

## CAMPOS DE BÚSQUEDA Y CONTROL DEL VOCABULARIO

La búsqueda por materias se realiza generalmente a través de dos tipos muy diferentes de campos:

a) Campos de texto libre: título del documento, resumen, el texto completo (en bases de datos que lo incluyan).

b) Campos formados por entradas de un índice alfabético de materias. Son campos sin la estructura sintáctica de un texto, formados por grupos nominales de una sola palabra (por ejemplo "Filosofía") o más de una ("Filosofía de la Ciencia"), que definen un tema de búsqueda por sí mismas o en combinación con otras entradas del índice. Estas entradas se denominan términos de indización y puede haber diferentes tipos:

-Encabezamientos y subencabezamientos de materia. Es el sistema utilizado en los catálogos de las bibliotecas para describir las materias tratadas en un documento. Se trata normalmente de términos procedentes de un lenguaje controlado establecido previamente (Listas de encabezamientos de materia). Cada entrada del índice se compone de una materia principal (encabezamiento), que puede estar acompañada de facetas de materia, lugar, tiempo o forma (subencabezamientos). Por ejemplo, la entrada 'Mujeres - Historia - Recursos bibliográficos - Cádiz' sería un único encabezamiento con tres subencabezamientos.

-Palabras clave: términos simples o compuestos extraídos de los documentos, sin que estén presentes en ningún lenguaje controlado preestablecido. En un sistema de indización por palabras clave pueden utilizarse tantas formas de expresar el mismo concepto como sean utilizadas por los autores de los documentos originales, por ejemplo "Política ambiental" y "Política de medio ambiente".

-Descriptor: son términos simples o compuestos extraídos de un lenguaje controlado que traducen los conceptos utilizados por los autores, para evitar la ambigüedad del lenguaje natural. Estos lenguajes documentales controlados pueden ser simples léxicos o listados alfabéticos o tratarse de tesauros con una estructura jerárquica entre sus términos. Los productores de una base de datos pueden elaborar su propio tesoro o adoptar uno elaborado por otra institución.

-Otros campos de descripción del contenido. En algunas bases de datos existen diferentes campos destinados a los términos de indización para facilitar el control del vocabulario y reducir las ambigüedades existentes entre nombres comunes (campo de descriptores), nombres propios (campo de identificadores), topónimos (identificadores geográficos), o cualquier otro definido para resolver necesidades específicas de información (compuestos químicos, coordenadas geográficas, legislación y jurisprudencia, periodo histórico). El contenido de estos campos también puede tener un control del vocabulario a través de listados de autoridades. Además puede existir un campo para reflejar aquellos términos que no se han incorporado aún al tesoro (campo de palabras candidatas).

La búsqueda debe adaptarse a las características de esta estructura de campos, adoptando dos tipos de estrategias complementarias:

- a) Cuando se realiza una búsqueda en campos de texto libre (título, resumen,...) o de palabras clave (lenguajes no controlados), se deben tener en cuenta las diferentes formas en las que puede aparecer un concepto en las referencias bibliográficas: sinónimos, conceptos afines, variaciones gramaticales,...
- b) Por el contrario, cuando la búsqueda se realiza en campos con control del vocabulario, debemos intentar localizar la forma seleccionada en el lenguaje controlado utilizado por los productores de la base de datos. Para ello, es conveniente consultar previamente el tesoro o léxico de indización o bien realizar la búsqueda a través de la visualización de los índices de cada campo.

## CAMPOS DE BUSQUEDA Y CONTROL DEL VOCABULARIO

La búsqueda por materias se realiza generalmente a través de dos tipos muy diferentes de campos:

- a) Campos de texto libre: título del documento, resumen, el texto completo (en bases de datos que lo incluyan).
- b) Campos formados por entradas de un índice alfabético de materias. Son campos sin la estructura sintáctica de un texto, formados por grupos nominales de una sola palabra (por ejemplo "Filosofía") o más de una ("Filosofía de la

Ciencia”), que definen un tema de búsqueda por sí mismas o en combinación con otras entradas del índice. Estas entradas se denominan términos de indización y puede haber diferentes tipos:

- Encabezamientos y subencabezamientos de materia. Es el sistema utilizado en los catálogos de las bibliotecas para describir las materias tratadas en un documento. Se trata normalmente de términos procedentes de un lenguaje controlado establecido previamente (Listas de encabezamientos de materia). Cada entrada del índice se compone de una materia principal (encabezamiento), que puede estar acompañada de facetas de materia, lugar, tiempo o forma (subencabezamientos). Por ejemplo, la entrada ‘Mujeres – Historia – Recursos bibliográficos – Cádiz’ sería un único encabezamiento con tres subencabezamientos.
- Palabras clave: términos simples o compuestos extraídos de los documentos, sin que estén presentes en ningún lenguaje controlado preestablecido. En un sistema de indización por palabras clave pueden utilizarse tantas formas de expresar el mismo concepto como sean utilizadas por los autores de los documentos originales, por ejemplo “Política ambiental” y “Política de medio ambiente”.
- Descriptores: son términos simples o compuestos extraídos de un lenguaje controlado que traducen los conceptos utilizados por los autores, para evitar la ambigüedad del lenguaje natural. Estos lenguajes documentales controlados pueden ser simples léxicos o listados alfabéticos o tratarse de tesauros con una estructura jerárquica entre sus términos. Los productores de una base de datos pueden elaborar su propio tesoro o adoptar uno elaborado por otra institución.
- Otros campos de descripción del contenido. En algunas bases de datos existen diferentes campos destinados a los términos de indización para facilitar el control del vocabulario y reducir las ambigüedades existentes entre nombres comunes (campo de descriptores), nombres propios (campo de identificadores), topónimos (identificadores geográficos), o cualquier otro definido para resolver necesidades específicas de información (compuestos químicos, coordenadas geográficas, legislación y jurisprudencia, periodo histórico). El contenido de estos campos también puede tener un control del vocabulario a través de listados de

autoridades. Además puede existir un campo para reflejar aquellos términos que no se han incorporado aún al tesoro (campo de palabras candidatas).

La búsqueda debe adaptarse a las características de esta estructura de campos, adoptando dos tipos de estrategias complementarias:

- a) Cuando se realiza una búsqueda en campos de texto libre (título, resumen,...) o de palabras clave (lenguajes no controlados), se deben tener en cuenta las diferentes formas en las que puede aparecer un concepto en las referencias bibliográficas: sinónimos, conceptos afines, variaciones gramaticales,...
- b) Por el contrario, cuando la búsqueda se realiza en campos con control del vocabulario, debemos intentar localizar la forma seleccionada en el lenguaje controlado utilizado por los productores de la base de datos. Para ello, es conveniente consultar previamente el tesoro o léxico de indización o bien realizar la búsqueda a través de la visualización de los índices de cada campo.