

Boletín oficial del Estado: análisis de metadatos, detección de errores y recomendaciones de mejora

The Spanish *Boletín oficial del Estado*: Metadata analysis, error detection and recommendations for improvement

Juan-Carlos Bailón-Elvira; Manuel-Jesús Cobo-Martín; Antonio-Gabriel López-Herrera

Cómo citar este artículo:

Bailón-Elvira, Juan-Carlos; Cobo-Martín, Manuel-Jesús; López-Herrera, Antonio-Gabriel (2020). "Boletín oficial del Estado: análisis de metadatos, detección de errores y recomendaciones de mejora". *El profesional de la información*, v. 29, n. 2, e290226.

<https://doi.org/10.3145/epi.2020.mar.26>

Artículo recibido el 08-08-2019
Aceptación definitiva: 11-12-2019



Juan-Carlos Bailón-Elvira ✉
<https://orcid.org/0000-0002-4745-9121>

Universidad de Granada
Departamento de Ciencias de la
Computación e Inteligencia Artificial
Daniel Saucedo Aranda, s/n.
18071, Granada, España
bailone@correo.ugr.es



Manuel-Jesús Cobo-Martín
<https://orcid.org/0000-0001-6575-803X>

Universidad de Cádiz
Escuela Superior de Ingenierías
Departamento de Ingeniería Informática
Avda. Universidad de Cádiz, 10.
11519 Puerto Real (Cádiz), España
manueljesus.cobo@uca.es



Antonio-Gabriel López-Herrera
<https://orcid.org/0000-0001-8424-275X>

Universidad de Granada
Departamento de Ciencias de la
Computación e Inteligencia Artificial
Daniel Saucedo Aranda, s/n.
18071 Granada, España
lopez-herrera@decsai.ugr.es

Resumen

El acceso a la información pública (la que generan las instituciones públicas) es un derecho reconocido tanto por la legislación europea (*Unión Europea*, 2001) como por la española (*España*, 1992; 1997; 1999; 2013). Los organismos del estado español están obligados a poner a disposición del ciudadano de manera gratuita y transparente todo lo que generan a menos que existan motivos de seguridad nacional. Muchos organismos estatales, autonómicos y locales han optado por crear sistemas de información o plataformas de información pública en forma de boletines informativos. Este artículo tiene como finalidad conocer cómo se encuentra a nivel documental el principal boletín informativo del estado español, el *Boletín oficial del Estado* (BOE) estudiando los metadatos de los documentos publicados y analizado en profundidad los que se utilizan para la descripción documental. Los resultados reflejan la ausencia de descriptores documentales en más del 89% del total del corpus documental del BOE. En los últimos años esta cifra apenas supera el 50%. El trabajo finaliza esbozando algunas líneas de mejora.

Palabras clave

Recuperación de información; Metadatos; Metaanálisis; Análisis documental; Recomendaciones; Difusión de información; Sistemas de alertas; *Boletín oficial del Estado*; BOE; España.

Financiación

Este trabajo está soportado por el *Ministerio de Economía y Competitividad* (Referencia TIN2016-75850-R), España.

Abstract

Access to public information (that generated by public institutions) is a right recognized by both European (*Unión Europea*, 2001) and Spanish legislation (*España*, 1992; 1997; 2013). The bodies of the Spanish State must to make available to the citizens, free of charge and in a transparent way, all the documents that they generate, unless there are national security reasons. Many national, regional and local governments have created information systems or public information platforms in the form of newsletters or gazettes. The purpose of this paper is to find out how is, on a documentary level, the main informative gazette of Spain, the *Boletín oficial del Estado (BOE)*. The metadata of the published documents are studied and the metadata used for documentary description are analyzed in a deeper way. The results reflect the absence of documentary descriptors in more than 89% of the total documentary *BOE'* corpus. In recent years this figure has barely exceeded 50%. The paper ends with a few recommendations for improvement.

Keywords

Information retrieval; Metadata; Meta-analysis; Documentary analysis; Gazette; Recommendations; Information dissemination; Alert systems; *Boletín oficial del Estado; BOE; Spain.*

1. Introducción

Desde el momento en que se constituye un organismo público, éste empieza a generar leyes, normas, resoluciones, licitaciones, anuncios o concursos, entre otros. Toda esta actividad es reflejada en documentos de carácter público, siempre y cuando no interfieran con motivos de seguridad nacional. Los documentos son custodiados por las administraciones y pueden ser consultados por los ciudadanos.

El acceso a esta documentación no siempre ha sido posible o no ha estado garantizada por ley. No fue hasta el año 1992 con la publicación de la *Ley 30/1992 (España, 1992)* en la que se reconoció el derecho de acceso (ver artículo 35.h). A esta ley le siguieron otras como la *Ley 6/1997 de organización y funcionamiento de la Administración General del Estado (España, 1997)* que reconocía el derecho del ciudadano a tener acceso a información sobre procesos administrativos. En 2013 se aprobó la *Ley 19/2013 de transparencia, acceso a la información pública y buen gobierno (España, 2013)* que defendía el derecho del ciudadano al acceso a documentos publicados por administraciones públicas.

Este libre acceso está garantizado en España tanto por la *Constitución Española* como por las leyes *30/1992 (España, 1992)*, *6/1997 (España, 1997)* o *19/2013 (España, 2013)* y se materializan en una serie de boletines como el *Boletín oficial del Estado*, el *Boletín oficial del registro mercantil*, los boletines oficiales de las comunidades autónomas y sus parlamentos, entre otros.

No sólo es importante la publicación de los documentos generados por instituciones públicas sino por supuesto también, que el acceso a estos sea efectivo, eficiente y permanente. Es necesario que los documentos estén bien formados, se publiquen íntegros, sean totalmente legibles por humanos y por máquinas, y estén bien descritos a nivel documental, de manera que el acceso sea rápido, preciso, exhaustivo y completo.

Este artículo tiene como finalidad conocer cómo se encuentra a nivel documental la principal fuente de información institucional del estado, el *Boletín oficial del Estado (BOE)*. <https://www.boe.es>

Página inicial de la web del *Boletín oficial del Estado* <https://www.boe.es>

Se estudian los metadatos asociados a los documentos publicados en el *BOE* entre el día 1 de septiembre de 1960 y el 30 de junio de 2019. Especialmente se analizan con mayor énfasis los metadatos que pudieran ser usados para su descripción documental, como son las alertas, materias y materias CPV (*Common procurement vocabulary, Unión Europea, 2008*). La idea es conocer la situación actual del *BOE* a nivel de descripción documental, detectar los principales problemas de acceso y recuperación documental, y finalmente la propuesta de una serie de recomendaciones.

El presente trabajo se divide de la siguiente manera. En la segunda sección se introduce la estructura del *BOE*. La sección tercera expone la metodología empleada. En la sección cuarta se presentan los principales resultados del estudio. Finalmente se exponen las principales conclusiones obtenidas a partir de los resultados.

2. Estructura del Boletín oficial del Estado

Se divide en seis secciones:

- I. Disposiciones generales
- II. Autoridades y personal
 - II.A. Nombramientos, situaciones e incidencias
 - II.B. Oposiciones y concursos
- III. Otras órdenes
- IV. Administración de justicia
- V. Anuncios
 - V.A. Contratación del sector público
 - V.B. Otros anuncios oficiales
 - V.C. Anuncios particulares
- VI. TC (Resoluciones del *Tribunal Constitucional*)

Los documentos publicados en el *BOE* se pueden consultar de forma gratuita y abierta en diferentes formatos electrónicos (pdf, epub y xml). Algunos de los documentos están también disponibles en varios idiomas de entre los cinco oficiales del estado (español, catalán, euskera, valenciano y gallego). Estos documentos pueden consultarse por medio de:

1. Un buscador en el que el usuario introduce su consulta a través de campos preestablecidos como la fecha de publicación, número del documento, el departamento que lo publica, etc.

2. Suscripción a un sistema de alertas gratuito llamado *Mi BOE*. En base a los intereses declarados por el usuario mediante una serie de términos preestablecidos, éste recibe en su correo electrónico los documentos que concuerdan con sus intereses. El listado de estos términos es grande e incluye materias como “Ayudas”, “Becas”, “Cambios de divisas”, “Convenios colectivos”, “Planes de estudios”, “Sentencias del Tribunal Constitucional”, entre otros.

https://www.boe.es/mi_boe

3. Descargando los boletines que se deseen por medio de un script programado en lenguaje php y disponible en:

https://www.boe.es/datosabiertos/ejemplo_script_boe.php

Con este programa se puede descargar la versión xml del diario y todos los documentos en versión pdf asociados a él.

A modo de ejemplo, en la figura 1 se muestra un documento cualquiera (BOE-A-2018-14948 perteneciente a la sec-

Figure 1 shows two side-by-side views of a document from the Boletín Oficial del Estado (BOE). The left view is a PDF page, and the right view is an XML representation of the same document.

PDF View (Left):

- Header: BOLETÍN OFICIAL DEL ESTADO, Núm. 264, Jueves 1 de noviembre de 2018, Sec. I. Pág. 106850.
- Section: I. DISPOSICIONES GENERALES, COMUNIDAD AUTÓNOMA DE CANARIAS.
- Article: 14948 Ley 2/2018, de 28 de septiembre, de régimen jurídico de la Inspección Técnica de Vehículos en Canarias.
- Text: Sea notorio a todos los ciudadanos que el Parlamento de Canarias ha aprobado y yo, en nombre del Rey y de acuerdo con lo que establece el artículo 12.8 del Estatuto de Autonomía, promulgo y ordeno la publicación de la Ley 2/2018, de 28 de septiembre, de Régimen Jurídico de la Inspección Técnica de Vehículos en Canarias.
- Section: PREÁMBULO.
- Text: La competencia relativa a la prestación del servicio de inspección técnica de vehículos (ITV) fue transferida en el año 1984 por el Estado a la Comunidad Autónoma de Canarias, la cual asumió directamente la construcción y explotación de diversas estaciones de ITV en Tenerife, Gran Canaria, La Palma, Lanzarote y Fuerteventura, además de proveerse de una estación móvil para cubrir el servicio en las islas que no disponían de estación de ITV.

XML View (Right):

```
<documento fecha_actualizacion="20190617182602">
  <metadatos>
    <identificador>BOE-A-2018-14948</identificador>
  </metadatos>
  <titulo>
    Ley 2/2018, de 28 de septiembre, de régimen jurídico de la Inspección Técnica de Vehículos en Canarias.
  </titulo>
  <diario codigo="BOE">Boletín Oficial del Estado</diario>
  <diario_numero>264</diario_numero>
  <seccion I</seccion>
  <subseccion>
    <departamento codigo="8030">Comunidad Autónoma de Canarias</departamento>
    <rango codigo="1300">Ley</rango>
    <numero_oficial>2/2018</numero_oficial>
    <fecha_disposicion>20180928</fecha_disposicion>
    <fecha_publicacion>20181101</fecha_publicacion>
    <fecha_vigencia>20181010</fecha_vigencia>
    <fecha_derogacion>
    <letra_imagen>A</letra_imagen>
    <pagina_inicial>106850</pagina_inicial>
    <pagina_final>106859</pagina_final>
    <suplemento letra_imagen>
    <suplemento pagina_inicial>
    <suplemento pagina_final>
    <estatus_legislativo>L</estatus_legislativo>
    <origen_legislativo codigo="2">Autonómico</origen_legislativo>
    <estado_consolidacion codigo="3">Finalizado</estado_consolidacion>
    <judicialmente anulada>N</judicialmente anulada>
    <vigencia_agotada>N</vigencia_agotada>
    <estatus_derogacion>N</estatus_derogacion>
    <url_apun>diario_boe/epub.php?id=BOE-A-2018-14948</url_apun>
    <url_pdf>boe/dias/2018/11/01.pdf#BOE-A-2018-14948.pdf</url_pdf>
    <url_pdf_catalan>
    <url_pdf_euskera>
    <url_pdf_gallego>
    <url_pdf_valenciano>
    <url_eli>https://www.boe.es/eli/es-cn/2018/09/28/2</url_eli>
  </metadatos>
  <metadatos-elib</metadatos-elib>
  <analisis>
    <notas></notas>
  </analisis>
  <materias>
    <materia codigo="397" orden="">Autorizaciones</materia>
    <materia codigo="599" orden="">Canarias</materia>
    <materia codigo="4289" orden="">Inspección Técnica de Vehículos</materia>
    <materia codigo="5281" orden="">Organización de las Comunidades Autónomas</materia>
  </materias>
  <alertas>
    <alerta codigo="109" orden="">Derecho Administrativo</alerta>
    <alerta codigo="121" orden="">Organización de la Administración</alerta>
    <alerta codigo="132" orden="">Transportes y tráfico</alerta>
  </alertas>
</documento>
```

Figura 1. Extracto de los detalles del documento BOE-A-2018-14948 en sus versiones pdf (izquierda) y xml (derecha).

ción I) en su versión pdf (a la izquierda una miniatura de la primera página) y un extracto de los metadatos del mismo documento en su versión xml (a la derecha). Un análisis de los metadatos de este documento permite ver que está documentalmentemente descrito a nivel de contenido mediante los metadatos “materia” y “alerta”. El documento trata de las materias: “Autorizaciones”, “Canarias”, “Inspección técnica de vehículos”, “Organización de las Comunidades Autónomas”; mientras que este documento está vinculado a las alertas: “Derecho administrativo”, “Organización de la administración”, “Transportes y tráfico”. Seleccionando estos descriptores temáticos en el sistema de alertas *Mi BOE*, el usuario/ciudadano sería notificado automáticamente de la publicación del mismo, y de otros que tuvieran la misma descripción.

En la tabla 1 se muestra el listado completo de los metadatos que se pueden consultar y obtener para las diferentes secciones del *BOE*. En el apartado de Metodología se explica en detalle cómo acceder a ellos.

Tabla 1. Listado completo de metadatos

Etiqueta / metadato	Descripción	Secciones
departamento	Departamento que publica el documento	Todas
diario	Diario al que pertenece el documento	Todas
diario_numero	Número del diario al que pertenece el documento	Todas
fecha_actualizacion	Última fecha de modificación del documento	Todas
fecha_publicacion	Fecha de publicación del documento	Todas
Identificador	Identificador único del documento	Todas
letra_imagen	Indica la letra de la imagen	Todas
numero_oficial	Número oficial asignado al documento	Todas
pagina_final	Página en la que termina el documento	Todas
pagina_inicial	Página en la que empieza el documento	Todas
seccion	Sección a la que pertenece el documento	Todas
texto	Texto íntegro del documento	Todas
titulo	Título del documento	Todas
url_pdf	Url de acceso al formato pdf	Todas
subseccion	Subsección a la que pertenece el documento	I, II, III, V, TC
alerta	Alerta que contiene el documento	I, II, III, TC
estado_consolidacion	Indica si el texto está consolidado	I, II, III, TC
estatus_derogacion	Indica si está derogado el documento	I, II, III, TC
estatus_legislativo	Estatus legislativo del documento	I, II, III, TC
fecha_derogacion	Fecha en la que se derogó el documento	I, II, III, TC
fecha_vigencia	Fecha en la que entra en vigor el documento	I, II, III, TC
judicialmente_anulada	Indica si está anulado el documento	I, II, III, TC
materia	Materia que contiene el documento	I, II, III, TC
nota	Notas del documento	I, II, III, TC
origen_legislativo	Origen legislativo del documento	I, II, III, TC
rango	Rango al que pertenece el documento	I, II, III, TC
referencia_anterior	Indica si el documento referencia a documentos anteriores a su publicación	I, II, III, TC
referencia_posterior	Indica si el documento referencia a documentos posteriores a su publicación	I, II, III, TC
suplemento_letra_imagen	Indica si existen suplementos a la letra de la imagen	I, II, III, TC
suplemento_pagina_final	Indica si existen suplementos a página final	I, II, III, TC
suplemento_pagina_inicial	Indica si existen suplementos a la página inicial	I, II, III, TC
url_epub	Url de acceso al formato epub	I, II, III, TC
url_pdf_catalan	Url de acceso al formato pdf en catalán	I, II, III, TC
url_pdf_euskera	Url de acceso al formato pdf en euskera	I, II, III, TC
url_pdf_gallego	Url de acceso al formato pdf en gallego	I, II, III, TC
url_pdf_valenciano	Url de acceso al formato pdf en valenciano	I, II, III, TC
vigencia_agotada	Indica si la vigencia del documento está agotada	I, II, III, TC
numero_anuncio	Número oficial asignado al anuncio	IV, V
ambito_geografico	Ámbito geográfico al que aplica el anuncio	V
fecha_apertura_ofertas	Fecha en la que se inician las ofertas	V
fecha_presentacion_ofertas	Fecha de presentación de ofertas	V
importe	Importe del anuncio	V
materias_cpv	Materias según códigos CPV (<i>Common procurement vocabulary</i> - Vocabulario común de contratación pública)	V
modalidad	Modalidad del anuncio o sentencia	V
observaciones	Texto con aclaraciones	V
precio	Precio del anuncio	V
procedimiento	Tipo de procedimiento del anuncio o sentencia	V
tipo	Tipo de anuncio o sentencia	V
tramitacion	Tipo de tramitación del anuncio o sentencia	V

Además de los metadatos listados en la tabla 1, en algunas secciones del *BOE* se pueden encontrar metadatos adicionales, como son “metadata_eli” y “url_eli”. ELI (*European legislation identifier*) es un sistema que permite el acceso online a la legislación en un formato normalizado. Es una iniciativa adoptada conjuntamente por los países y las instituciones de la UE (*Unión Europea*, 2017). Este sistema está siendo implementado por la agencia del *BOE* para aplicar una capa semántica a los documentos por medio de una ontología.

El sistema de información del *BOE* es el encargado de poner a disposición del ciudadano todo lo referente a normas, leyes, resoluciones judiciales, convocatorias, concursos públicos, etc., de aplicación en todo el estado español

3. Metodología

Dado que el objetivo de este trabajo es conocer el estado de descripción documental que presentan los documentos del *BOE*, necesitamos acceder a todos los metadatos, y especialmente a los que aportan descripción de contenido (“materia”, “materia_cpv” y “alerta”). Se llevaron a cabo los siguientes pasos:

1. Descarga de datos: ninguno de los tres métodos de consulta que facilita oficialmente el *BOE* permite el acceso íntegro a todos los metadatos disponibles. Se diseñó y programó en lenguaje java una araña web para iterar por todas las fechas de publicación (desde 1 de septiembre de 1960 hasta el 30 de junio de 2019) y acceder a la versión xml de cada documento. De los tres formatos disponibles (pdf, epub y xml) sólo el formato xml proporciona la estructura necesaria para realizar el análisis de metadatos. Para procesarlos correctamente fue necesario conocer el documento esquema asociado.

En nuestro caso usamos la versión xsd de este esquema. Un archivo xsd es aquel en el que se define la metainformación de un xml declarando los campos obligatorios, las opciones, cardinalidades, así como el tipo de datos que contendrán, lo que permite validar la estructura de los ficheros xml asociados.

El *BOE* sólo ofrece el archivo xsd para los sumarios y no para el resto de los documentos (resto de secciones), para los cuales fue necesario crear xsd *ad-hoc*. Debido a la diferente estructura de metadatos entre las secciones, fue necesario usar varios archivos xsd. A modo de ejemplo, en la figura 2 se muestran un fragmento del fichero xsd (en la parte superior) junto con su fragmento del fichero xml (parte inferior), ambos correspondientes a un fragmento de un documento cualquiera perteneciente a la Sección I.

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="documento">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="metadatos">
          <xs:complexType>
            <xs:sequence>
              <xs:element type="xs:string" name="identificador"/>
              <xs:element type="xs:string" name="titulo"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```
<documento fecha_actualizacion="20181219102601">
  <metadatos>
    <identificador>BOE-A-2018-14948</identificador>
    <titulo>Ley 2/2018, de 28 de septiembre, de régimen jurídico de la Inspección
Técnica de Vehículos en Canarias.</titulo>
    <diario codigo="BOE">Boletín Oficial del Estado</diario>
    </diario_numero>
    <seccion>1</seccion>
    <subseccion/>
    <departamento codigo="8030">Comunidad Autónoma de Canarias</departamento>
    <rango codigo="1300">Ley</rango>
    <numero_oficial>2/2018</numero_oficial>
  </metadatos>
</documento>
```

Figura 2. Fragmento de un fichero xsd (arriba) y el fichero xml (abajo) asociado al documento con referencia BOE-A-2018-14948

2. Almacenamiento de datos: todos los metadatos de los documentos descargados fueron almacenados en una base de datos de tipo relacional (*MySQL*). Esta base de datos está compuesta por 21 tablas con un tamaño total de 8,8 gigabytes. La tabla principal llamada “documento” contiene los metadatos más relevantes de cada documento como la fecha de publicación, sección a la que pertenece, si está derogado o no, etc. Esta tabla está conectada a un conjunto de tablas auxiliares, como las que recogen los metadatos “materias”, “alertas”, “referencia_anterior” y “referencia_posterior”, etc. Un listado completo de los metadatos se muestra en la tabla 1. Cada tabla auxiliar puede tener diferentes relaciones con la tabla principal, por ejemplo, un documento sólo puede tener cardinalidad 1:1 respecto al departamento o la sección a la que pertenece, ya que sólo se publica en una sección y por un único departamento. Por el contrario, existen otros metadatos que pueden tener cardinalidad n:m, como las alertas, las materias o las referencias a otros documentos que contiene el documento en cuestión. Un documento puede estar asociado a ninguna o varias materias, ninguna o varias alertas, y puede referenciar a cero o varios documentos.

3. Análisis de datos: sobre la colección completa de documentos descargados se realizaron distintos análisis cuantitativos; para ello se emplearon consultas en lenguaje sql contra la base de datos. Estas consultas se integraron desde el lenguaje de programación R (*R Core Team*, 2014) con el que se realizó el resto de los gráficos y análisis.

La figura 3 muestra a modo de ejemplo una consulta sql empleada en uno de los análisis realizados (cálculo de la distribución de documentos publicados por año –los resultados se muestran en la figura 4). Empleando sentencias parecidas a la anterior se realizaron diversos análisis para dar respuesta a cuestiones como: ¿Qué cantidad de materias y alertas han sido

utilizadas en el *BOE* para describir estos documentos? ¿Todos los documentos hacen uso real de esos metadatos? ¿Cómo

```
SELECT distinct (documento.identificador) as 'ndoc', year(documento.fecha_publicacion) as 'year'
FROM documento
WHERE documento.fecha_publicacion BETWEEN '1969-09-01' AND '2019-06-30'
GROUP BY year(documento.fecha_publicacion)
```

Figura 3. Ejemplo de sentencia SQL

se reparte el uso de estos metadatos en los diferentes tipos de documentos y a lo largo de las diferentes secciones, departamentos, años?, etc. La respuesta a estas preguntas se detalla de manera más extensa en el apartado de Resultados.

4. Resultados

Se descargaron 2.396.485 documentos publicados en el *BOE* desde el 1 de septiembre de 1960 hasta el 30 de junio de 2019 para todas las secciones.

En la figura 4 se muestra la distribución de documentos publicados en cada año. El año en el que menos se publicó fue 1960 (7.233), ya que sólo se recogen los primeros 4 meses (septiembre a diciembre) de vigencia del *BOE*. De todo el conjunto de años destacan los años 2005 a 2010 (ambos inclusive) con más de 50.000 items cada uno, siendo 2007 el año con más publicaciones (61.093) de la serie histórica. El año completo (doce meses) con menos documentos publicados fue 1996 con 22.007. Por su parte, los primeros seis meses de 2019 acumularon un total de 21.398 publicaciones.

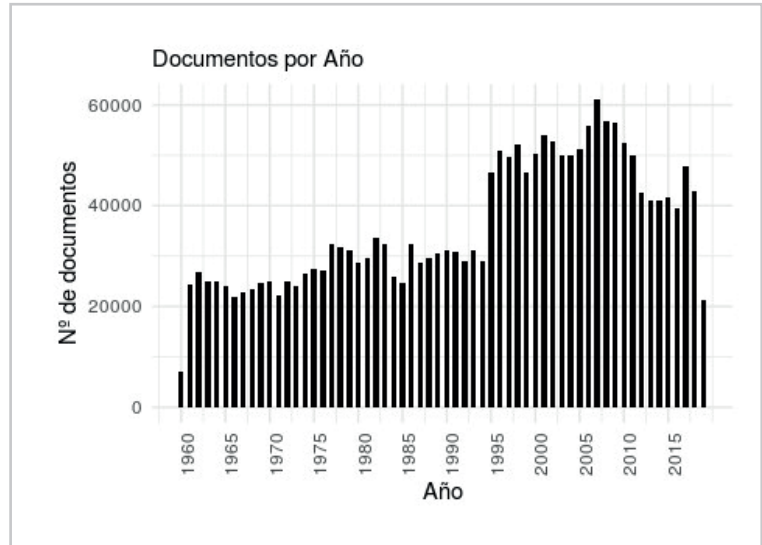


Figura 4. Cantidad de documentos publicados cada año en el *BOE*

Se estudió también el significado de cada metadato y en qué secciones se emplean (tabla 1). Hay 14 metadatos comunes a todas las secciones, como son el título, datos identificativos del documento, url a la versión pdf del propio documento, o el texto completo. Existen metadatos que sólo están presentes en una única sección, como el importe o el ámbito geográfico, que sólo aplican a documentos de la Sección V.

Del estudio de la tabla 1 se concluye que los metadatos “alerta”, “materia”, “materias_cpv” y “metadata_eli” son los únicos que pueden usarse para la descripción de contenido de los documentos. Es de resaltar que “alerta” y “materia” sólo se aplican a las secciones I, II, III y TC, mientras que “materias_cpv” solo se emplean en la Sección V. Según hemos podido comprobar por nuestro análisis, aunque en el xml de los documentos adscritos a la Sección V aparece el metadato “materia”, éste siempre está vacío. Por su parte, “metadata_eli” sólo se encuentra en uso en algunas secciones.

Ningún documento de los publicados en la Sección IV presenta descripción documental. No le son de aplicación los metadatos “materia”, “alerta” ni “materia_cpv”. La única manera de recuperar estos documentos es por título, departamento, número del *BOE* o fecha de publicación, entre otros. De modo que los documentos de la Sección IV quedan excluidos de este análisis.

Aunque el metadato “materia” siempre aparece como parte de la sección de metadatos de los xml recuperados para los documentos de la Sección V, éste siempre está vacío, por lo que entendemos que es un error de la estructura de datos del *BOE*. El metadato “materia_cpv” es de aplicación exclusiva para la Sección V.

Se han realizado tres análisis:

- uno para las Secciones I, II, III y TC (que comparten el mismo conjunto de metadatos descriptores de contenido: “alerta” y “materia”);
- un segundo análisis para la Sección V (focalizado en “materias_cpv”);
- un último análisis centrado en la implementación de la iniciativa ELI.

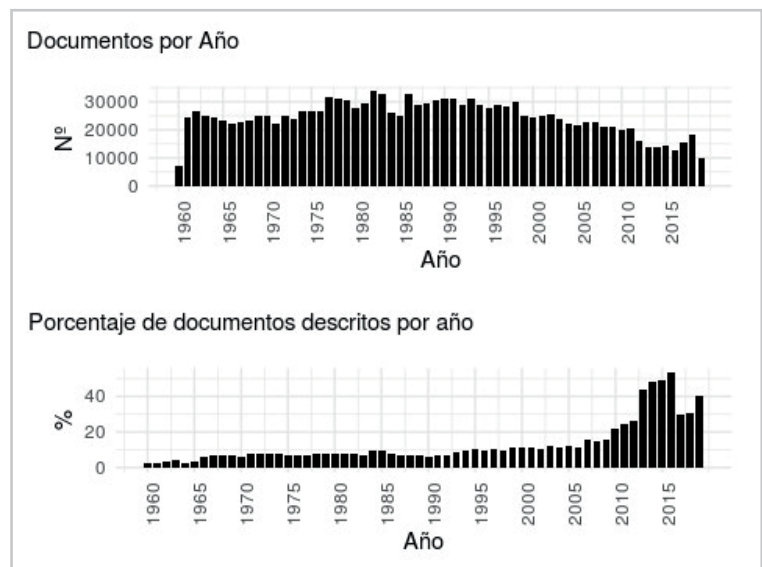


Figura 5. Documentos publicados y su porcentaje descrito anualmente para las Secciones I, II, III y TC

4.1. Análisis de las secciones I, II, III y TC

Del total de documentos publicados en el *BOE* entre las fechas indicadas, las Secciones I, II, III y TC representan el 61,25% del total (1.467.805 documentos).

En la figura 5 (arriba) se muestra la distribución de documentos publicados anualmente para las Secciones I, II, III y TC. El año con menos publicaciones fue 1960 (7.233), ya que sólo se recogen los primeros 4 meses de vigencia del *BOE*, mientras que el año con más ítems publicados en las secciones analizadas fue 1982 (33.794). A partir del año 2011 se observa un descenso notable de publicaciones, 14.495,611 documentos de media entre 2011 y 2019, mientras que en los años anteriores (1960 a 2010) la media fue de 26.129,549. De todo el conjunto de años destacan otros años como 1977, 1978, 1979, 1982, 1983, 1986, 1989, 1990, 1991, 1993, 1998 en los que se publicaron más de treinta mil documentos cada año. El año íntegro con menos publicaciones fue 2016 (12.573).

En la figura 5 (abajo) se muestra el porcentaje de documentos descritos anualmente por alguna alerta y/o materia. Se observa que hasta el año 2010 no se supera el 20% de documentos descritos por año. Los publicados entre 2013 y 2016 presentan una descripción superior al 40%, llegando su máximo en 2016 con casi un 55%. 2016 fue el año en el que menos ítems se publicaron en estas secciones en la historia del *BOE*, sin embargo se aprecia que es el año en el que porcentualmente se han descrito más. Las descripciones documentales más frecuentes en 2016 fueron:

- "Oposiciones" (2.085 documentos);
- la alerta "Concursos de personal público" (637 documentos);
- "Planes de estudios" (692 documentos).

Destacan también los seis primeros meses de 2019, donde el 40% de los 9.764 documentos han sido descritos, siendo "Convenios colectivos sindicales" (21 documentos), "Organización de las Comunidades Autónomas" (10 documentos) y "Planes de estudios" (10 documentos) las 3 materias más recurrentes. A nivel global sólo el 10,84% de los documentos publicados por las Secciones I, II, III y TC han hecho uso real de los metadatos materias y/o alertas para la descripción de su contenido.

Se han cuantificado 6.634 materias y 43 alertas. Algunos otros ejemplos de materias utilizadas son: "Abogados", "Iberia, Líneas aéreas de España", "Igualdad de oportunidades". Ejemplo de alertas son: "Energía", "Telecomunicaciones", y "Trabajo y empleo".

En la parte superior de la figura 6 se muestra el análisis de las diez materias más utilizadas a nivel histórico. En la parte inferior de esta misma figura se observan las diez materias más utilizadas en los últimos diez años. Se aprecia cómo coinciden en casi la totalidad de términos.

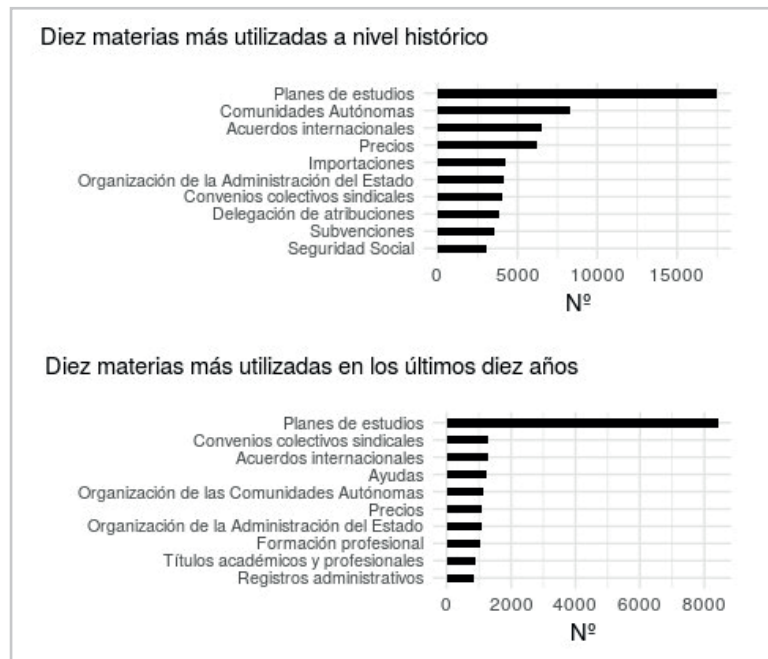


Figura 6. Diez materias más usadas a nivel histórico (arriba) y en los últimos diez años (abajo) por las Secciones I, II, III y TC

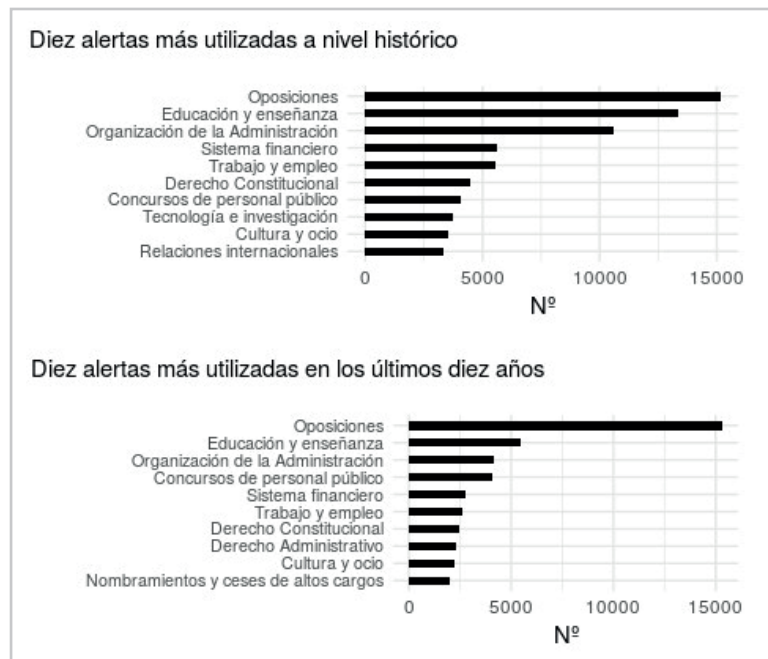


Figura 7. Las diez alertas más usadas a nivel histórico (arriba) y en los últimos diez años (abajo) por las Secciones I, II, III y TC

Hasta 2010 el porcentaje de documentos descritos anualmente por alguna alerta y/o materia no supera el 20%, mientras que de 2013 a 2016 llegan al 40%, alcanzando el máximo en 2016 con casi un 55%

Destaca la materia “Planes de estudios” en ambos casos con un uso total de 17.484 veces frente a 8.379 veces usada en los últimos diez años. Prácticamente en los últimos diez años se han publicado casi la mitad de los documentos relativos a “Planes de estudios” de toda la serie histórica.

La figura 7 muestra el mismo tipo de análisis de la figura 6, pero aplicado en este caso al metadato alertas. En la parte superior se muestran las diez alertas más utilizadas a nivel histórico, mientras que en la parte inferior se muestran las más utilizadas en los últimos diez años. En ambas coinciden gran parte de éstas. Siendo “Oposiciones” la que más destaca, empleada 14.295 veces por todos los documentos de las Secciones I, II, III y TC, frente a los 14.293 usos en los últimos diez años, lo cual muestra que hasta hace diez años aproximadamente no se etiquetaban documentos con ella.

Conociendo estos porcentajes se estudiaron las alertas y materias empleadas anualmente. En la parte superior de la figura 8 se puede observar el uso total de materias, el cual ha sido irregular con el paso del tiempo. Destaca el año 2011, cuando alcanza su máximo, utilizándose 19.678 términos. En los últimos tres años completos (2016, 2017 y 2018) su uso desciende drásticamente, con 11.926, 7.690 y 5.469 usos respectivamente. Los seis primeros meses de 2019 se hicieron un total de 4.219 usos.

En la parte inferior de la figura 8 se muestra el uso total del metadato alertas. Hasta 2002 inclusive este uso se sitúa por debajo de 1.250 por año; esto quiere decir que en este período de tiempo en el que se publicaron cerca de 25.000 documentos de media anual, apenas el 0.05% tenían una alerta asignada. A partir de 2003 empezó a aumentar muy significativamente, siendo 2015 cuando más usos se realizaron (9.544).

En la parte superior de la figura 9 se observa la cantidad de materias utilizadas en cada año. Destaca el año 2003 en el que se emplearon 2.094 términos diferentes, un 31,56% del total de las 6.634 materias diferentes presentes en la serie histórica.

En la parte inferior de la figura 9 se muestran las alertas diferentes empleadas por año. Se aprecia cómo de 1979 a 2015 (inclusive) el uso de estos términos se sitúa entre los 30 y 40. A partir de 2016 ya se superan los 40 términos diferentes empleados por año.

Sabiendo el número de documentos publicados junto con las alertas y materias utilizadas se realizó una comparación entre ellas. Tal y como se aprecia en la figura 10 el uso de las materias fue siempre muy superior al de alertas. Se puede observar que la tendencia de publicación de documentos ha sido cada vez menor en los últimos años. En el caso de las materias fueron en aumento de manera gradual, pero no de la misma forma respecto a las alertas, las cuales presentan un aumento exponencial a partir del año 2000.

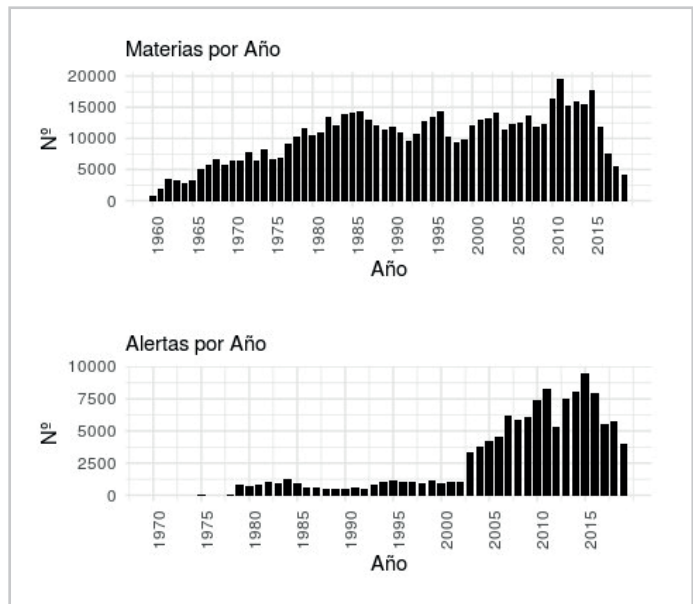


Figura 8. Cantidad de usos de materias (arriba) y alertas (abajo) empleadas en los documentos de las Secciones I, II, III y TC a lo largo de los años

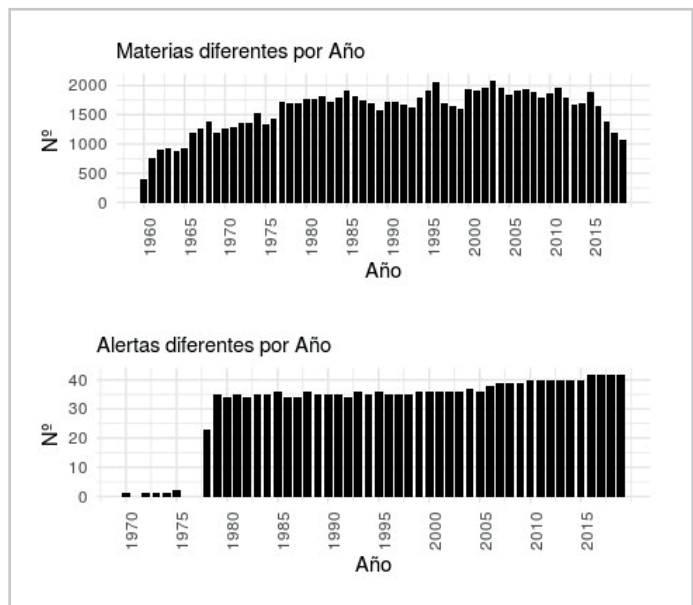


Figura 9: Cantidad de materias (arriba) y alertas (abajo) diferentes empleadas en los documentos de las Secciones I, II, III y TC a lo largo de los años

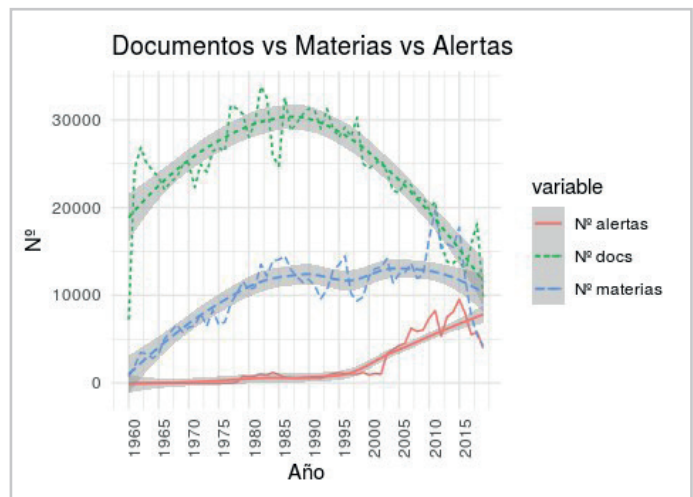


Figura 10. Evolución anual de publicaciones, materias y alertas para los documentos de las secciones I, II, III y TC

En la parte superior de la figura 11 se muestra la media anual de las materias empleadas en los documentos descritos. Cada documento etiquetado se describe con al menos tres términos y con un máximo de seis en el mejor de los casos. Los documentos no etiquetados no son tenidos en cuenta en este análisis.

En la parte inferior de la figura 11 se muestra la media anual de alertas por documento. Se aprecia que nunca superó la media de dos alertas por documento descrito, y existen tres años (1971, 1976 y 1977) en los que no se utilizó ninguna. Igualmente, para este cálculo no se han tenido en cuenta los documentos sin este metadato.

En la parte izquierda de la figura 12 se muestra la media de materias empleadas por cada sección y subsección de los documentos descritos. Destaca que los documentos se describen con entre 2 y 6 términos, siendo la Sección TC la que más emplea (5) y la Sección II la que menos (3). En la parte derecha de la figura 12 se muestra la media de alertas que se emplean por documento para cada sección y subsección. Se observa que se utilizan entre una y tres por documento, siendo la Sección TC la que más emplea (3) y las subsecciones II.A y II.B las que menos (1).

Para tener una idea más precisa sobre cuánto se publica anualmente, la cantidad de materias y alertas empleadas, o el porcentaje de documentos recuperables por medio de estos descriptores, se presenta la tabla 2 en la que se puede observar la media anual de estas variables y su desviación estándar para las Secciones I, II, III y TC.

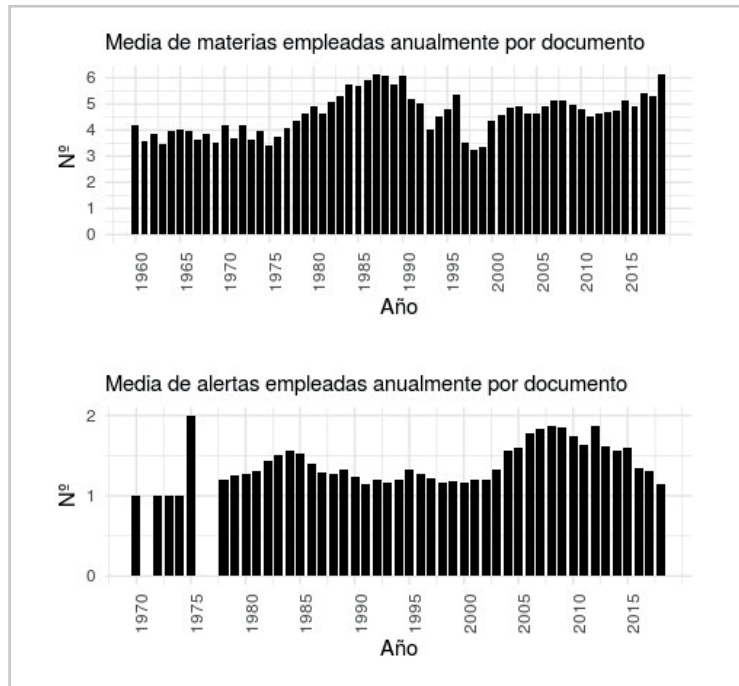


Figura 11. Media de uso de materias (arriba) y alertas (abajo) anuales de los documentos descritos para las secciones I, II, III y TC a lo largo de los años

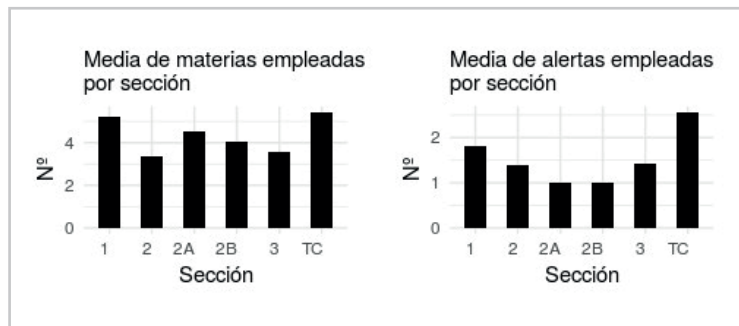


Figura 12. Media de materias (izquierda) y alertas (derecha) por sección de documentos descritos a lo largo de los años

En la tabla 2 se aprecia que anualmente se publicaron una media de 24.702,51 documentos. De ellos, una media de 1.345,91 fueron descritos por alguna alerta y 2.226,32 contienen alguna materia. En la tabla destacamos que de media se publicaron al año 22.023,88 documentos sin ningún descriptor frente a 2.678,62 que sí tenían alguno (materias o alertas). De manera porcentual, las Secciones I, II, III y TC del BOE presentan un 87,46% de documentos sin describir (ni materias ni alertas), frente a un 12,54% que sí tienen valores para alguno o ambos de esos metadatos.

Adicionalmente se realizó un análisis a través del resto de metadatos, con los siguientes resultados:

- Los cinco departamentos que más publicaron fueron: "Administración Local" (176.676 documentos; 96,36% no descritos), "Universidades" (117.410 documentos; 81,59%), "Ministerio de Educación y Ciencia" (114.099 documentos; 93,26%), "Ministerio de Justicia" (88.614 documentos; 95,48%), y "Ministerio de Economía y Hacienda" (64.616 documentos; 88,69%).
- Los tres rangos con más publicaciones han sido: "Resoluciones" (817.507 documento; 92,57% no descritos), "Órdenes" (411.240 documentos; 90,49%), y "Real Decreto" (95.734 documentos; 80,62%).

Tabla 2. Media y desviación estándar de documentos y descriptores anuales

Muestra analizada	Media	Desviación estándar
Documentos anuales	24.702,51	5.587,30
Documentos anuales con alertas	1.345,91	1.747,24
Documentos anuales con materias	2.226,32	790,00
Documentos anuales sin materias ni alertas	22.023,88	6.434,61
Documentos anuales con materias o alertas	2.678,62	1.468,56
Porcentaje de documentos anuales con materias o alertas	12,54	11,52
Porcentaje de documentos anuales sin materias ni alertas	87,46	11,52

- El 89,39% de los documentos de código legislativo estatal está sin describir, frente a un 64,67% de documentos de código legislativo autonómico.

4.2. Análisis de la Sección V

Este análisis se ha realizado sobre los metadatos de un total de 672.548 documentos (o anuncios) correspondientes a la Sección V. En parte superior de la figura 13 se muestran los documentos asociados a la Sección V publicados por cada año. Antes de 1995 apenas se publicaron anuncios, en total fueron 4.965 entre los años 1964 a 1980 (ambos inclusive). A partir de 1995 aumenta considerablemente la publicación de este tipo de documentos con 18.522 ese mismo año. En 2007 se alcanzó el máximo histórico (38.520). De los 672.548 documentos de la Sección V sólo hay descritos el 20,28% (136.441), para ello se han empleado sólo 57 términos diferentes como materias CPV de las más de 9.400 que existen en el *Vocabulario común de contratación pública* (Unión Europea, 2008)

En la parte inferior de la figura 13 se muestra el porcentaje de documentos descritos por año de la Sección V. En 2012 este porcentaje empezó a ser representativo. Se observa cómo en 2017 el 81,56% de los documentos estaban descritos mientras que en el último año completo (2018) esta cifra bajó hasta el 71,78%. Los seis primeros meses de 2019 acumularon un 11,49% de publicaciones descritas (1.122 de los 9.764 anuncios publicados).

En la figura 14 se muestran los documentos publicados frente a las materias CPV utilizadas para describirlos. Se aprecia cómo a pesar de que en los últimos años se han publicado menos, su descripción documental sí aumenta significativamente a partir del año 2000.

En la figura 15 se pueden apreciar las materias CPV empleadas en los documentos de la Sección V. En el período anterior al año 2012 apenas se alcanzaron las 8 materias. Fue en 2017 cuando se alcanzó el máximo con 30.852 usos para describir los documentos publicados. Desde 2013 se han hecho al menos 20.000 usos de materias CPV para describirlos.

El número medio de materias usadas para describir cada documento se muestra en la figura 16, en la que se observa que de media se utilizan entre una y dos materias CPV. Siendo los años 2008 y 2009 cuando más materias se han empleado de media y los años 2000, 2010 y 2011 cuando menos (sólo 1 por documento).

Las diez materias CPV más utilizadas tanto a nivel histórico como en los últimos diez años se muestran en la figura 17. Destacan las materias “Servicios de alcantarillado, basura, limpieza y medioambiente” (con 14.161 anuncios) y los “Trabajos de construcción” (con 13.935 anuncios).

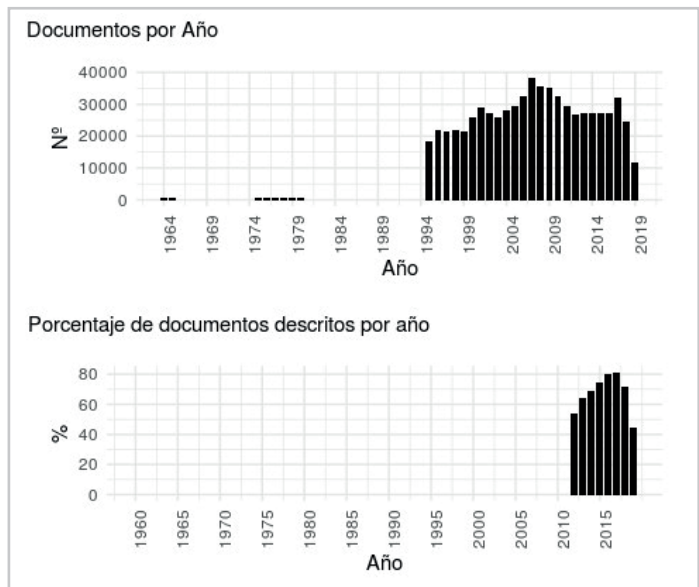


Figura 13. Anuncios publicados (arriba) y porcentaje de descritos (abajo) por año

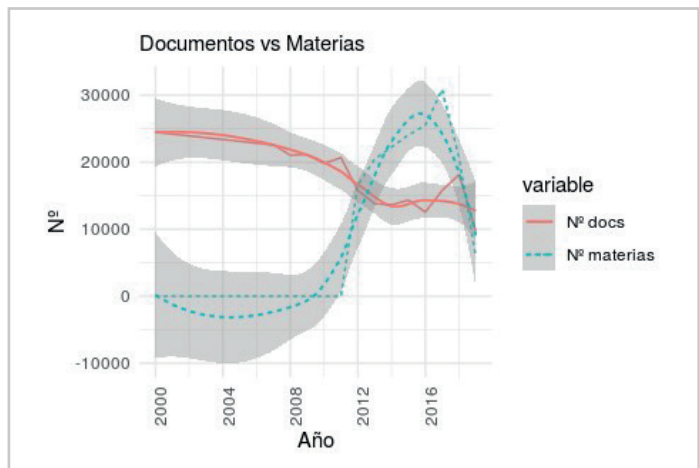


Figura 14. Documentos publicados vs materias empleadas para describirlos

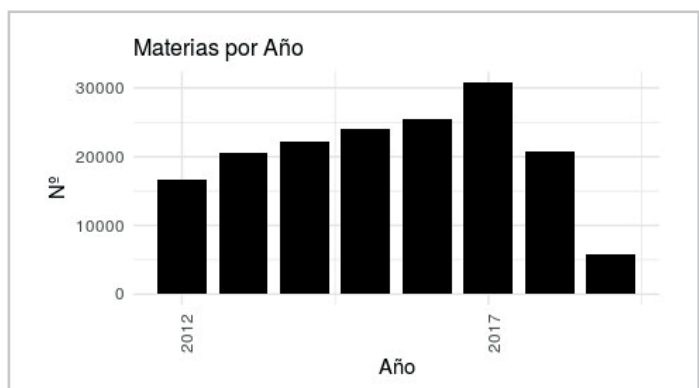


Figura 15. Materias empleadas por año

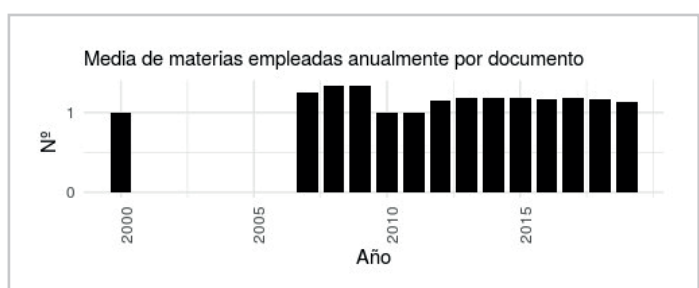


Figura 16. Media de materias CPV empleadas anualmente para describir cada documento

Al igual que en el análisis de las anteriores secciones podemos observar en la tabla 3 la media anual de estas variables analizadas y su desviación estándar para la Sección V.

En la tabla 3 se aprecia que anualmente se publicaron una media de 11.399,12 documentos. De ellos, una media de 2.312,56 anuales fue descrita con alguna materia CPV. De manera porcentual la Sección V presenta un 81,68% de media de anuncios sin describir con materias CPV. Sólo un 8,32% de los documentos están descritos. Casi todos los documentos descritos pertenecen a la subsección V.A, ya que las subsecciones V.B y V.C se encuentran vacías al 99,99% y 100% respectivamente.

Del análisis del resto de metadatos de esta sección V se concluye:

- Los cinco departamentos que más anuncios publicaron, y su proporción sin describir fueron: "Administración Local" (74.947 documentos; 57,62% sin describir), "Ministerio de Defensa" (66.891; 84,50% sin describir), "Ministerio de Fomento" (63.749; 85% sin describir), "Universidades" (58.525; 86,14% sin describir) y "Anuncios particulares" (27.140; 100% vacíos).
- El tipo de procedimiento con más documentos (122.803) es el abierto, y sólo tiene un 1,08% de ellos sin describir.
- En cuanto al tipo de anuncio se ha observado que los correspondientes a "Servicios" sólo un 0,96% no tienen descripción (de un total de 76.523), "Concurso de Servicios" el 99,45% de sus 57.724 de documentos tienen ausencia de descriptores; y "Suministros" con 54.745 documentos, presenta un 99,63% con descriptores vacíos.
- La tramitación mayoritaria es la "Ordinaria" con 113.021 y donde sólo se ha encontrado un 1,20% de documentos no descritos.
- A nivel cuantitativo la modalidad "Licitaciones" (con 73.054 documentos) presenta sólo un 0,97% de anuncios sin describir, en línea con lo sucede a las "Formalizaciones de contrato" (57.279 anuncios), con sólo un 1,10% de documentos sin descripción documental.

4.3. Análisis de los documentos con metadatos ELI

En este artículo hacemos un análisis final sobre la cobertura actual que presenta la implementación de la iniciativa ELI, que es bastante reciente. El BOE la empezó a implementar en diciembre de 2018 (Ministerio de Hacienda y Función Pública, 2018).

Del conjunto de documentos publicados en el BOE hasta el 30 de junio de 2019, sólo 79.156 estaban descritos con estos metadatos ("metadata_eli" y "url_eli"), lo que representa un 3% del total.

Como se puede apreciar en la figura 18, a fecha 30 de junio de 2019 se habían etiquetado documentos desde 1960 hasta 2019, siendo 1982 cuando más documentos se han descrito (2.403) y 1961 cuando menos, tan sólo un documento.

Un análisis por secciones nos muestra cómo la Sección I es en la que más documentos se han etiquetado (56.075), seguida por la Sección III (22.967). Aglutinan ambas el 99,86% de los documentos descritos mediante ELI.



Figura 17. Materias CPV más utilizadas a nivel histórico

Tabla 3. Media y desviación estándar de documentos y descriptores anuales

Muestra analizada	Media	Desviación estándar
Documentos anuales	11.399,12	14.060,04
Documentos anuales con materias	2.312,559	6.468,00
Documentos anuales sin materias	9.086,559	12.752,83
Porcentaje de documentos anuales con materias	8,32	23,08
Porcentaje de documentos anuales sin materias	81,68	23,08

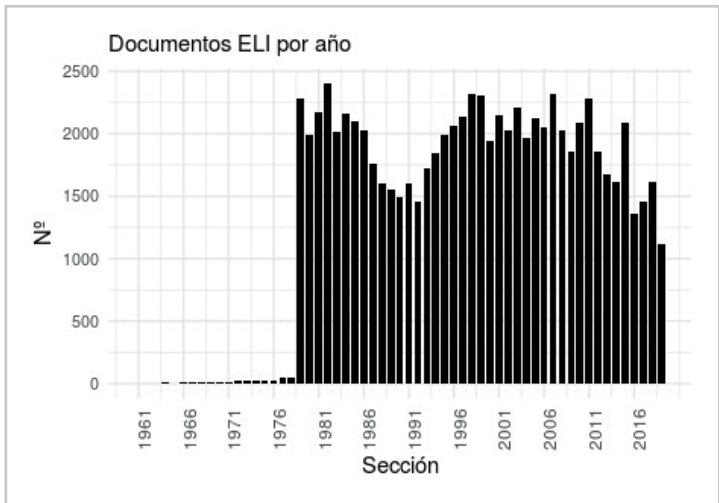


Figura 18. Documentos por año que contienen metadatos ELI

Por último, se analizaron los rangos de estos documentos. Las “órdenes” (22.355), “resoluciones” (21.981), “reales decretos” (15.601), “correcciones de erratas” (7.381) y “leyes” (7.362) son los cinco rangos que mayor cantidad de documentos aglutinan, el 94,35% (74.683 documentos) del total de 79.156 documentos descritos por ELI.

4. Conclusiones

El sistema de información del *Boletín oficial del Estado (BOE)* es el encargado de poner a disposición del ciudadano todo lo referente a normas, leyes, resoluciones judiciales, convocatorias, concursos públicos, etc., de aplicación y difusión a todo el estado español.

En este artículo se ha realizado una revisión documental sobre 2.396.485 documentos publicados por el *Boletín* entre el 1 de septiembre de 1960 y 30 de junio de 2019.

Se han estudiado los metadatos asociados a estos documentos, haciendo mayor énfasis en los que realizan la tarea de descripción documental del contenido [alertas, materias, materias CPV (*Common procurement vocabulary*) y metadatos ELI (*European legislation identifier*)]. El análisis nos muestra cómo tan sólo el 12,68% del total de los documentos hacen realmente uso de algún descriptor de corte documental.

Aunque la descripción documental ha ido aumentando con el paso del tiempo, apenas supera en los últimos años el 50% anual en el mejor de los casos (año 2016).

“ Aunque la descripción documental ha ido aumentando con el paso del tiempo, apenas supera en los últimos años el 50% anual en el mejor de los casos, el año 2016 ”

Del estudio podemos concluir que existe un determinado grupo de documentos que son más propensos a no tener descriptores de contenido, y por tanto a no ser recuperados de manera adecuada ni por el sistema web de búsqueda, ni tampoco ser sugeridos por el sistema de alertas. Estos grupos de documentos corresponden fundamentalmente a los publicados por los departamentos de la Administración Local, el *Ministerio de Educación y Ciencia*, las universidades y el *Ministerio de Justicia*, entre otros. También son muy altas las ausencias de descripciones documentales en documentos cuyos rangos son órdenes, resoluciones o reales decretos.

Se ha realizado, además, un análisis sobre la implementación de la iniciativa ELI por parte del *BOE*. Según los resultados obtenidos ésta alcanza (apenas) al 3% de los documentos, concentrándose por ahora sólo en documentos de las Secciones I y III.

También se han encontrado problemas a nuestro entender con la vasta cantidad de términos poco acertados usados como materias y/o alertas. A modo de ejemplo, se han encontrado los términos “Empleo” (1.375 documentos) y “Trabajo” (748), y las alertas “Concursos de personal público” (4.050) y “Oposiciones” (15.158). Estas materias y alertas que, si bien semánticamente están muy próximas entre sí y en algún caso son totalmente sinónimas (“Empleo” y “Trabajo”), el *BOE* los trata de manera indistinta. Se han encontrado documentos donde sólo aparece alguna de ellas, como por ejemplo “Empleo” (en el documento BOE-A-2016-7931) o “Trabajo” (BOE-A-2016-573), e incluso en algunos casos aparecen ambos en el mismo documento (documento BOE-A-2015-9735).

La alerta “Concursos de personal público” puede aparecer de manera individual en documentos como el BOE-A-2019-6111; “Oposiciones” por su parte aparece en documentos como el BOE-A-2019-8584; mientras que en otros documentos como en BOE-A-2019-7926 aparecen conjuntamente ambas.

Llamativo es también que los diferentes nombres que de manera histórica han tenido los entes gubernamentales (departamentos, ministerios, organismos, etc.) sean usados como materias. A modo de ejemplo se han encontrado asociado al metadato “materia” los textos: “*Ministerio de Economía, Industria y Competitividad*”, “*Ministerio de Economía y Hacienda*”, “*Ministerio de Economía y Empresa*”, “*Ministerio de Economía y Competitividad*”, “*Ministerio de Economía y Comercio*”, “*Ministerio de Economía Nacional*”. Estos textos usados como materia aportan poco o nada al contenido de los documentos donde aparecen, no describen el contenido subyacente al documento si no el ministerio que publica dicho documento. Se han encontrado más de 6.000 valores diferentes para el metadato “materia”, muchos de los cuales son de este tipo, semánticamente nulos y/o asociados a la fuente u origen del documento y no tanto al contenido.

El acumulado de estos errores (muy alto porcentaje de documentos vacíos, términos semánticamente idénticos o muy parecidos, términos poco acertados usados como materias, entre otros) hace que los sistemas de información y alertas del *BOE* se estén infrutilizando, y se esté mermando su capacidad de búsqueda, recuperación y difusión de información. La interacción se vuelve dificultosa, poco ágil, y en el mejor de los casos limitada, entre personas (con derecho a estar informadas) y los sistemas de información del *BOE*.

“ Existe un grupo de documentos más propensos a no tener descriptores, y por tanto a no ser recuperados por el sistema de búsqueda, por ejemplo, los publicados por la Administración Local y los ministerios de *Educación y Ciencia* y de *Justicia* ”

Hay demasiada desconexión entre los términos que un usuario puede seleccionar en el sistema de alertas *Mi BOE* y los términos (que según los resultados) son inexistentes o poco acertados en una inmensa mayoría de documentos. Hay una alta probabilidad de que la persona no sea adecuadamente informada por el principal boletín del país, lo que de alguna manera cercenaría sus derechos a estar informado y a acceder de manera transparente a la información.

En fechas recientes (19 de junio de 2019) se anunciaba a la sociedad la remodelación de la web del *BOE* (Hellín, 2019; EFE, 2019). Se han hecho progresos, se ha mejorado la web, la usabilidad y la navegación, su adaptabilidad a diferentes dispositivos, como la mejora de la interfaz de consulta, y se ha volcado nueva legislación y jurisprudencia autonómica, estatal y europea. Se han incorporado algunas nuevas opciones de búsqueda, y se ha iniciado la implementación de la iniciativa ELI. No obstante, aún queda trabajo por realizar. Son necesarios la revisión y el completado de la descripción documental de contenido de muchos documentos, de manera que esta descripción se engarce mejor con las opciones de búsqueda del boletín y con el listado de términos a elegir como alertas en el sistema de alertas.

Además de aumentar el etiquetado de los documentos que carecen de materias y/o alertas, se sugiere el uso de ontologías para unificar términos semánticamente iguales

Además de aumentar el etiquetado de los documentos que carecen de materias y/o alertas se sugiere el uso de ontologías para unificar términos semánticamente iguales, y la incorporación de una capa semántica que permita la búsqueda conceptual. De este modo se reduciría la cantidad de términos empleados como materias y alertas (más de 6.000). Con una capa semántica implementada, una persona podría, por ejemplo, solicitar documentos relacionados con el concepto “Empleo”, la capa semántica permitiría a los sistemas del *BOE* orientar automáticamente a esta persona hacia documentos relacionados con ese concepto. La persona en cuestión sólo tendría que introducir en el sistema el concepto “Empleo” y el sistema haría todo lo demás, facilitando así la interacción, y garantizando al mismo tiempo los derechos de acceso público y transparente a la información.

5. Referencias

España (1992). “Ley 30/1992, de 26 de noviembre, de régimen jurídico de las administraciones públicas y del procedimiento administrativo común”. *BOE*, n. 285, 27 noviembre, pp. 40300-40319.
<https://www.boe.es/buscar/doc.php?id=BOE-A-1992-26318>

España (1997). “Ley 6/1997, de 14 de abril, de organización y funcionamiento de la Administración General del Estado”. *BOE*, n. 90, 15 abril, pp. 11755-11773.
<https://www.boe.es/buscar/doc.php?id=BOE-A-1997-7878>

España (2013). “Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno”. *BOE*, n. 295, 1 diciembre, pp. 97922-97952.
<https://www.boe.es/buscar/doc.php?id=BOE-A-2013-12887>

EFE (2019). “Calvo presenta la nueva web del ‘mejor Boletín Oficial del Estado del mundo’”. *Eldiario.es*, 19 junio.
https://www.eldiario.es/politica/Calvo-presenta-nueva-Boletin-Oficial_0_911659585.html

Hellín, Jesús (2019). “Carmen Calvo celebra que el BOE renueve su ‘cara’ en internet y dice que es ‘el mejor del mundo’”. *Europa press*, 19 junio.
<https://www.europapress.es/nacional/noticia-carmen-calvo-celebra-boe-renueve-cara-internet-dice-mejor-mundo-20190619150600.html>

Ministerio de Hacienda y Función Pública (2018). *Proyecto ELI: European legislation identifier. Especificación técnica para la implementación del identificador europeo de legislación en España (Fase 1)*. NIPO: 169 18 041 8
https://www.elidata.es/documentacion_tecnica/especificacion_ELI_fase_1.pdf

R Core Team (2014). *The R project for statistical computing*.
<http://www.R-project.org>

Unión Europea (2008). “Reglamento (CE) No 213/2008 de la Comisión de 28 de noviembre de 2007 que modifica el Reglamento (CE) n. 2195/2002 del Parlamento Europeo y del Consejo, por el que se aprueba el vocabulario común de contratos públicos (CPV), y las Directivas 2004/17/CE y 2004/18/CE del Parlamento Europeo y del Consejo sobre los procedimientos de los contratos públicos, en lo referente a la revisión del CPV”. *Diario oficial de la Unión Europea*, L 74/1, 15 marzo.
<https://bit.ly/355Ma6o>

Unión Europea (2017). “Conclusiones del Consejo de 6 de noviembre de 2017 sobre el Identificador Europeo de Legislación”. *Diario oficial de la Unión Europea*, C 441/8, 22 diciembre.
[https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52017XG1222\(02\)&from=ES](https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52017XG1222(02)&from=ES)