, un índice polijerárquico en el sentido que las relaciones establecidas entre los términos pueden ser entre descriptores pertenecientes a diferentes facetas.

3. Uso del Tesauro Plurilingue POPINS

3.1. Introducción

En un sistema de información, es de vital importancia la consistencia en la información; es decir, debe tenderse a que los descriptores asignados a un mismo tipo de documentos sean los mismos, sin importar quien sea, dentro del grupo de compendadores/ indizadores/ la persona que realice la indización. En el caso de DOCPAL, la consistencia es aún de mayor importancia y necesidad puesto que, por una parte, en el futuro se procederá al intercambio de cintas con otras instituciones similares y, por otra parte, el producto del análisis de los documentos ingresados al sistema constituye el Índice temático de la Revista DOCPAL Resúmenes.

Como una ayuda al indizador para alcanzar esta consistencia, que posteriormente permitirá además una mejor recuperación de información, DOCPAL preparó guías que se incluyen en páginas III #76 de su Manual.

3.2. Indización de documentos

En la selección de los descriptores se deben seguir las siguientes pautas y reglas:

- a) Examinar el resumen y extraer los términos que describen aquellos conceptos de importancia en el documento.
- b) Examinar el documento y seleccionar los términos que describen su contenido. Ver cuidadosamente el título, revisar el índice, prólogo, introducción, conclusiones, resumen del autor, etc. Leer unos párrafos para comprobar si los conceptos seleccionados en los pasos a) y b) corresponden a lo tratado en el texto.
- c) Convertir estos términos en descriptores usando el Tesauro Multilingue POPINS. Para esto, consultar el listado alfabético del tesauro y seleccionar aquellos descriptores que más se aproximen al concepto expresado. Comprobar su alcance temático dirigiéndose a la faceta que le corresponde.

d) Utilizar siempre el descriptor que describe en forma más específica el concepto, es decir, el término específico (NT) que mejor aplica. Tratar de evitar descriptores de carácter general tales como

⟨FECUNDIDAD⟩ ⟨MORTALIDAD⟩ etc.

e) Evaluar los descriptores seleccionados. Con este fin, considerar si el usuario/^{potencial}del documento utilizaría o no estos términos al solicitar información sobre el tema tratado en el documento.

Si la respuesta es negativa se recomienda sustituir los descriptores por otros que describan en forma más precisa el concepto tratado.

f) Una buena indización representa un equilibrio muy delicado.

Al asignar un descriptor, es necesario asegurarse que el concepto representado es de suficiente importancia en el texto como para dejar satisfecho a un usuario que recupere el documento. La sobreindización a menudo es más dañina que la sub-indización, ya que produce recuperaciones falsas. Sin embargo, si es apropiado, se pueden asignar a un mismo documento, tanto descriptores genéricos como específicos. En otras palabras, el indizador debe tener en cuenta al analizar un documento, básicamente las siguientes normas:

i) Exhaustividad: cada concepto de importancia en un documento debe estar representado al menos por un descriptor. Se considera que un concepto es de importancia si corresponde a un criterio potencial de recuperación.

ii) Especificidad: los conceptos que se usarán en la indización deben representar el descriptor más específico de que se dispone en el tesoro. Agregar conceptos más generales, sólo en aquellos casos en que se considera de utilidad para su posible recuperación.

- g) Verificar que los descriptores con asteriscos (no más de 4 y normalmente 2 ó 3) estén pre-coordinados ya que permitirán la recuperación del documento a través del Índice Temático de la Revista. Recordar que son los descriptores que aparecen siempre en la Revista y por tanto, importa que haya consistencia.
- h) Revisar también qué documentos de temas frecuentemente tratados (proyecciones, etc.) estén indizados utilizando los mismos términos (comparar con documentos similares en los números anteriores de la Revista).

4.3. Manejo del Tesouro POPINS

Al indizar un documento, la búsqueda de los términos que describen los conceptos tratados normalmente, tiene que comenzar por el Índice alfabético, para determinar si el término es legal, es decir, que ha sido aceptado como el más acertado por el Tesouro para normalizar el ingreso y recuperación de información.

Una vez localizado el término que represente parcial o totalmente el concepto a describir, debe irse necesariamente al Índice por facetas para verificar si el término es el más adecuado y el más específico para describir el concepto.

Si en el Índice por facetas se descubre que el término seleccionado en el Índice alfabético no es el adecuado y otro término (NT, BT o RT) parece mejor, debe volverse nuevamente al Índice alfabético y de ahí al Índice por facetas, repitiéndose este proceso tantas veces como sea necesario.

Sin embargo, es todavía posible considerar que el Tesouro no incluye ningún término realmente adecuado para describir fielmente la materia, o concepto específico, que el autor expresa en su documento. Si este es el caso, el compendiador deberá hacerlo saber al personal de DOCPAL, proponiendo un término adecuado. Se sugiere dejar además nota adicional en la HDB correspondiente.

Las sugerencias ^{de inclusión} de nueva terminología serán cuidadosamente consideradas por DOCPAL para enriquecer el Tesouro POPINS y convertirlo en un instrumento eficaz para analizar información sobre población en América Latina.