



---

## NOTAS Y EXPERIENCIAS / NOTES AND EXPERIENCES

---

### Análisis terminológico a través de la producción científica relativa a los libros electrónicos

Raquel Gómez-Díaz\*, José Antonio Cordon-García\*,  
Antonia Olivia Jarvio-Fernández\*\*, Julio Alonso-Arévalo\*

\* Universidad de Salamanca, Facultad de Traducción y Documentación, Salamanca (España). GRUPO E-LECTRA.

Correo-e: [rgomez@usal.es](mailto:rgomez@usal.es); [jcordon@usal.es](mailto:jcordon@usal.es); [alar@usal.es](mailto:alar@usal.es)

\*\* Universidad de Veracruz (México). Correo-e: [ojarvio@uv.mx](mailto:ojarvio@uv.mx)

Recibido: 19-09-2011; 2ª versión: 16-01-2012; 3ª versión: 24-05-2012; Aceptado: 29-05-2012.

**Cómo citar este artículo/ Citation:** Gómez-Díaz, R.; Cordon-García, J.A.; Jarvio-Fernández, A.O.; Alonso-Arévalo, J. (2013). Análisis terminológico a través de la producción científica relativa a los libros electrónicos. *Revista Española de Documentación Científica*, 36(1):en002. doi: <http://dx.doi.org/10.3989/redc.2013.1.889>

**Resumen:** Los libros electrónicos, a pesar de ser un fenómeno relativamente reciente, poseen un recorrido cambiante sujeto a las mutaciones tecnológicas, al desarrollo de dispositivos de lectura, y a las distintas prácticas científicas y comerciales. Todos estos cambios que van de lo puramente terminológico a lo técnico y conceptual tienen su reflejo en la producción científica recogida en las distintas bases de datos, por lo que un análisis diacrónico de la misma permitirá detectar las transformaciones que se han producido así como la contextualización de dicho fenómeno desde un punto de vista científico, social y cultural. Para ello se procede a la combinación de distintas herramientas informáticas que permiten hacer el análisis bibliométrico y textual, así como la obtención de resultados extrapolables a estudios similares.

**Palabras clave:** Libros electrónicos; análisis terminológico; análisis conceptual; análisis bibliométrico.

#### Conceptual and terminological analysis of scientific production of electronic books

**Abstract:** Despite being a relatively recent phenomenon, e-books have been subject to various changes: in technology, in the development of reading devices, and in various scientific and commercial practices. All these changes ranging from the purely terminological to the technical and conceptual are reflected in the scientific collections of different databases. Therefore, a diachronic analysis of this production will permit detection of the transformations produced, as well as the contextualization of this phenomenon from a scientific, social and cultural point of view. This leads to a combination of different tools that allow textual and bibliometric analysis and that produce results that can be extrapolated to similar studies.

**Keywords:** eBooks; terminological analysis; conceptual analysis; bibliometric analysis.

## 1. INTRODUCCIÓN

El 21 de julio de 2011 la comisión de cultura del Congreso de los Diputados aprobó la nueva Ley de Depósito Legal, en sustitución de la normativa de 1971, para adaptar el depósito del patrimonio bibliográfico -y así su conservación y difusión- a los cambios producidos en el mundo de la edición como consecuencia de la aparición de las nuevas tecnologías y, especialmente, de las publicaciones en red. La iniciativa respondía, además, a la Estrategia Europa 2020 y sus iniciativas más emblemáticas, entre otras la Agenda Digital Europea ([http://ec.europa.eu/information\\_society/digital-agenda/index\\_en.htm](http://ec.europa.eu/information_society/digital-agenda/index_en.htm)), que recoge el impulso que desde las principales instituciones europeas se pretende dar a la conservación de contenidos digitales con vistas a garantizar su acceso a las generaciones futuras. En el Capítulo de Disposiciones Generales, en su artículo 1 sobre definiciones, referido al libro se expone: "*Libro: Obra científica, artística, literaria o de cualquier otra índole que constituye una publicación unitaria en uno o varios volúmenes y que puede aparecer impresa o en cualquier soporte susceptible de lectura. Se entienden incluidos en la definición de libro, a los efectos de esta Ley, los libros electrónicos y los libros que se publiquen o se difundan por Internet o en otro soporte que pueda aparecer en el futuro, los materiales complementarios de carácter impreso, visual, audiovisual o sonoro que sean editados conjuntamente con el libro y que participen del carácter unitario del mismo, así como cualquier otra manifestación editorial*". A la definición de libro como objeto se le añade otra de carácter funcional relativa a las nueva condición desempeñada por el mismo, aludiendo a los dos escenarios en los que se desarrolla esta función en la red: la producción y la difusión. Con esto se acomoda la ley a lo que podíamos considerar como una definición estándar de libro electrónico, esto es el archivo de cualquier tipo de obra realizado en un formato legible por ordenador o por los diferentes dispositivos de lectura, que puede contener elementos textuales, multimedia y audiovisuales (Cordón, 2011).

Esta incorporación da fe de un fenómeno cada vez más importante en el ámbito documental, esto es, la irrupción del libro electrónico como nuevo soporte de comunicación científica y cultural (Cordón y otros, 2011). El 8 de septiembre de 2011 moría Michael Hart, quien en 1971 creara un sitio pionero en el desarrollo de los libros electrónicos: el Proyecto Gutenberg. Poco después lo hacía Steve Jobs, artífice de Apple y de otro de los dispositivos que han revolucionado el mundo digital: el Ipad (Issacson, 2011). En el mismo mes Safari, una de las empresas emblemáticas en la agregación de libros electrónicos de carácter académico cumplía una década de andadura. Los libros electrónicos no son un fenómeno reciente, como ponen de manifiesto los hechos referenciados, lo que es reciente es la pujanza que están alcanzando so-

cialmente y los movimientos estratégicos que están obligando a desplegar a la industria editorial de todos los países (Benhamou y Guillon, 2010). Amazon abrió su plataforma en España el 15 de septiembre de 2011 (<http://www.amazon.es/>), Barnes and Noble presentó la suya para libros en español en 2010 (<http://www.barnesandnoble.com/u/Libros-en-espanol-Spanish-Books-Learn-Spanish/379001126/>), Google Ebooks ha seguido los mismos pasos, y el sector editorial reacciona al hilo de un movimiento cada vez más importante (Thompson, 2010).

En la Feria Internacional de Contenidos Digitales (FICOD) celebrada en Madrid en noviembre de 2011, se anunció que, por primera vez en España, estos habían superado a los contenidos analógicos (ONTSI 2011). Según este:

- La industria de los contenidos digitales facturó 9.125 millones de euros en 2010, un 14,1% más que el año anterior.
- La cifra de negocio de la venta de software para videojuegos en España ascendió a 631 millones de euros y se mantuvo en los niveles de 2009.
- La música digital experimentó un incremento del 10% en su cifra de facturación, alcanzando los 192 millones de euros.
- El mercado digital de las actividades cinematográficas, de vídeo y de programas de televisión supuso 2.571 millones de euros, un 5,1% más que en 2009.
- El sector audiovisual se consolida como el motor de la industria digital en España con un 44% de la facturación total de 2010.
- La edición digital de la prensa y los libros alcanzó los 780 millones de euros, un 6,3% más que en 2009.
- La inversión publicitaria en medios digitales fue de 799 millones de euros en el año 2010 (IAB, 2011).
- El 91,5% de la población española consume contenidos digitales a través de Internet o mediante un dispositivo electrónico no conectado a la Red.

Los libros representan una fuente de información en permanente evolución, constantemente sujeta a cambios y transformaciones que han incidido en sus sistemas de producción, distribución y recepción, al hilo de las modificaciones tecnológicas y sociales que se han ido sucediendo. La industria editorial ha constituido el fiel reflejo de estos cambios acogiendo las transformaciones sociales implícitas en las nuevas formas de lectura y escritura y generando productos que han revalidado estas prácticas (Stein, 2010). Un ejemplo de ello lo constituyen los modernos sistemas de lectura en la nube deudores de la filosofía del *Cloud Computing*

y del *Cloud Housting*, que están creando un modelo de negocio singular y privativo en el ámbito del libro (Shen y Koch, 2011).

La historia muestra que el libro no es una entidad fija, sino un accidente de la historia, de la tecnología que, en su configuración actual, tiene una duración relativamente breve en relación con otras formas y soportes (Manguel, 2005; Cavallo y Chartier, 2001).

Ernst Posner (O'Toole, 2011) explica que más de la mitad del tiempo en el que la humanidad se ha comunicado por escrito, lo ha hecho sobre polvo de ladrillo, y esta producción habría superado con creces la de toda Europa durante la Edad Media. En el contexto de la línea de tiempo que sostiene la lecto-escritura, la imprenta y el códice, constituyen acontecimientos importantes, pero coyunturales en el contexto global.

Los libros electrónicos poseen una historia, breve pero intensa, sujeta a los cambios tecnológicos, al desarrollo de dispositivos de lectura, y a las prácticas científicas y comerciales. Su naturaleza, estructura, prestaciones y configuración múltiple son objeto de análisis desde una gran variedad de disciplinas científicas, habida cuenta de la repercusión que su implantación reviste para el ámbito de la cultura, la economía y la sociedad (Bon, 2011; Dacos, 2010).

Son numerosos los datos contextuales que sirven para situar la importancia del fenómeno a nivel internacional y nacional. En Estados Unidos representan ya el 15% del conjunto de la producción editorial (IDPF, 2011), Amazon vende más libros electrónicos que en papel, lo mismo que ocurre con Barnes and Noble. En la última feria de Fráncfort las previsiones más conservadoras apuntaban a que en 2020 el 50% del mercado mundial del libro sería digital. En Europa la tendencia, aunque con menor importancia cuantitativa, es similar (Cordón y otros, 2011), y en España la evolución corre en paralelo con el resto de los países, con un crecimiento progresivo que arroja cifras de incremento de un 155% en la producción de libros electrónicos del año 2010 respecto al año anterior (Panorámica, 2011). Frente a los 12.514 ISBN concedidos en 2009, en 2010 éstos han ascendido a 18.500, producidos por pequeñas y medianas editoriales que aglutinan el 60% del mercado, y con un significativo 8,2% relativo a la autoedición que se incrementa en casi un 30% con respecto al año 2009.

Según la publicación *Hábitos de Lectura y Compra de Libros* (2011), en 2010 cerca de la mitad de la población de 14 o más años (47,8%) es lectora de formato digital. De estos, un porcentaje todavía elevado (48%) lee en el ordenador, un 6,6% en el móvil, y un 0,8% en el e-reader, porcentaje que se eleva al 1,1% en la encuesta del 2º trimestre de 2010. El perfil de los lectores digitales es interesante porque ilustra acerca de las tendencias y los desarrollos futuros. Los lectores digitales

hombres (54,3%) superan en 13,3 puntos a las lectoras (41,0%), y por tramos de edad, la diferencia más acusada, 70 puntos, se da entre los jóvenes de 14 a 24 años (78,2%) y los mayores de 65 años (7,7%). El nivel de estudios también agudiza las diferencias entre los lectores digitales, así el 75,2% de los que tienen estudios universitarios utilizan los nuevos soportes tecnológicos, porcentaje que desciende hasta el 21,5% entre los que sólo tienen estudios primarios.

Una evidencia de la progresión a que se hace referencia está en que en los datos aportados para el primer semestre de 2011, la lectura digital ha pasado de 47,8% al 52,5%. Así pues, los lectores en soporte digital han crecido en los primeros 6 meses del año 4,7 puntos respecto al año 2010. En libros electrónicos, 2,1 puntos respecto al 1er trimestre de 2010, y 1,5 puntos respecto a la media del pasado año.

La importancia del fenómeno ha sido recogida en la literatura científica cuyas bases de datos registran estos cambios y transformaciones que van de lo puramente terminológico a lo técnico y conceptual.

El objetivo de este artículo es analizar la evolución de esta nueva área de investigación en las principales bases de datos internacionales, con objeto de detectar la evolución y cambios en el tiempo de diferentes indicadores que permitan contextualizar la importancia del fenómeno. Para ello se procederá a:

1. Comprobar el índice de solapamiento de las principales bases de datos internacionales, generales y especializadas, en sus registros sobre libros electrónicos.
2. Conocer la evolución cronológica de las publicaciones relativas a los libros electrónicos, indicando tanto el volumen, como la temporalización y la distribución geográfica de las publicaciones.
3. Examinar el ámbito científico al que se circunscribe el tema, analizando su grado de multidisciplinariedad, para lo que el estudio se basará en las frecuencias de los términos que utilizan.
4. Detectar los autores más productivos y analizar los sistemas de autoría.
5. Analizar los temas emergentes a través del análisis de contenido de los artículos.
6. Examinar la evolución temática inherente a los mismos.
7. Analizar los términos recogidos en los trabajos publicados en las revistas españolas de la temática objeto de estudio que están incluidos en las bases de datos internacionales utilizadas en el estudio.

## 2. METODOLOGÍA

Para dar respuesta a los objetivos marcados en el apartado anterior se ha tratado de recoger una amplia muestra de la producción científica sobre los libros electrónicos. Para ello, se han realizado búsquedas en tres bases de datos distintas: WOS, SCOPUS Y LISTA. Las dos primeras son generales y la tercera especializada en el área de Biblioteconomía y Documentación. La elección de la combinación de bases de datos generales y específicas obedece a la necesidad de abarcar el mayor número posible de aspectos relacionados con los libros electrónicos ya que es posible que enfoques distintos al documental pudieran perderse en el caso de utilizar solo bases de datos especializadas.

En las tres bases de datos se ejecutó la misma sentencia de búsqueda y sobre los mismos campos, con el objetivo de establecer comparaciones entre ellas.

Tras diversas pruebas en las que haciendo búsquedas parciales se obtenían resultados con los términos elegidos, se ejecutó la sentencia que se muestra a continuación, buscando aunar las variantes de número mediante los truncamientos y haciendo combinaciones con el operador OR para conseguir, con una sola sentencia de búsqueda, un listado de referencias libre de duplicados. Se buscó exclusivamente en el campo título dado que el campo 'descriptor' no está presente en la base de datos WOS, y el campo 'resumen' generaba mucho ruido.

La sentencia de búsqueda fue la siguiente: Title=(E-book\* OR Ebook\* OR electronic book\* OR e-reader\* OR ereader\* OR electronic Reading OR digital book\* OR digital ink OR e-ink OR tactile book\* OR libro\* electrónico\* OR libro\* digital\*).

Como puede apreciarse en la sentencia de búsqueda la mayor parte de los términos elegidos están en inglés, lengua predominante en las tres bases de datos, aunque también se incluyeron algunos términos en español, dado que SCOPUS incluye un gran número de revistas en esta lengua.

Asimismo se pensó en incluir los descriptores "lectura electrónica", "lectura digital", "dispositivo\* de lectura", pero en las pruebas preliminares de las búsquedas no produjeron ningún resultado, por lo que finalmente se decidió no tenerlos en cuenta.

La utilización de las tres bases de datos persigue varias finalidades: por una parte extraer una muestra lo más amplia posible y, por otra, conocer el grado de solapamiento entre las distintas bases de datos con el fin de conocer si para el tema de los libros electrónicos es suficiente con consultar una base de datos general, como es el caso de WOS o de SCOPUS, y en este caso, conocer si ambas contienen las mismas revistas. La combinación de bases de datos también permitirá comprobar si por el contrario, es suficiente con realizar las búsquedas en una base de datos especializada, como es

el caso de LISTA, y ver qué grado de solapamiento tiene esta última base de datos con las dos bases de datos anteriores.

La ejecución de las búsquedas aportó 942 registros en el caso de la base de datos WOS, 702 en el caso de SCOPUS y 1045 para LISTA.

Como se puede apreciar el número de registros conseguidos con esta última base de datos es superior al de las dos bases de datos generales, lo cual ya es un indicador que apunta a que el tema de los libros electrónicos es recurrente en el ámbito de la Biblioteconomía y la Documentación.

Una vez obtenida la muestra fue necesario realizar distintas acciones. En primer lugar eliminar duplicados, al hacerlo se pudo comprobar cómo el índice de solapamiento era menor del esperado ya que son pocas las revistas que están presentes en las tres bases de datos, con la particularidad añadida de que la cobertura cronológica no es igual, por lo que la primera conclusión del estudio es que realmente es necesario consultar las tres bases de datos para tener una visión completa de la producción científica relativa a los libros electrónicos.

Tras la eliminación de los registros duplicados, los 2.689 registros se vieron reducidos a 2.567 artículos diferentes.

Aunque las bases de datos WOS y SCOPUS permiten hacer algunos análisis estadísticos sobre autores, temas, años... , estos análisis solo engloban a una de las bases de datos, por lo que para poder realizar el estudio completo ha sido necesario unificar los registros únicos en una base de datos de elaboración propia. La unificación de la base de datos se realizó en tres pasos:

1. Exportación de los registros desde cada base de datos al gestor de referencias *Refworks*.
2. Importación de los registros a la base de datos *Microsoft Access* para eliminar duplicados y hacer los primeros análisis.
3. Exportación de los resultados a *Microsoft Excel* para completar el análisis estadístico.

Como uno de los objetivos es la realización de un análisis terminológico, ha sido necesario construir un corpus textual. Este se ha realizado utilizando los títulos y los resúmenes de los artículos, por lo que fue necesario descartar aquellos artículos que no incluían resumen. El corpus textual contiene 1.860 artículos de la muestra inicial donde se han identificado 12.451 palabras únicas que tienen un total de 189.589 ocurrencias. En el tratamiento se eliminaron también las palabras vacías y aquellas palabras con una frecuencia de aparición inferior a 11 ocurrencias, por lo que el total de palabras a tener en cuenta en el análisis fue de 1.794.

En el proceso de depuración del corpus textual se suprimieron caracteres como "", espacios, ,, {an-non}, +, \$, acentos, etc. También se eliminaron

los espacios entre palabras unidas por guión como por ejemplo, "Hiper - book = Hiper-book," o "Electronic - book = Electronic-book". En lo que se refiere a los nombres de los autores se unificó bajo una sola forma gráfica, de este modo las variantes 'Gibbons S', 'Gibbons Sue' y 'GibbonsSusan' se recogieron bajo la forma Gibbons\_Susan. En todos los nombres de los autores, la coma en la columna autor se sustituyó por un guión bajo. Este mismo procedimiento se realizó con los títulos de las revistas.

En el análisis del corpus textual se agregaron las columnas: década, cinco años, categoría revista y categoría autor.

Para poder realizar el análisis del corpus textual se utilizó el sistema informático SPAD.T (Lebart y otros, 1998, 2000) que permite estudiar la información derivada del recuento de palabras desde una perspectiva multivariante. Para su implementación fue necesario construir una base de datos con los campos que se querían analizar. Gracias a este programa se pudo efectuar un análisis, por ejemplo, de la autoría y descriptores, que aporta mayor información que el análisis hecho con cualquiera de las herramientas citadas anteriormente, ya que el programa genera un listado en orden alfabético y otro de acuerdo a la frecuencia con que se repite cada palabra. Asimismo permite filtrar el

vocabulario eliminando las palabras vacías y marcando aquellas otras que presentan una frecuencia mayor.

### 3. ANÁLISIS DE LOS DATOS

Del conjunto de información extraída se eligieron algunas variables que permiten analizar distintos aspectos de la producción sobre libros electrónicos.

#### 3.1. Grado de solapamiento de las revistas

Una de las cuestiones que resultan más prácticas a los investigadores es conocer cuál es la base de datos que contiene mayor número de publicaciones sobre su ámbito de investigación, y sobre todo, cuando el tema que trata de explorar es susceptible de ser tratado desde varias disciplinas, como es el caso de los libros electrónicos. Por ello nos pareció interesante conocer la distribución de publicaciones de cada una de las bases de datos, y ver en qué base de datos se encuentran las revistas que contienen un mayor número de artículos. Así, de las 846 publicaciones, 398 están en WOS, 359 en SCOPUS y 224 en LISTA.

En la tabla I se muestran las publicaciones con un mayor número de artículos publicados y su distribución según las distintas bases de datos.

**Tabla I.** Revistas con mayor número de artículos

| REVISTA   | WOS | SCOPUS | LISTA | TODOS |
|---|-----|--------|-------|-------|
| <i>Library Journal</i>                                      | 41  | 26     | 114   | 151   |
| <i>Publishers Weekly</i>                                    | 0   | 0      | 110   | 110   |
| <i>Information Today</i>                                    | 0   | 0      | 92    | 92    |
| <i>Electronic Library</i>                                   | 43  | 24     | 23    | 69    |
| <i>Bookseller</i>   | 0   | 0      | 60    | 60    |
| <i>Against the Grain</i>                                    | 0   | 0      | 42    | 42    |
| <i>Econtent</i>   | 28  | 11     | 11    | 41    |
| <i>Online</i>   | 16  | 0      | 14    | 30    |
| <i>Computers in Libraries</i>                               | 0   | 8      | 21    | 29    |
| <i>Serials</i>  | 0   | 18     | 10    | 28    |
| <i>El Profesional de la Información</i>                     | 12  | 12     | 9     | 21    |
| <i>Journal of Academic Librarianship</i>                    | 9   | 5      | 12    | 21    |
| <i>Abstracts of Papers of the American Chemical Society</i> | 19  | 0      | 0     | 19    |
| <i>Learned Publishing</i>                                   | 12  | 5      | 0     | 19    |

Analizando los títulos de las revistas se puede comprobar cómo las publicaciones se inscriben dentro del área de la Biblioteconomía y Documentación. La diferencia entre el número de artículos de las revistas en las distintas bases de datos se debe a la diferente cobertura cronológica.

En cuanto a las revistas más productivas, la tabla I muestra cómo algunas de las revistas que tienen un mayor número de artículos no están en todas las bases de datos, como es el caso de *Publisher Weekly* o *Information Today*, que solo están en LISTA. También destaca que, entre las publicaciones más productivas, se encuentra una española: *El profesional de la información* que aparece en undécimo lugar y se encuentra presente en las tres bases de datos.

Si en lugar de hacer la evolución cronológica según el número total de artículos, se establece agrupando los años junto con las revistas, se puede observar cómo la revista más productiva ha sido *Publishers Weekly*, que en el año 2001 publicó un total de 41 trabajos, revista que como se acaba de indicar solo aparece recogida en

LISTA. Es necesario insistir que aquí solo está contemplado el primer trimestre del año 2011.

Otra cuestión que resulta de interés es el grado de solapamiento de las tres bases de datos por lo que se procedió a cruzar los nombres de las revistas y se pudo comprobar cómo sólo hay doce revistas que están presentes en las tres bases de datos, aunque esto no significa que en las tres bases de datos aparezcan los mismos artículos. Esta información permite conocer si es necesario consultar una sola base de datos o por el contrario es más adecuado consultarlas todas para estar al día en el tema en cuestión, y dado lo que muestra la tabla II comprobamos cómo a pesar de que las publicaciones pertenecen al área de la Biblioteconomía y Documentación, en algunos casos encontramos un mayor número de publicaciones (caso de *Aslib Proceedings*) en una base de datos general, como es el caso de WOS, mientras que en otros casos (*Journal of Academic Librarianship*) aparecen más publicaciones en una base de datos especializada como es el caso de LISTA.

**Tabla II.** Revistas presentes en las tres bases de datos con indicación del número de artículos

| REVISTA                                  | WOS | SCOPUS | LISTA | Nº TOTAL DE ARTÍCULOS |
|--|-----|--------|-------|-----------------------|
| <i>Aslib Proceedings</i>                 | 8   | 1      | 2     | 10                    |
| <i>Communications of the ACM</i>         | 4   | 1      | 1     | 4                     |
| <i>Econtent</i>                          | 28  | 11     | 11    | 41                    |
| <i>Electronic Library</i>                | 43  | 24     | 23    | 69                    |
| <i>Journal of Academic Librarianship</i> | 9   | 5      | 12    | 21                    |
| <i>Journal of Documentation</i>          | 6   | 1      | 1     | 8                     |
| <i>Learned Publishing</i>                | 12  | 5      | 6     | 19                    |
| <i>Library Hi Tech</i>                   | 5   | 5      | 4     | 9                     |
| <i>Library Journal</i>                   | 41  | 26     | 114   | 151                   |
| <i>Online Information Review</i>         | 8   | 2      | 1     | 9                     |
| <i>Publishing Research Quarterly</i>     | 5   | 3      | 6     | 14                    |
| <i>Serials Review</i>                    | 1   | 1      | 1     | 2                     |

### 3.2. Evolución de la producción

Otro de los datos de interés es conocer cómo ha sido la evolución de la producción. Para ello se analizó la fecha de edición de las distintas publicaciones tal y como se muestra en la figura 1, viendo también la evolución de la presencia en las distintas bases de datos.

Según los datos obtenidos la mayor producción se da en 2010, es necesario tener en cuenta que en la búsqueda sólo están incluidos los tres primeros meses del 2011 por lo que es de esperar que viendo la tendencia creciente de las publicaciones, en este último año haya más que en 2010. Asimismo se observa que la tendencia de evolución en las tres bases de datos es similar. La figura muestra también cómo el orden de producción de las bases de datos es LISTA, WOS y SCOPUS, aunque en los últimos años SCOPUS tiene más artículos de la temática estudiada que WOS.

Los trabajos más antiguos referenciados se encuentran en la base de datos WOS, pero tras el análisis se pudo comprobar que estos trabajos no se refieren ni a libros electrónicos ni a lectura electrónica en el sentido actual, sino que se refieren

a artículos o bien de medicina o bien relacionados con la industria en los que se hace referencia a la lectura que ofrecen determinados aparatos utilizados en mediciones como puede ser un "sonar".

En cuanto a los picos en la producción de artículos coinciden con tres hechos importantes: en 2001, en los dos años anteriores aparecieron los primeros dispositivos portátiles de lectura, en 2004 se inventó la tinta electrónica, en 2007 Amazon lanzó el Kindle, y en 2010 apareció el IPAD, el dispositivo que ha revolucionado el mundo de los dispositivos de lectura.

### 3.3. Tipologías documentales

En cuanto a las tipologías documentales (ver tabla III), como es de esperar por la información que recogen las bases de datos utilizadas, la tipología más abundante corresponde a los artículos de revistas. Además cabe señalar que Scopus también recoge información presentada en congresos (*Proceedings*). Cabe señalar cómo en LISTA no se incluyen monografías pero sí informes, tipología que solo está presente en esta base de datos. En la categoría otros se recoge lo que en la base de datos LISTA aparece recogido como *Generic*.

Figura 1. Evolución de la producción editorial de los libros electrónicos

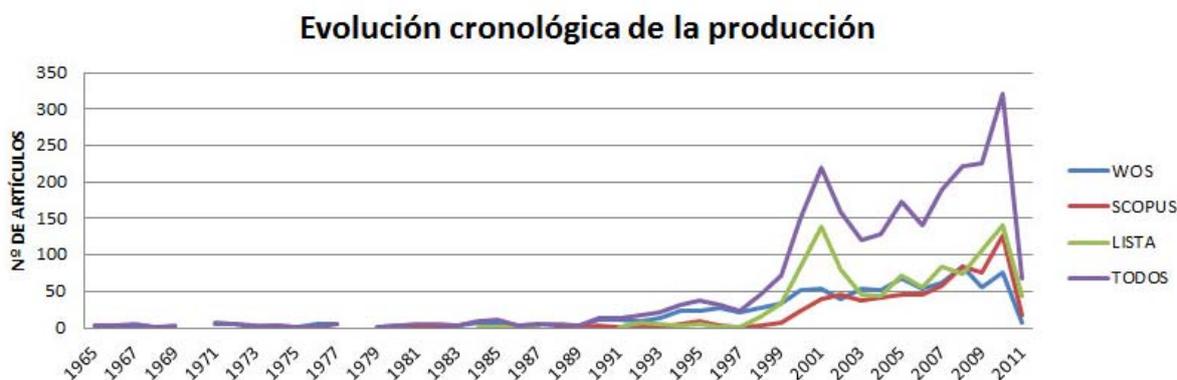


Tabla III. Tipologías documentales

| Reference Type     | WOS | SCOPUS | LISTA | TODOS |
|--------------------|-----|--------|-------|-------|
| Actas de congresos |     | 171    |       | 171   |
| Artículos          | 816 | 508    | 1010  | 2185  |
| Informes           |     |        | 2     | 2     |
| Libros             | 126 | 23     |       | 149   |
| Otros              |     |        | 33    | 33    |

Como cabía esperar para el tema de los libros electrónicos el artículo científico es el principal vehículo de transmisión del conocimiento.

### 3.4. Áreas temáticas de las revistas

Aunque tal y como se recoge con anterioridad, la base de datos con más artículos sobre los libros electrónicos es LISTA, especializada en Biblioteconomía y Documentación. Se analizó también la base de datos WOS para comprobar en qué otras áreas temáticas se recogen trabajos de este ámbito de investigación. En el caso de la base de datos Scopus, no se pudo analizar puesto que no proporciona dicha información.

Según la base de datos WOS, la mayor parte de los trabajos (un 42%) pertenecen al área de Biblioteconomía y Documentación seguidos por los de Informática y Sistemas Informáticos (15%). El resto de las áreas tiene una escasa representación.

### 3.5. Evolución en los temas a través de los términos

El corpus textual ha servido para estudiar los términos en función de las fechas de los artículos. En este sentido es importante señalar que al eliminar del corpus textual aquellos artículos que no tenían resumen, los artículos más antiguos han sido directamente descartados. Para este estudio se han tomado los datos a partir de la década de los 80, porque en los años anteriores la muestra es muy pequeña y por tanto se suponía que los resultados no serían significativos. A este hecho hay que unir, tal y como se ha explicado con anterioridad, que las referencias a la lectura electrónica se refieren a otros tipos de lectura que no son objeto de este estudio.

Al hacer el análisis de los términos más utilizados por décadas, se encontró que en los años 80 términos como 'e-book' o 'ebooks' están infrarepresentados. Es a partir de la década siguiente, cuando a pesar de que la producción en los temas objeto de estudio sigue siendo baja, el análisis de los términos comienza a mostrar aspectos de interés: empiezan a aparecer términos como 'hyper-book', 'electronic book', 'electronic publishing', y también los primeros nombres de modelos como es el 'Rocket' y 'Sony electronic book'. En esta misma década encontramos los términos 'ebooks', 'e-book' y 'e-books', aunque su frecuencia de aparición es baja.

En la década de los 2000 es cuando, tal y como muestra la figura 1, la producción científica aumenta considerablemente. Es en esta época cuando se producen distintos hechos de gran relevancia para el libro electrónico como la aparición de la tinta electrónica, el lanzamiento del Kindle y el del IPAD. En la producción científica de esta época aparece el término 'e-book' como uno de los términos más frecuentes, frente a lo que ocurría en la década anterior que era uno de los términos infrarepresentados. Asimismo se comienza a hablar de plataformas como 'Ebrary'

PDA (porque también hay aplicaciones de lectura para este tipo de dispositivos) o de 'formatos'. Aparecen sintagmas como 'e-book cataloging', 'hyper-book model', 'kindle', 'e-book piracy', 'ipad', 'piracy', 'electronic-book readers', 'e-readers', 'ereader' como palabras infrarepresentadas.

Finalmente en los años 2010 y primer trimestre de 2011 las palabras que aparecen con una mayor frecuencia coinciden con algunas menos representadas que aparecían en la década anterior: 'piracy', 'e-book cataloging'. También aparecen en esta época 'ereader', 'ebook piracy', 'Kindle', 'electronic book reader' o 'amazon' por indicar algunas.

En la tabla IV se muestra el número de artículos en los que aparecen los términos indicados. Los datos aparecen agrupados en intervalos de 5 años, y se ha considerado solo desde los años 80 puesto que a pesar de que nuestro corpus recoge textos desde los años 40 realmente el término electrónico book no aparece hasta mediados de los 80 (tres ocurrencias para el término en singular y una para el plural) y el resto de los términos no aparecen.

En la tabla IV hemos seleccionado aquellos términos que tienen una frecuencia de aparición superior a 20 ocurrencias.

Entre los términos más frecuentes utilizados encontramos 'Kindle' el dispositivo de la empresa Amazon.

Si en lugar de hacer el análisis de todos los términos que aparecen en el campo resumen nos limitamos al primer descriptor que marca el artículo, podemos hacernos una idea también de los temas que tratan los artículos.

Es importante indicar que la base de datos WOS no incluye descriptores, y además, en la tabla también se muestran como descriptores términos que en realidad indican el área de conocimiento al que se circunscribe el artículo. Aún así, resulta interesante destacar las áreas de conocimiento con un mayor número de artículos *Information Science and Lybrary*, seguido de *Computer Science information*.

### 3.6. Autores

Otro de los objetivos de este trabajo era conocer quiénes son los autores más productivos. Para el tratamiento de la variable autor se normalizaron los nombres para que quedaran en una sola forma gráfica, y eliminar así las variaciones que pudieran distorsionar los datos.

En cuanto a la productividad de los autores, en la tabla V se indican los diez más productivos señalando su evolución en el tiempo. El análisis de los autores revela cómo sólo hay 10 autores que escriben 10 artículos o más, mientras que la mayor parte de los trabajos están escritos por autores que sólo participan en la redacción de un solo trabajo. El autor más productivo es Calvin Reid que ha colaborado en 27 artículos, de los cuales 17 fueron entre los años 2000 y 2004.

Otra de las cuestiones interesantes es conocer si los autores escriben en solitario o lo hacen junto con otros colegas, en este sentido cabe señalar cómo más de la mitad de los artículos están escritos por un solo autor. Frente a los 1.342 artículos escritos por un solo autor, hay 852 que están escritos por 2 ó más autores.

### 3.7. Instituciones

Ligado al dato anterior está conocer las instituciones a las que están adscritos los autores. Aunque este dato no está disponible en la base de datos LISTA, a través de WOS y SCOPUS tratamos de hacernos una idea del tipo de institución. El análisis de las instituciones ha servido para ver cómo

la universidad es el ámbito donde se inscriben los autores que realizan trabajos relativos a los libros electrónicos, y también para comprobar cómo la Universidad Carlos III de Madrid se encuentra entre las instituciones más productivas en este tema.

Otras universidades españolas presentes en las bases de datos a través de los trabajos de sus autores son: la Universidad de Barcelona, la Complutense de Madrid, la Oberta de Cataluña, la Politécnica de Cataluña y la Rey Juan Carlos, aunque todas con un número de artículos inferior.

Tras este análisis hemos podido extraer las siguientes conclusiones que tratan de responder a los objetivos planteados en el artículo.

**Tabla IV.** Número de apariciones de los términos más frecuentes por años

|                       | 80-84 | 84-89 | 90-94 | 95-99 | 00-04 | 05-06 | 10-11 | TOTAL |
|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| E-books               | 0     | 0     | 0     | 16    | 322   | 682   | 232   | 1.252 |
| E-book                | 0     | 0     | 3     | 20    | 267   | 448   | 262   | 1.000 |
| Electronic-books      | 1     | 1     | 23    | 48    | 310   | 225   | 97    | 705   |
| Ebook                 | 0     | 0     | 0     | 16    | 133   | 189   | 84    | 422   |
| Ebooks                | 0     | 0     | 0     | 5     | 98    | 201   | 81    | 385   |
| Electronic-book       | 0     | 3     | 28    | 48    | 156   | 91    | 57    | 383   |
| Electronic publishing | 0     | 2     | 10    | 12    | 35    | 7     | 4     | 70    |
| Kindle                | 0     | 0     | 0     | 0     | 0     | 32    | 36    | 68    |
| E-readers             | 0     | 0     | 0     | 0     | 0     | 11    | 35    | 46    |
| E-book readers        | 0     | 0     | 0     | 0     | 9     | 16    | 18    | 43    |
| E-reader              | 0     | 0     | 0     | 0     | 0     | 8     | 33    | 41    |

**Tabla V.** 10 Autores más productivos por quinquenios

| Autor             | 90-94 | 95-99 | 00-04 | 05-09 | 10-11 | TOTAL |
|-------------------|-------|-------|-------|-------|-------|-------|
| REID_CALVIN       | 0     | 1     | 17    | 6     | 3     | 27    |
| ROGERS_MICHAEL    | 1     | 0     | 14    | 6     | 0     | 21    |
| HANE_PAULA_J_     | 0     | 3     | 9     | 3     | 5     | 20    |
| HILTS_PAUL        | 0     | 8     | 11    | 0     | 0     | 19    |
| BARKER_PHILIP     | 5     | 4     | 0     | 4     | 0     | 13    |
| MILLIOT_JIM       | 0     | 0     | 8     | 2     | 3     | 13    |
| OLEARY_MICK       | 0     | 0     | 12    | 1     | 0     | 13    |
| FALK_HOWARD       | 0     | 2     | 9     | 0     | 0     | 11    |
| WILLIAMS_WILDA_W_ | 0     | 0     | 10    | 0     | 0     | 10    |
| HAWKINS_DONALD_T_ | 0     | 0     | 9     | 0     | 0     | 9     |

#### 4. CONCLUSIONES

- Para tener una visión completa de la producción científica relativa a los libros electrónicos es necesario consultar las bases de datos, tanto de ámbito general como especializado, en el campo de la Biblioteconomía y la Documentación.
- Se muestra una tendencia al alza en la producción de artículos, coincidiendo con la aparición de los primeros dispositivos portátiles de lectura (año 2001), la aparición de la tinta electrónica (2004), el lanzamiento del Kindle (2007), y un incremento notable del número de artículos en el año 2010 donde destaca la aparición del IPAD.
- La literatura científica muestra como los cambios experimentados en los desarrollos tecnológicos y editoriales tienen una representación significativa en las bases de datos generales y especializadas. La aparición de nuevos formatos, de software de lectura, de dispositivos, de prácticas de uso diferenciadas, ha dado lugar a estudios de campo cada vez más abundantes.
- La terminología en el ámbito del libro electrónico ha ido evolucionando al tiempo que lo han ido haciendo los procesos de producción y recepción del mismo, incluyendo términos asociados a todos los ámbitos de la cadena editorial electrónica.
- Las publicaciones sobre libros electrónicos tienen una fuerte presencia en el ámbito de la Biblioteconomía y Documentación, tal y como lo demuestra el hecho de que se encuentren tantos trabajos incluidos en LISTA, base de datos especializada en este campo, pero también hay publicaciones adscritas al ámbito de la Informática, tal y como señala el análisis extraído de la base de datos WOS.
- En el caso de la investigación en España sobre este tema los resultados indican cómo la presencia más importante en el panorama internacional ha sido a través de la Universidad Carlos III de Madrid, y en el resto del panorama español la investigación se reparte entre las universidades ubicadas en Madrid y en Barcelona.

#### 5. BIBLIOGRAFÍA

- Benhamou, F.; Guillon, O. (2010). Modèles économiques d'un marché naissant: *le livre numérique. Prospective: économie de la culture et de la communication*, 2010, n. 2, pp. 6.
- Bon, F. (2011). *Après le livre*. Paris, Publie net.
- Cavallo, G.; Chartier, R. (Dir.). (2001). *Historia de la lectura en el mundo occidental*, Madrid; Taurus.
- Cordón García, J. A. (2011). *La revolución del libro electrónico*. Barcelona, El Profesional de la Información-Universitat Oberta de Catalunya, 105 pp.
- Cordón García, J. A.; Gómez Díaz, R.; Alonso Arévalo, J. (2011). *Gutenberg 2.0: la revolución de los libros electrónicos*. Gijón; Trea.
- Cordón García, J. A.; Gómez Díaz, R.; Alonso Arévalo, J. (2011). *Las plataformas de venta de libros electrónicos: modelos de negocio y estrategias de mercado*. BiD: textos universitaris de biblioteconomia i documentació, juny, núm. 26. <http://www.ub.edu/bid/26/cordon2.htm>
- Dacos, Marín (2010). *Read, Write Book*. Marsella, Cleo.
- Hábitos de lectura y compra de libros* (2011). Madrid, Federación de Gremios de editores.
- IAB (2011). Estudio sobre Inversión Publicitaria en medios digitales: resultados del año 2011. [http://www.iabspain.net/wpcontent/uploads/downloads/2012/05/Estudio\\_inversion\\_publicidad\\_IAB\\_Spain\\_2011.pdf](http://www.iabspain.net/wpcontent/uploads/downloads/2012/05/Estudio_inversion_publicidad_IAB_Spain_2011.pdf)
- International Digital Publishers Forum (IDPF) (2011). Industry Statistics. <http://idpf.org/about-us/industry-statistics>
- Isaacson, W. (2011). *Steve Jobs*. Barcelona, Debate.
- Lebart, L.; Salem, A.; Bécue, M. (2000). *Análisis Estadístico de Textos*. Lleida: Milenio.
- Lebart, L.; Salem, A.; Berry, E. (1998). *Exploring Textual Data*. Dordrecht: Kluwer Academic Published.
- Manguel, A. (2005). *Una historia de la lectura*. Barcelona; Lumen.
- Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información (ONTSI) (2011). Informe anual de los contenidos digitales en España 2011: industria y hábitos de consumo. <http://www.ontsi.red.es/contenidos-digitales/articulos/id/5542/informe-anual-los-contenidos-digitales-espana-2011.html>
- O'Toole, J. M. (2004). Back to the Future: Ernst Posner's Archives in the Ancient World. *The American Archivist*, Vol. 67, 2004, pp. 161-175.
- Panorámica de la Edición Española de Libros* (2011). Madrid, Ministerio de Cultura.
- Shen, W.; Koch, U. (2011). *eBooks in the Cloud: Desirable Features and Current Challenges for a Cloud-based Academic eBook Infrastructure*. Digital Publishing and Mobile Technologies 15th International Conference on Electronic Publishing, June 22-24, 2011, İstanbul, Turkey, p. 81.
- Stein, B. (2010). A Unified Field Theory of Publishing in the Networked Era. En: Dacos, M. *Read, Write Book*. Marsella, Cleo. Pp. 65-71.
- Thompson, J. B. (2010). *Merchants of culture. The publishing business in the Twenty First Century*. Londres; Polity Press, 440 pp.
- Wischenbart, R.; Kaldonek, S. (2011). *The global ebook market: current conditions and future projections*. Cambridge, Oreilly Media. 43 pp.