

DATOS ABIERTOS

Introducción

El Senado indiza describiendo con palabras clave todas sus colecciones documentales y bibliográficas utilizando para ello los términos del [Tesauro del Senado](#), publicado en 1993. Las colecciones quedaban así íntegramente indizadas, pero no estaban clasificadas en grandes bloques de conocimientos.

Para conseguir este objetivo y no hacerlo de forma manual, documento a documento, se proyectó que se podía aprovechar la asignación de los términos del Tesauro a los documentos y conseguir que los contenidos se clasificaran [en 83 Áreas Temáticas \(AATT\), agrupadas en 22 grandes Bloques Temáticos \(BBTT\)](#) automáticamente. Es decir, lograr agrupaciones clasificatorias, construidas a partir del etiquetado con descriptores, asignándoles a las colecciones las AATT y BBTT de forma automática, mediante un desarrollo informático.

Por tanto, el sistema se basa en un trabajo intelectual realizado por los servicios documentales del Senado de agrupación primaria de términos del Tesauro en grandes bloques de conocimientos.

Creemos que este sistema clasificatorio, multidisciplinar y transversal creado en el Senado, que ha supuesto un gran esfuerzo de coordinación interna, de gestión del conocimiento y una gran carga de trabajo, puede ser útil al resto de usuarios de sistemas de indización, y especialmente a los que utilizan el Tesauro Eurovoc, por lo que ponemos a su disposición en formato .XML, en el [Portal de Datos Abiertos](#), esta clasificación de términos del Tesauro en Áreas y Bloques temáticos, para que pueda ser utilizado libremente y aplicado por quien lo desee.

Estructura

Este documento está estructurado de la siguiente manera

- Cabecera

- Bloques temáticos (número 22)

- Título del bloque (descripción)

- Áreas temáticas (número 83)

- Título del área (descripción)

- Materias que comprende el área (notas)

Áreas relacionadas (relacionadas) (n número)
Título del área relacionada (descripción)
Términos del Tesouro asociado al Área (término) (n número)
Título del Término (descripción)
Microtesauro
Código y título del microtesauro al que pertenece

```
<bloques>  
  <bloques num="1">  
    <descripcion>TÍTULO DEL BLOQUE</descripcion>  
    <areas>  
      <area num="1">  
        <descripcion>TÍTULO DEL ÁREA</descripcion>  
        <notas>Enumeración de las materias que comprende.</notas>  
        <relacionadas>  
          <relacionada num="1">  
            <descripcion>Otras Áreas con las que está relacionada</descripcion>  
          </relacionada num="n">  
        </relacionadas>  
        <terminos>  
          <termino num="1">  
            <descripcion>Título del término</descripcion>  
            <microtesauros>  
              <microtesauro num="1">  
                <descripcion>0000 Título del microtesauros al que pertenece</descripcion>  
              </microtesauro>  
            </microtesauros>  
          </termino num="1">  
        </terminos>  
      </area num="1">  
    </areas>  
  </bloques num="1">  
</bloques>
```

</bloques>
</bloque>
</areas>
</area>
</terminos>
</termino>

Algunas aclaraciones para comprender mejor el contenido:

- 1) Aunque el Tesauro del Senado está compuesto por el Tesauro Eurovoc, versión 4.3 + una lista complementaria de términos elaborada por el Senado + la relación de municipios de España facilitada por el Instituto Nacional de Estadística (INE), sólo clasifican en AA.TT **algunos** de los términos del Eurovoc y **algunos** de la lista complementaria.
- 2) Sólo aparecen términos contenidos hasta la versión 4.3.
- 3) La estructura del documento corresponde a BBTT> AATT>Términos del Tesauro.
- 4) Si necesita la traducción a idioma inglés de los BBTT y AATT solicítelas contactando con sugerenciasbuscador@senado.es