

## ANÁLISIS DEL MERCADO DE LA PRODUCCIÓN DE BASES DE DATOS EN LINEA

- Autores:** José Antonio Salvador Oliván  
Depto. Ciencias de la Documentación e Historia de la Ciencia.  
Universidad de Zaragoza  
jaso@posta.unizar.es
- José María Angós Ullate  
Depto. Ciencias de la Documentación e Historia de la Ciencia.  
Universidad de Zaragoza  
angos@posta.unizar.es
- María Jesús Fernández Ruiz  
Centro de Documentación. Ayuntamiento de Zaragoza  
mjferuiz@posta.unizar.es)
- Resumen:** Esta comunicación describe las características de uno de los actores fundamentales de la industria de la información como son los productores de bases de datos en línea. Se han seleccionado los productores que licencian sus bases de datos a 11 hosts importantes a nivel mundial, analizando qué productores individuales y tipo de organizaciones predominan en el mercado de la información, así como un análisis por países y continentes y el solapamiento existente entre los diferentes distribuidores. Los resultados muestran que un núcleo muy pequeño de productores cubre el 25% de las bases de datos, el predominio de Estados Unidos como primera potencia en la producción de información y que el tipo de productores más frecuente corresponde a las empresas privadas. Los distribuidores con mayor porcentaje de productores únicos son Dialog, DIMDI y FT Profile.
- Palabras Clave:** Productores de bases de datos; industria de la información; bases de datos en línea; estadísticas.
- Abstract:** This paper describe the features of the online databases producers which represent a key role in the information industry. We selected the producers licensing their databases to mundially important hosts eleven for examining the leading producers and organizations type in the information market, as well as analyzing te production by countries and continents, and the overlap between the hosts. The results show that a producers core covers the 25% of databases, the hegemony of USA as information producer, and the private companies are the type of organization most frequently. Dialog, DIMID and FT Profile are the hosts which present the percentage highest of unique producers.
- Keywords:** Database producers; information industry; online databases; statistics.

## 1. Introducción

Uno de los elementos más importantes en la industria de la información actual son los productores de bases de datos, organismos públicos o privados que se encargan de organizar, representar y almacenar la información y el conocimiento. Para hacer accesible esta información a los usuarios utilizan dos vías: alquilan sus bases de datos a empresas (conocidas como distribuidores o hosts) que se encargan de comercializarlas, o bien son ellos mismos los que facilitan el acceso a sus bases de datos, pudiendo coexistir las dos formas de explotación.

La segunda opción se ha visto favorecida por el impacto que ha tenido la World Wide Web en los productores, ya que además de tener casi todos una presencia de márketing en la red, algunos de ellos proporcionan un acceso directo a sus bases de datos, como es el caso de la National Library of Medicine, Information Access Company y Cambridge Scientific Abstracts, siendo este último el primer editor científico en proporcionar sus bases de datos en la web<sup>1</sup>.

Se espera que cada vez sean más los productores que se beneficien del bajo coste de edición en Internet y proporcionen un acceso web directo a sus bases de datos, aunque también es cierto que muchos productores se centran más en coleccionar información que en la difusión y que los costes de promoción y organización que implican estos cambios pueden desanimar a las organizaciones más pequeñas y conservadoras.

## 2. Objetivos

De forma global, la presente comunicación pretende ofrecer un mayor y mejor conocimiento de este eslabón fundamental en la cadena de producción y consumo de información, animados por la escasísima bibliografía existente al respecto y porque gracias a ellos podemos acceder a información de calidad y satisfacer con eficiencia la mayoría de las necesidades de información de los usuarios de un servicio de información.

Los objetivos específicos se resumen en los siguientes:

- Conocer qué productores predominan en el mercado de la información, según una base cuantitativa que mida el número de bases de datos creadas y su presencia en los principales servicios comerciales online (hosts).
- Análisis de productores por países y continentes.
- Análisis de la producción según el tipo de productor.

- Analizar el solapamiento de los productores en los diferentes distribuidores.

### 3. Material y Método

Se han seleccionado los productores de aquellas bases de datos que son accesibles a través de Dialog, STN, Orbit, FT Profile, Data-Star, OCLC, Ovid, Dimdi, EINS, HW Wilson y National Library of Medicine, que representan a la mayoría de los servicios mundiales más importantes y utilizados en las bibliotecas y centros de documentación. Todos los distribuidores debían de cumplir al menos uno de los siguientes criterios: que las bases de datos comercializadas proviniesen de más de un productor, o que fueran también productores y alquilaran sus bases de datos a más de un distribuidor (como es el caso de HW Wilson).

El número de productores incluidos en el estudio es de 562, que producen un total de 1.111 bases de datos. Los datos de los productores se han obtenido de las Hojas de Datos (FactSheets) proporcionadas por los diferentes distribuidores e introducidos en una base de datos diseñada para tal efecto con FileMaker 5.0.

Las variables fundamentales del estudio son: nombre del productor, bases de datos, distribuidores, país del productor y tipo de productor. El tipo de productor se ha clasificado en tres categorías analizando el dominio principal de su dirección URL:

- empresas privadas: dominio .com, .co o .net
- organismos gubernamentales: dominio .gov
- organizaciones o asociaciones orientadas a la investigación y difusión de la información sin ánimo de lucro: dominio .org

En aquellos productores que no aparecían ninguno de los dominios mencionados anteriormente, se ha entrado en su sitio web y a través de diferentes enlaces como “Quiénes somos”, “Misión”, etc., se ha conseguido suficiente información para clasificarlos en alguna de las tres categorías. El análisis estadístico se ha realizado con el programa SPSS 8.0.

### 4. Resultados

**Principales productores de bases de datos:** La Tabla I muestra los principales productores de bases de datos según el número de bases de datos, con un punto de corte en 10 o más bases de datos creadas. En total son 15 los productores que cumplen esta condición, lo que representa que sólo el 2,6% de los productores acaparan el 24,4% (271 de 1.111) de todas las bases de datos ofrecidas por los distribuidores; 122 productores (el 21,7%) crean entre 2 y 9

bases de datos, lo que supone un total de 414 bases (37,2%) y 426 productores (75,8%) sólo crean 1 base de datos (38,3%).

De estos 15 productores, destaca que 4 son también distribuidores elegidos en este estudio: H.W.Wilson Company, U.S. National Library of Medicine, The Dialog Corporation, y Chemical Abstracts Service (como parte de STN). Otros productores también ofrecen acceso a sus propias bases de datos a través de la web, como The Gale Group, Reuters, Cambridge Scientific Abstracts, Bell & Howell, Disclosure, Elsevier e INPI.

Los productores que licencian sus bases de datos a más distribuidores son U.S. National Library of Medicine, Elsevier Science y Cambridge Scientific Abstracts. El primero de ellos cubre el área de la medicina, siendo uno de los pioneros en la industria online y uno de los pocos que ofrece gratis todas sus bases de datos en la web; el segundo cubre los campos de la ciencia, tecnología y medicina, y el último se ocupa principalmente de literatura científica y técnica. Los tres son pesos pesados en la industria de la información y han desempeñado un papel fundamental en la creación y difusión de conocimiento.

*Tabla I. Principales productores de bases de datos según número de bases de datos y distribuidores que las comercializan.*

	Nº Bas es	Dial og	Data -Star	STN	DIM DI	OCL C	Ovid	FT Profi le	Orbit	EIN S	Wils on	NLM
H.W. Wilson Company	32	10				14	10				31	
The Gale Group	30	25	18	2		3		3	1			
Reuters Ltd.	24							24				
U.S. National Library of Medicine	22	4	7	4	10	1	4			1		21
The Dialog Corporation	20	18	4					1				
Cambridge Scientific Abstracts	19	12	3	15	1	2	3		1	6		
Chemical Abstracts Service	18	4	3	12	1	1						
Dun & Bradstreet	17	8	12					1				
Compu-Mark	15	13							13			
Derwent Information Ltd.	15	6	9	12	3		1		4			
Bell & Howell Information and Learning	13	11	2			6	5	2				

Disclosure, Inc.	13	10				3						
Financial Times Business Ltd.	12	1	1					11				
Elsevier Science	11	8	5	6	5	2	3		1	2		
Institut National de la Propriété Industrielle (INPI)	10	1							10			

También es importante destacar que Dialog contiene bases de datos de todos estos productores excepto de Reuters, que por otra parte sólo licencia sus bases de datos a FT Profile, cubriendo 131 bases de datos de las 271 (48,3%) lo que indica algo ya conocido por todos y es que Dialog es el servicio comercial mundial con mayor cobertura de bases de datos. Igualmente reseñable que tanto Wilson como la National Library of Medicine, aunque como productores licencian sus bases de datos a más distribuidores, cuando actúan como distribuidores sólo ponen al servicio del público sus propias bases de datos.

**Procedencia de los productores:** En la Tabla II se muestra el país al que pertenece cada productor. Indiscutiblemente, casi la mitad de los productores son de Estados Unidos, siguiéndole el Reino Unido con un 19%, Alemania con un 12,3% y Francia con un 5%, que son los países que siempre han dominado la producción de información.

En Europa están representados todos los países excepto Andorra, Croacia, Finlandia, Croacia, Hungría, Islandia, Mónaco, Noruega, Portugal, Rumania, Turquía y Yugoslavia.

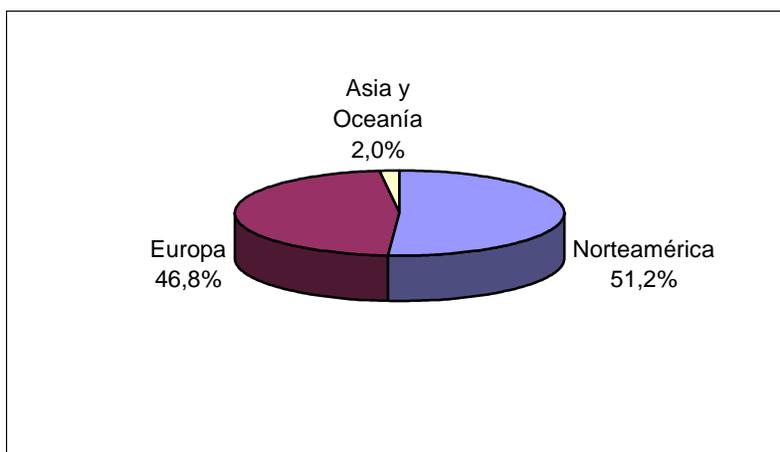
*Tabla II. Países de procedencia de los productores*

País	Nº Productores	Porcentaje
Estados Unidos	276	49,1%
Canadá	12	2,1%
Reino Unido	107	19,0%
Alemania	69	12,3%
Francia	28	5,0%
Suiza	10	1,8%
Italia	7	1,2%
Bélgica	7	1,2%
Holanda	6	1,1%
Austria	4	0,7%

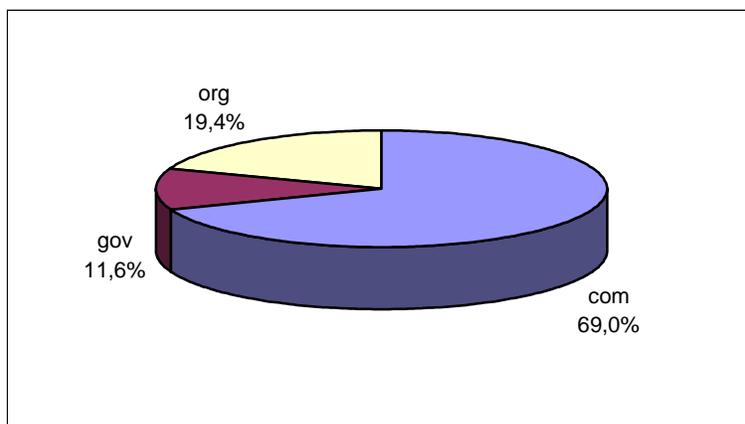
España	4	0,7%
Rusia	4	0,7%
Dinamarca	3	0,5%
Irlanda	3	0,5%
Suecia	3	0,5%
Escocia	2	0,4%
Eslovenia	1	0,2%
Estonia	1	0,2%
Luxemburgo	1	0,2%
Polonia	1	0,2%
República Checa	1	0,2%
Ucrania	1	0,2%
Japón	3	0,5%
Australia	2	0,4%
Israel	2	0,4%
Hong Kong	1	0,2%
India	1	0,2%
Nueva Zelanda	1	0,2%
Taiwán	1	0,2%

En el análisis por continentes (Gráfico I), destaca la igualdad existente entre Norteamérica (principalmente Estados Unidos con un 49,1%) y Europa (46,8%), y aunque limitado por el ámbito de este estudio, se reducen las distancias que existían en años anteriores cuando por ejemplo en 1989 había un total de 1.024 productores en Estados Unidos y de 610 en Europa, como indica Recoder y cols.<sup>2</sup> según datos extraídos del Directory of Online Databases.

Esta amenaza de Europa en su intento de usurpar a Estados Unidos su liderazgo en la producción de bases de datos, ya es advertida por algunos autores<sup>3,4</sup>. Robert E. Aber denuncia que la Directiva de la Protección Legal de Bases de Datos de la Unión Europea (adoptada en marzo de 1996 y entrada en vigor el 1 de enero de 1998) plantea una importante amenaza a los productores de bases de datos de Estados Unidos, ya que fue diseñada para promover una mayor inversión en la industria de las bases de datos en Europa y, de esta manera, superar sus deficiencias en competencia con la industria de Estados Unidos, mientras que las leyes de USA han desanimado al sector privado para crear, mantener y difundir colecciones valiosas de información.

*Gráfico I. Distribución de productores por continentes*

**Tipo de productores:** El tipo de productores que predomina en el mercado es el correspondiente a empresas privadas (69%), seguido de organizaciones sin ánimo de lucro (19,4%), y sólo el 11,6% de los productores son agencias u organismos de los gobiernos de los diferentes países.

*Gráfico II. Tipo de productores*

Al analizar el tipo de productores por continentes (Tabla III), se observa que en los tres continentes predominan las empresas privadas, aunque en mayor medida en Norteamérica y Asia-Oceanía, y que no existe ningún productor de gubernamental en este último. Sin embargo, debido al escaso número de productores en Asia y Oceanía, si comparamos sólo Europa frente a Norteamérica, es destacable que en el primero existe un mayor porcentaje de

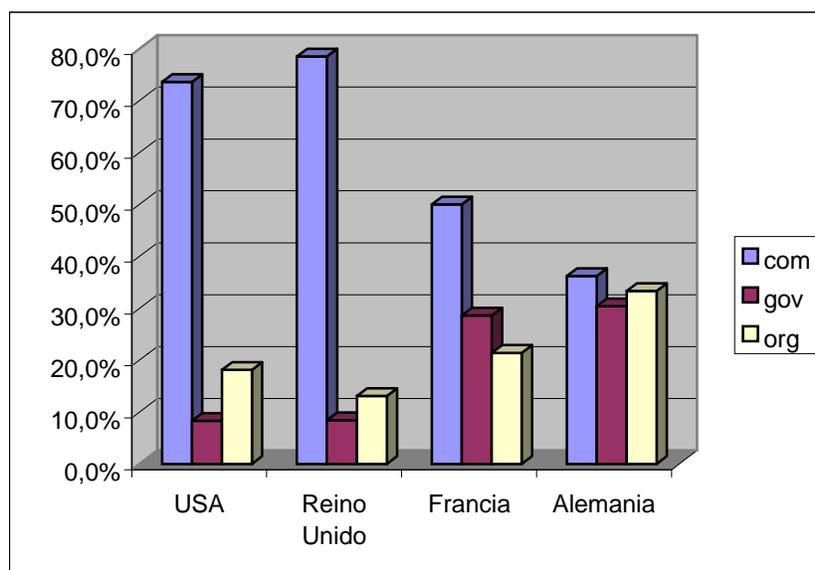
organizaciones sin ánimo de lucro y de agencias gubernamentales que en Estados Unidos.

*Tabla III. Tipo de productores por continentes*

	<b>Com</b>	<b>Gov</b>	<b>Org</b>
Norteamérica	213 (74,0%)	25 (8,7%)	50 (17,4%)
Europa	167 (63,5%)	40 (15,2%)	56 (21,3%)
Asia y Oceanía	8 (72,7%)	0 (0,0%)	3 (27,3%)

Por otra parte, no todos los países europeos presentan la misma tendencia en los porcentajes del tipo de productores, ya que mientras en el Reino Unido el porcentaje de productores de tipo comercial es el más alto (78,5%), y la proporción del tipo de productores es similar a la ofrecida por Estados Unidos, en Francia disminuye el porcentaje de productores comerciales a un 50% y todavía baja más en Alemania que presenta un porcentaje de productores muy similar en los tres tipos (Gráfico III).

*Gráfico III. Tipo de productores en Estados Unidos, Reino Unido, Francia y Alemania*



**Solapamiento de los productores en los distribuidores:** En la Tabla IV se muestra el solapamiento de los productores entre los distribuidores. Aunque puede ocurrir, y de hecho ocurre sobre todo en los grandes productores, que un productor coincida en varios distribuidores sin que por ello signifique que

ofrezca todas las mismas bases de datos, un alto grado de solapamiento de productores entre diferentes distribuidores indicará que existe una alta probabilidad de encontrar la misma información, o bases de datos, en estos últimos.

La Tabla hay que interpretarla de la siguiente manera: en la diagonal aparecen el número de productores en cada distribuidor, y los porcentajes corresponden al solapamiento de los distribuidores que aparecen en las filas con los distribuidores que aparecen en las columnas. De esta manera, un 25,4% de los productores con presencia en Dialog también lo están en Data-Star; casi la mitad de los productores en Data-Star (45,7%) aparecen en Dialog; STN, DIMDI, OCLC, Ovid y ORBIT también presentan el mayor grado de solapamiento con Dialog. EINS tiene el mismo grado de solapamiento con Dialog y con STN, y FT Profile presenta mayor solapamiento con Data-Star. Debido a que Wilson sólo cuenta con 1 productor, no puede ser tenido en cuenta en la valoración del solapamiento.

*Tabla IV. Solapamiento de productores en los diferentes distribuidores*

	Dialog	Data-Star	STN	DIMDI	OCLC	Ovid	FT Profile	ORBIT	EINS	Wilson	NLM
Dialog	295 (25,4%)	75 (16,6%)	49 (6,8%)	20 (7,8%)	23 (9,5%)	28 (9,5%)	28 (10,2%)	30 (6,4%)	19 (0,3%)	1 (0,7%)	2 (0,7%)
Data-Star	75 (45,7%)	164 (20,1%)	33 (11%)	18 (9,1%)	15 (14,6%)	24 (23,6%)	38 (11,6%)	19 (9,8%)	16 (0%)	0 (1,8%)	3 (1,8%)
STN	49 (50%)	33 (33,7%)	98 (14,3%)	14 (12,2%)	12 (15,3%)	15 (2%)	2 (21,4%)	21 (19,4%)	19 (0%)	0 (2%)	2 (2%)
DIMDI	20 (42,6%)	18 (38,3%)	14 (29,8%)	47 (17%)	8 (27,7%)	13 (0%)	0 (8,5%)	4 (23,4%)	11 (0%)	0 (6,4%)	3 (6,4%)
OCLC	23 (60,5%)	15 (39,5%)	12 (31,6%)	8 (21,1%)	38 (50%)	19 (7,9%)	3 (13,2%)	5 (15,8%)	6 (2,6%)	1 (2,6%)	1 (2,6%)
Ovid	28 (70%)	24 (60%)	15 (37,5%)	13 (32,5%)	19 (47,5%)	40 (2,5%)	1 (15%)	6 (22,5%)	9 (2,5%)	1 (7,5%)	3 (7,5%)
FT Profile	28 (28,3%)	38 (38,4%)	2 (2%)	0 (0%)	3 (3%)	1 (1%)	99 (5,1%)	5 (0%)	0 (0%)	0 (0%)	0 (0%)
ORBIT	30 (57,7%)	19 (36,5%)	21 (40,4%)	4 (7,7%)	5 (9,6%)	6 (11,5%)	5 (9,6%)	52 (23,1%)	12 (0%)	0 (0%)	0 (0%)
EINS	19 (44,2%)	16 (37,2%)	19 (44,2%)	11 (25,6%)	6 (14%)	9 (20,9%)	0 (0%)	12 (27,9%)	43 (0%)	0 (0%)	1 (2,3%)
Wilson	1 (100%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)	0 (0%)
NLM	2 (33,3%)	3 (50%)	2 (33,3%)	3 (50%)	1 (16,7%)	3 (50%)	0 (0%)	0 (0%)	1 (16,7%)	0 (0%)	6 (0%)

Complementando la Tabla anterior, se muestra en la Tabla V qué porcentaje de productores únicos aparece en cada distribuidor, constituyendo un indicador de en qué medida un distribuidor puede proporcionar información o bases de datos que no podremos encontrar en otro distribuidor. Así, destaca que Dialog trabaja con un 60% de productores cuyas bases de datos sólo se podrán encontrar en dicho distribuidor; también presentan altos porcentajes DIMDI y FT Profile, que superan ligeramente el 50%. Por el contrario, Ovid es el que ofrece menor porcentaje de productores únicos con un 17,5%.

*Tabla V. Productores únicos en cada distribuidor*

	<b>Total Productores</b>	<b>Productores únicos</b>	
Dialog	295	177	60,0%
Data-Star	164	62	37,8%
STN	98	38	38,7%
DIMDI	47	25	53,2%
OCLC	38	10	26,3%
Ovid	40	7	17,5%
FT Profile	99	52	52,5%
Orbit	52	14	26,9%
EINS	43	15	34,9%
Wilson	1	0	0,0%
NLM	6	3	50,0%

## 5. Conclusiones

El índice de progreso científico no depende sólo de coleccionar cantidades ingentes de datos, sino también de la calidad de estos datos, de su facilidad de uso y de la difusión de información en una base de datos<sup>4</sup>. Los métodos y actividades para organizar los datos en una base de datos que tenga una estructura, formato, presentación y documentación apropiados y con un control de calidad adecuado (producto final de valor añadido) recaen en aquellas entidades conocidas como productores o creadores de bases de datos.

En este trabajo se han analizado algunos aspectos o características de estos productores, y aun conscientes de que el ámbito de la población a estudio (productores de bases de datos en línea presentes en los distribuidores más importantes) impide generalizar los resultados a todos los productores, las conclusiones son las siguientes:

- El mercado de las bases de datos en línea está dominado por unos pocos productores que cubren aproximadamente el 25% de las bases de

datos, que corresponden a organizaciones con una gran tradición en la industria de la información y que en la mayoría de los casos alquilan sus bases de datos a varios distribuidores pero también son accesible a través de ellos. Por otra parte, hay un gran número de productores que sólo crean una base de datos.

- Estados Unidos sigue siendo con diferencia el país con mayor número de productores de bases de datos, seguido del Reino Unido, Alemania y Francia, países con mayor nivel tecnológico y científico que el resto. Sin embargo, cuando se tienen en cuenta todos los países europeos, los resultados revelan que se han reducido las distancias con respecto a Estados Unidos.
- De forma global, el tipo de productores más frecuente es el correspondiente a empresas de tipo privado. Este predominio es más marcado en Estados Unidos y en el Reino Unido que en Francia y Alemania, en los que aumenta el número de productores del tipo de organizaciones sin ánimo de lucro y gubernamentales hasta el punto de que en este último país el porcentaje es muy similar entre los diferentes tipos de productores
- En Dialog, DIMDI y FT Profile encontramos más productores únicos que en el resto de distribuidores, lo que indica que contienen más bases de datos únicas y que ofrecen información que no vamos a encontrar en otros distribuidores.

## 6. Bibliografía

1. Wiley, D.L. (1996). "Will the Internet turn database producers into information providers". *Business Information Alert: sources, strategies and signposts for information professionals*, vol. 8, nº 9, pp. 1-5. [accesible online en <http://www.consultnw.com/about/businf.html>]. (consultado el 12 de Enero del 2001).
2. Recoder, M.J.; Abadal, E.; Codina, L. (1991). *Información electrónica y nuevas tecnologías*. Promoción y Publicaciones Universitarias, S.A.: Barcelona.
3. Aber, R.E. (1998). H.R. 2652, the Collections of Information Antipiracy Act. [accesible online en <http://www.hyperlaw.com/aber.htm>] (consultado el 12 de Enero del 2001).
4. Committee for a Study on Promoting Access to Scientific and Technical Data for the Public Internet; Commission on Physical Sciences Mathematics and Applications; National Research Council. (2000). *A Question of Balance: Private Rights and the Public Interest in Scientific and Technical Databases*. National Academy Press. [accesible online en <http://www.nap.edu/books/0309068258/html>] (consultado el 12 de Enero del 2001).