

Recuperación y visualización de información en *Web of Science* y *Scopus*: una aproximación práctica

Cristina Jiménez Noblejas *

Antonio Perianes Rodríguez **

Artículo recibido:
28 de noviembre de 2012.

Artículo aceptado:
15 de enero de 2014.

RESUMEN

La aparición de *Scopus* (Elsevier) en noviembre de 2004 supuso un cambio en el mercado internacional de bases de datos y una presión añadida a la supremacía ejercida hasta ese momento por *Web of Science* (Thomson Reuters). El presente trabajo tiene como propósito la comparación de las posibilidades para recuperar y visualizar información en pantalla a partir del análisis de las interfaces gráficas de usuario y de los principales sistemas de búsqueda de ambas bases de datos. La metodología se basa en el estudio de ambos productos y en la realización de consultas seleccionadas con el exclusivo propósito de conocer el compor-

* Universidad Carlos III de Madrid, España. cristina.jnoblejas@alumnos.uc3m.es

** Universidad Carlos III de Madrid; SCImago Research Group; España.
antonio.perianes@uc3m.es

tamiento de *Scopus* y *Web of Science* durante el proceso de búsqueda, recuperación y visualización de información. Los resultados obtenidos permiten concluir que si bien las dos bases de datos presentan defectos en la descripción de los registros que afectan su localización, *Scopus* ofrece mayor eficacia en la obtención de información.

Palabras clave: *Web of Science*; *Scopus*; Recuperación de Información; Interfaces Gráficas de Usuario.

ABSTRACT

Information retrieval and visualization in *Web of Science* and *Scopus*: a practical approach

Cristina Jiménez-Noblejas and Antonio Perianes-Rodríguez

The emergence of *Scopus* (Elsevier) in November 2004 marked a major shift in the international market of databases exerting pressure on the monopoly traditionally held by *Web of Science* (Thomson Reuters). This paper aims to carry out an analysis of the graphical interfaces of both products as a key factor in the process of information retrieval and visualization. The methodology entails the execution of a series of searches in order to understand the respective performances of *Scopus* and *Web of Science* when put through the paces of searching, locating and retrieving information. Although the two databases exhibit deficiencies, results show that *Scopus* more effectively meets informational requirements in the indexing process by author.

Keywords: *Web of Science*; *Scopus*; Information Retrieval; Graphical User Interfaces.

INTRODUCCIÓN

Desde la aparición de *Science Citation Index (SCI)* en los primeros años de la década de los sesenta, germen de *Web of Science* (en adelante *WoS*) (Garfield, 1955; 1964), se convirtió en la herramienta de referencia para la satisfacción de las necesidades bibliográficas (*citation tracking*) de la comuni-

dad científica. En la actualidad, recoge más de 10 000 revistas de investigación y 38 millones registros, y está integrada en la plataforma *Web of Knowledge*, propiedad de la empresa Thomson Reuters.

En noviembre de 2004 la empresa Elsevier lanzó al mercado la base de datos multidisciplinar *Scopus*. En sus inicios, *Scopus* proporcionaba al usuario la posibilidad de consultar el número de veces que un trabajo había sido citado. En la actualidad recoge más de 18 000 títulos y 41 millones de registros y se ha convertido en la gran competidora de *WoS* por el control del mercado internacional de bases de datos de carácter científico.

El surgimiento de *Scopus* alentó la aparición de una gran cantidad de trabajos con el objeto de identificar cuál de los dos productos era el que mejor respondía a las necesidades de los investigadores. Estos estudios se agruparon en diferentes categorías: aquellos centrados en el análisis individual de *Scopus* (Fingerman, 2005), aquellos dirigidos a la comparación de las dos bases de datos (Deis & Goodman, 2005; Fingerman, 2006; LaGuardia, 2005) y aquellos que comparaban ambas con *Google Scholar* (Jacso, 2005; Bar-Ilan, 2010; Mingers & Lipitakis, 2010).

En este sentido, Susan Fingerman (2005) realizó un análisis sintético de las principales características de *Scopus*. En él la autora destacó la cobertura temática de la base de datos, en la que las Ciencias de la Vida y de la Salud copaban el mayor número de títulos. Otro de los aspectos analizados fue su cobertura temporal. El producto de Elsevier permite acceder a los resúmenes de los artículos desde 1966 y al listado completo de referencias bibliográficas desde 1996. Estos datos deben ponerse en relación con *WoS*, cuya cobertura temporal es significativamente mayor a la de *Scopus*. Por último, Fingerman señaló el hecho de que el 60 % del contenido recogido en la base de datos no proviene de los Estados Unidos.

Deis & Goodman (2005) efectuaron un análisis comparativo de ambos productos, dando una severa crítica a *Scopus* con base en el incumplimiento de la prometida actualización diaria de contenidos y en la falta de algunos artículos, números e incluso años completos de revistas cuya cobertura temporal debía estar cubierta. Este aspecto también ha sido señalado en trabajos posteriores como el de Vieira & Gomes (2009), en el que se detectó la cobertura parcial de algunos títulos de revistas.

Cheryl LaGuardia (2005) es una firme defensora de *Scopus* ya que considera que su interfaz está mejor orientada a la búsqueda, utiliza una terminología menos técnica y, por tanto, es más accesible para los usuarios. En relación a *WoS* reconoce su superioridad en Ciencias Sociales y en Artes y Humanidades. A pesar de ello, recientes estudios como el de Mingers & Lipitakis (2010) concluyen que la cobertura de *WoS* es pobre en Ciencias Socia-

les y, más concretamente, en el área de administración y gestión de empresas, siendo incluso superada por *Google Scholar*.

En un nuevo estudio comparativo, Fingerman (2006) consideró como una ventaja de *Scopus* sobre *WoS* el hecho de incluir tipos documentales como libros, actas de congresos o estudios de mercado, además de los artículos habituales. Posteriormente, la versión 4.10 de *WoS* incluyó dos nuevas bases de datos referenciales de actas de congresos internacionales: *Conference Proceedings Citation Index-Science (CPCI-S)* y *Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH)*. Santa & Herrero-Solana (2010) analizaron la cobertura de las revistas, la producción científica y la visibilidad de los principales países de América Latina y el Caribe en estas dos bases de datos, determinando la mayor cobertura de *Scopus* en cuanto a revistas, aunque no necesariamente traducida en mayor volumen de trabajos debido a la reciente incorporación de muchos de esos títulos a la base de datos holandesa.

Peter Jacso (2005) fue uno de los primeros autores en realizar un análisis pormenorizado de *WoS*, *Scopus* y *Google Scholar*, prestando especial atención al aumento y a la mejora de las citas así como a los trabajos que de ellas se derivan. Con similares propósitos Judith Bar-Ilan (2010) comparó esas tres mismas fuentes a partir de las citas recibidas por el libro *Introduction to informetrics* de Leo Egghe y Ronald Rousseau. Una de las principales conclusiones fue que la cobertura de *Scopus* y *WoS* es similar si el año de publicación es posterior a 1996.

En esta misma línea se encuentra la principal conclusión del trabajo de Escalona, Lagar & Pulgarín (2010). Los autores centraron su estudio en el campo de la Ingeniería Química y realizaron un análisis comparativo de ambas bases desde un punto de vista cuantitativo. El resultado de su estudio reflejó la similitud existente entre ellas en lo que a contenido y dinámica se refiere.

Leydesdorff (2012) llevó a cabo una comparación de tendencias de publicación entre Estados Unidos, la Unión Europea (UE-27) y China gracias a una de las últimas novedades de la versión 5.5 de *WoS*, la cual permite recuperar más de 100 000 registros en una única consulta, facilidad ya disponible en *Scopus*.

A pesar de que han sido numerosos los estudios comparando ambos sistemas de búsqueda, el presente trabajo trata de cubrir, por un lado, el aspecto insuficientemente tratado de la realización de búsquedas específicas para detectar las fortalezas y debilidades de ambos productos; por otro, actualizar resultados previos dado que las novedades en las últimas versiones de ambas bases de datos suponen saltos cualitativos reseñables respecto a las precedentes.

METODOLOGÍA

El principal propósito de este trabajo es establecer la idoneidad de los procesos de obtención de información en *Scopus* y *WoS* como herramientas esenciales en los estudios relacionados con la evaluación de la ciencia y la tecnología. Para ello se ha realizado un exhaustivo análisis de ambas interfaces y de los mecanismos de búsqueda que ofrecen, previo a su comparación sistemática, con el fin de establecer las principales fortalezas y debilidades de ambos productos desde la perspectiva de la recuperación y la visualización de los resultados en pantalla.

Para lograr los objetivos planteados se consultaron *WoS* y *Scopus*, la primera gracias al contrato suscrito por la Fundación Española para la Ciencia y la Tecnología que la ofrece como servicio público a todas las instituciones públicas de investigación españolas desde enero de 2004 (FECYT, 2009); la segunda a través del servicio de referencia de la Biblioteca Tomás Navarro Tomás del Centro de Ciencias Humanas y Sociales del Consejo Superior de Investigaciones Científicas.

El primer paso consistió en analizar las interfaces de cada una de las bases de datos. A continuación se compararon sus distintas funcionalidades con el propósito de establecer equivalencias y desigualdades desde el punto de vista de la recuperación de información, y finalmente se realizaron consultas en ambas fuentes con idénticos criterios de búsqueda (de autores y palabras en títulos). Para las pesquisas sobre autores se utilizaron las opciones de búsqueda sencilla *Search* (*WoS*) y *Document Search* (*Scopus*), para las búsquedas de autores se emplearon *Author Search* (*Scopus*) y *Author Finder* (*WoS*). La búsqueda de palabras en títulos se centró en la localización de términos concretos a través de la búsqueda sencilla (*Search* de *WoS* y *Document Search* de *Scopus*) y la búsqueda avanzada de *WoS* y *Scopus*.

Con ello se pretende, en primer lugar, dar respuesta a cuestiones relacionadas con la calidad de las herramientas gráficas para facilitar el acceso de los usuarios a su materia prima; en segundo lugar, y de forma coyuntural, establecer problemas de descripción que afectan a la localización y visualización de registros en pantalla. En definitiva, se trata de cuestiones que tienden a determinar si estas bases de datos cuentan con instrumentos para resolver los problemas de sinonimia y desambiguación de autores y conceptos, su influencia sobre la descripción de registros o su sensibilidad al uso de idiomas diferentes y signos diacríticos distintivos en las tareas de recuperación de documentos.

RESULTADOS

Interfaces y visualización de resultados

Ambas bases de datos estructuran su diseño mediante una barra de menús a través de la cual es posible acceder a las principales opciones de búsqueda. Por otra parte, y por lo que al idioma de consulta se refiere, las búsquedas en *WoS* deben realizarse obligatoriamente en inglés, quedando constancia de ello en las distintas pantallas que constituyen el sitio. *Scopus*, en cambio, permite el uso de varios idiomas en las consultas.

Las opciones de búsqueda de *WoS* son las siguientes:

- *Search* o búsqueda sencilla. La versión más reciente de *WoS* (5.5) facilita la búsqueda por ResearchID e incluye dos nuevas opciones de refinamiento. Una de ellas es la lematización, que tiene como propósito la extracción del lema o forma consensuada de un término que tiene posibles variantes (plural, masculino o femenino, conjugaciones verbales). La otra opción permite indicar el número de resultados que se visualizarán por página (Figura 1). El campo de búsqueda *Author* lleva asociado un enlace al asistente de nombre de autor *Author Finder*.

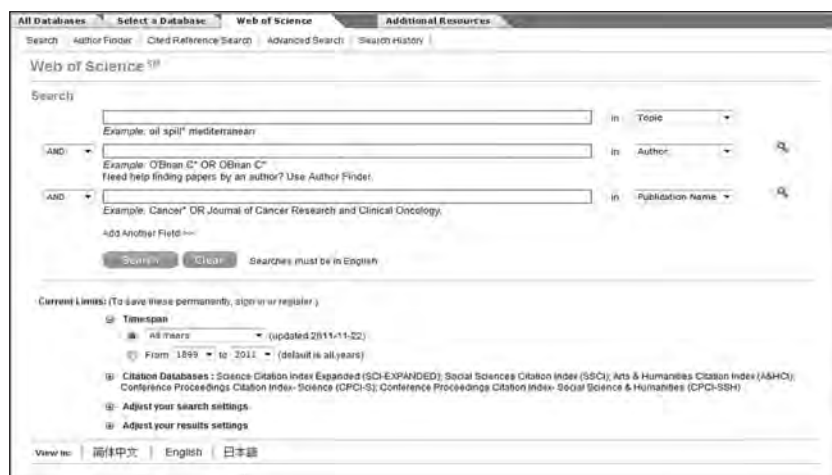


Figura 1. Búsqueda sencilla en *WoS* (versión 5.5). Fuente: *Web of Science*.

- *Author Finder*. En la versión anterior sólo se podía acceder a este asistente desde el campo *Author* en *Search*. Actualmente esta opción se incluye en la barra de navegación principal de la base de datos.

- *Cited reference*. Localiza artículos que citan un trabajo.
- *Advanced Search*. Permite hacer consultas complejas mediante el uso de etiquetas de campo, operadores booleanos y paréntesis.

Las principales opciones de búsqueda de *Scopus* son:

- *Document Search*. Es la opción de búsqueda sencilla. La pantalla está dividida en dos partes, en la superior se encuentra la zona de consulta y en la inferior se localizan las opciones de refinamiento (*Figura 2*).
- *Author Search*. Permite localizar todas las variantes del nombre de un autor registradas en *Scopus*.
- *Affiliation Search*. Localiza el nombre y las posibles variantes de una institución.
- *Advanced Search*. Como en el caso de *WoS*, es la opción más completa, permitiendo al usuario el uso personalizado de términos, campos y operadores booleanos.

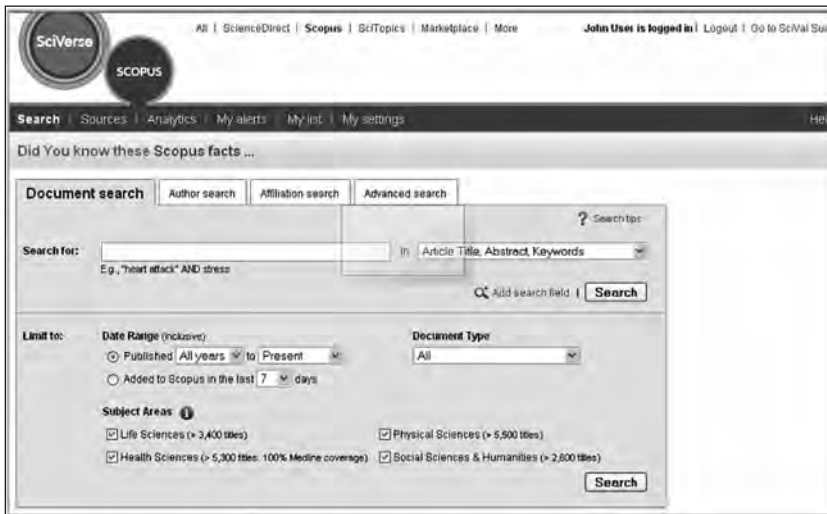


Figura 2. Búsqueda sencilla en *Scopus*. Fuente: *Scopus*.

El historial de búsqueda de *Scopus* (*Search history*) está presente en la parte inferior de la pantalla de todos los tipos de búsqueda, facilitando así su consulta en todo momento.



Figura 3. Presentación de resultados en WoS. Fuente: Web of Science.

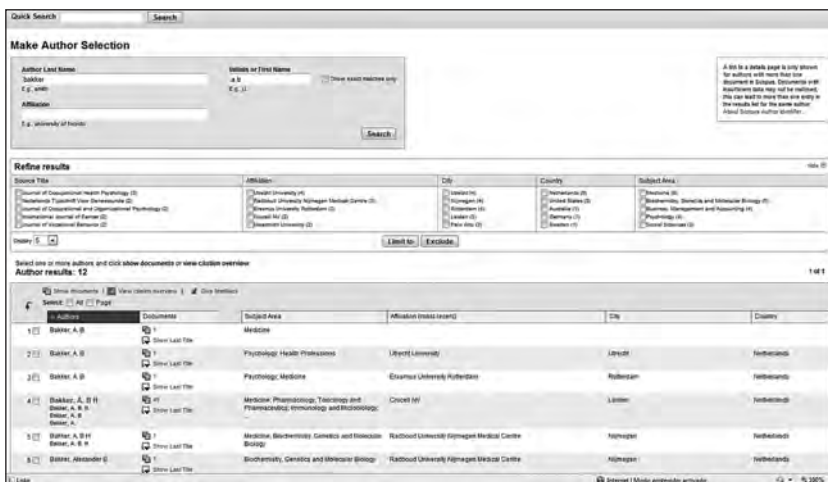


Figura 4. Presentación de resultados en Scopus. Fuente: Scopus.

La página de resultados de ambas bases de datos difiere significativamente en la distribución de la información en la pantalla (Figuras 3 y 4). Mientras que en WoS los resultados se distribuyen en forma de lista, en Scopus se encuentran en la parte inferior, distribuidos en forma tabular, junto con sus funciones asociadas. Por el contrario, las pantallas *Full Record* (WoS) y *Document Details* (Scopus) presentan una distribución de contenidos similar en ambas bases de datos.

Consultas y recuperación de información

Para probar ambos productos fue necesario el diseño de diversas estrategias de búsqueda que permitieran conocer el grado de adecuación de las interfaces a los usos y necesidades de los usuarios. Estas consultas no son exhaustivas y su única finalidad fue la detección de posibles problemas en la descripción de registros que pudieran afectar de forma determinante la localización y visualización de registros en pantalla. Para ello se plantearon ejemplos de interés para este propósito, con independencia de la temática de las publicaciones, la nacionalidad de los autores o sus filiaciones institucionales.

En la *Tabla I* se recogen los resultados de las búsquedas de autores realizadas en *Search* y *Author Finder*, en el caso de *WoS*, y *Document Search* y *Author Search* de *Scopus*.

En el caso del primer autor (Arnold B. Bakker) el número de documentos recuperados mediante búsquedas sencillas es similar. En *WoS* se recuperaron 168 resultados, de los que 44 no le correspondían. En *Scopus* 35 de los 177 registros no eran del autor.

En el segundo ejemplo (Evangelia Demerouti) el número de documentos recuperados también es parecido. En *WoS* se recuperaron 55 documentos de los que 7 no eran de la autora. En *Scopus* se recuperaron 64 documentos, de ellos 4 no eran de la profesora Demerouti.

En el tercer ejemplo (Ana Isabel Sanz Vergel) sí se aprecian variaciones en el número de resultados obtenidos en *WoS* en función de los términos empleados. Por el contrario, *Document Search* de *Scopus* recuperó el mismo número de documentos independientemente de los términos de búsqueda utilizados.

Para el cuarto autor (Daniel Peña) se realizaron consultas utilizando el término “peña” y “pena”. Mientras que *Scopus* recuperaba documentos con el primero, no ocurría lo mismo en *WoS*. Esto se debe a que la búsqueda en *Scopus* permite identificar signos diacríticos, hecho que redundaba en la mejora de los resultados de las búsquedas y evitan, por ejemplo, la inclusión de registros de autores apellidados “Pena” en el mismo conjunto de resultados.

En el quinto caso (Mariano Barbacid), tanto en *Author Finder* como en *Author Search* el número de resultados es similar, en ambos casos algunos resultados correspondían a la autora María del Mar Barbacid. En el sexto caso (Elizabeth Helen Blackburn), en las dos bases de datos todos los resultados correspondían a la autora cuando se incluían las dos iniciales del nombre.

Finalmente, para los dos últimos autores (François Barré-Sinoussi y José Manuel Sánchez Ron) la opción de búsqueda *Document Search* (*Scopus*) recuperó documentos empleando términos que contuvieran tilde, mientras que *Search* (*WoS*) no lo hizo.

Tabla I. Búsquedas comparadas de autores en WoSy Scopus.

Autores	WoS	Scopus	WoS	Scopus
	Search	Document Search	Author Finder	Author Search
Arnold B. Bakker	bakker a*b* 168 resultados 44 no del autor	bakker, a.b. 177 resultados 35 no del autor	bakker a 398 author sets	bakker a.b. 6 result. (190 doc) 2 result. (147 doc)
Evangelia Demerouti	demerouti e* 55 resultados 7 no de la autora	demerouti e. 64 resultados 4 no de la autora	demerouti e 17 author sets 10 de la autora	demerouti e. 4 result. (64 docs.) 2 result. (60 docs.)
Ana Isabel Sanz-Vergel	sanz vergel a* 14 resultados sanz-vergel a*i* 7 resultados sanz vergel a*i* 7 resultados	sanz vergel, a. sanz-vergel, a.i. sanz vergel, a.i. 12 resultados	sanz-vergel, a 7 author sets sanz-vergel a i sanz vergel a i 7 author sets todos de la autora	sanz-vergel a.i. sanz vergel a.i. 1 result. (12 docs.)
Daniel Peña	peña d* 0 resultados pena d* 532 resultados 69 del autor	peña, d. 973 resultados Sólo 60 del autor	peña d 0 author sets pena d 195 author sets 13 del autor	peña d. 91 result. (1242) 2 result. (78 docs.)
Mariano Barbacid	barbacid m* 269 resultados 9 resultados no del autor barbacid 279 resultados 11 resultados no del autor	barbacid, m. barbacid 235 resultados 10 no del autor	barbacid m 35 author sets 31 del autor	barbacid m 1 result. (232 docs.)
Elizabeth Helen Blackburn	blackburn e* 432 resultados 332 de la autora blackburn e*h* 259 resultados todos de la autora	blackburn, e. 507 doc. 227 de autora blackburn, e.h. 195 todos de la autora	blackburn eh 481 author sets 317 de la autora	blackburn e.h. 2 result. (225 docs.)
Françoise Barré-Sinoussi	barré-sinoussi f* 0 resultados barre-sinoussi f* 135 resultados 129 de la autora barre f* 93 resultados 0 de la autora	barré-sinoussi, f 250 doc. todos de la autora barré, f 339 docs. 246 de la autora	barre-sinoussi f 135 author sets todos de la autora	barré-sinoussi f 4 result. (261 docs.) 3 de la autora y uno dentro de la entrada Van Hung, Pham Pierre
José Manuel Sánchez Ron	sánchez ron j*m* 0 resultados sanchez ron j*m* 21 resultados sanchez-ron j*m* 26 resultados sanchez ron j*	sánchez-ron, j.m. sánchez ron, j.m. sánchez-ron,j. 18 docs. sánchez,j.m. 3634 docs. sanchez, j.	sanchez-ron j m 106 author sets 26 del autor	sánchez-ron j.m. 1 result. (16 docs.) todos del autor

	24 resultados sanchez j* 9 726 resultados	22 359 docs.		
--	---	--------------	--	--

La *Tabla II* presenta los resultados de las consultas en títulos realizadas a través de las opciones avanzadas en *Search (WoS)* y *Document Search (Scopus)*. El primer ejemplo tuvo como objeto la recuperación de documentos que contuvieran los términos “biological”, “damage” y “animals”. Ambas bases de datos recuperaron el mismo documento.

En el segundo ejemplo se trataba de localizar documentos que incluyeran las palabras “recovery” y “burnout” en el título, *abstract* y palabras clave, y cuyo autor no fuese Arnold B. Bakker. En *WoS* se obtuvieron 146 documentos, en *Scopus* 165.

El tercer ejemplo tuvo como propósito recuperar documentos que incluyeran el acrónimo “AIDS” en el título y no incluyeran el acrónimo “HIV”. *WoS* superó los 99 000 documentos y *Scopus* los 54 000.

El cuarto ejemplo trató de refinar la consulta anterior, recuperando documentos que contuvieran AIDS y “disease”, pero no HIV. El número de resultados se redujo ostensiblemente y la diferencia entre ambas bases de datos también.

El último ejemplo se centró en la recuperación de 8 trabajos de la profesora Sanz Vergel a través de sus títulos. En *Scopus* se recuperaron 7 documentos, mientras que en *WoS* sólo fue posible recuperarlos si la consulta incluía los títulos en inglés de los artículos en español.

Tabla II. Búsquedas comparadas de títulos en *WoS* y *Scopus*.

Ejemplos	Criterios de búsqueda	WoS	Scopus
1	WoS: TI=(biological) AND TI=(damage) AND TI=(animals) Scopus: TITLE=(BIOLOGICAL) AND TITLE=(DAMAGE) AND TITLE (ANIMALS)	1	1
	WoS: TI=(biological) AND TI=(damage) AND TI=(animal) Scopus: TITLE=(BIOLOGICAL) AND TITLE=(DAMAGE) AND TITLE (ANIMAL)	1	1
2	WoS: TS=(recovery) NOT AU=(BAKKER AB*) AND TS=(burnout) Scopus: TITLE-ABS-KEY (RECOVERY) AND TITLE-ABS-KEY (BURNOUT) AND NOT AUTH (BAKKER AB*)	146	165
3	WoS: TI=(AIDS) NOT TI=(HIV) Scopus: TITLE (AIDS) AND NOT TITLE (HIV)	99 160	54 007
4	WoS: TI=(AIDS) AND TI=(disease) NOT TI=(HIV) Scopus: TITLE (AIDS) AND TITLE (DISEASE) AND NOT TITLE (HIV)	1 897	1 398

5	Topic=		
	1. Work-family balance and energy: A day-level study on recovery conditions	1	1
	2. Effects of work-family on employees well-being: The moderating role of recovery strategies?	1	1
	3. Propiedades psicométricas de la versión española del cuestionario de interacción trabajo-familia	0	1
	4. The moderating effects of psychological detachment and thoughts of revenge in workplace bullying	1	1
	5. El efecto moderador de la autoestima en el acoso psicológico en el trabajo	0	1
	6. El efecto del acoso psicológico en el insomnio: el papel del distanciamiento psicológico y la rumiación	0	1
	7. La influencia del estrés de rol y la competencia emocional en el desgaste profesional y en el engagement	0	0
	8. Estrés traumático secundario: Personalidad y sacudida de creencias	0	1

DISCUSIÓN

Fortalezas y debilidades de las interfaces de Scopus y WoS

La interfaz de *Scopus* mantiene una apariencia uniforme y consistente en todas sus pantallas, presididas por la misma barra de navegación. La última versión de *Web of Science* analizada presenta, a diferencia de las anteriores, una interfaz mejorada y más integrada, como demuestra la inclusión de las opciones *Author Finder* y *Marked List* en la barra principal. Además, muestra de forma más accesible tanto la opción *Feedback* como el formulario de sugerencias. Sin embargo, el buscador *Scientific Web Plus* remite a una pantalla cuya apariencia difiere significativamente del resto de la interfaz, afectando la consistencia visual y el uso de las destrezas adquiridas previamente por el usuario.

La página de inicio de *Scopus* es muy simple, conforme al estilo de los principales buscadores web, con un solo campo de texto por defecto, sin operadores booleanos, con el botón *Search* fácilmente localizable y la información distribuida de forma sencilla y accesible en la pantalla. Además, el historial de búsqueda de *Scopus* está al alcance en todo momento, con independencia del tipo de búsqueda empleado, mientras que en *WoS* es necesario abandonar la página de consultas para poder acceder a las pesquisas realizadas a lo largo de la sesión.

La forma de realizar las consultas es común en todas las opciones de búsqueda de ambas bases de datos. La última versión de *WoS* demuestra el gran esfuerzo realizado por Thomson Reuters para que esto sea así, quizá en respuesta a las facilidades de uso que ofrece el producto de su competidor. Muestra de ello es la inclusión de *Author Finder* como una opción más de búsqueda y no como mero asistente cuyo manejo difería significativamente del requerido por el resto de opciones en versiones anteriores.

Sobre los resultados en búsquedas de autores

En relación al primer autor, la opción de búsqueda *Author Search (Scopus)* recuperó 6 *sets* de documentos de los que sólo 2 pertenecían al autor. El resto de variantes de nombre no se correspondían con el autor buscado:

<p>Bakker, Alexander B H (entrada principal) Bakker, A. B H Bakker, A. B. Bakker, A.</p>
--

La entrada principal, en negrita, obtuvo 29 documentos de los que ninguno era del autor. Un fenómeno similar se observó en los casos de Barbacid y Demerouti. Por su parte, la opción *Author Search* de *WoS* recuperó un conjunto de resultados muy elevado (398), dificultando la recuperación de la obra debido a la excesiva dispersión de los trabajos.

En el caso de la profesora Demerouti, las opciones de refinamiento de resultados de *WoS* no funcionaron debido a que bajo la entrada DEMEROUTI, E. se encuentran indizadas dos autoras distintas. En este caso, sólo la combinación de autor e institución permitiría lograr la desambiguación de ambas autoras.

En relación al caso de Sanz Vergel, no es posible recuperar todos sus trabajos en *Scopus* por una descripción incorrecta de la autora (los documentos fueron recogidos bajo la forma Vergel, A.I.S., en la que Sanz se considera como un tercer nombre y no como apellido). En el caso de *WoS* la incorrecta descripción de la autora hace imposible la recuperación de 7 de sus estudios, de los que hay constancia de localización sin inconvenientes en versiones anteriores de la base de datos.

La consulta en *WoS* del profesor Peña a partir del término “Peña” recuperó documentos de autores con apellidos “Peña” y “Pena” indistintamente. En el caso de Barbacid, la única entrada a este autor en *Author Search (Scopus)* es Barbacid, Mariano Mar M., por lo que incluye dentro de este conjunto las obras tanto de Mariano como de María del Mar.

En relación con la profesora Blackburn, la precisión de los resultados obtenidos varía en función de si se utiliza o no la inicial del segundo nombre. En la búsqueda sencilla los resultados correspondían en su totalidad a la autora si se incluían las dos iniciales del nombre de pila. Por otro lado, el número de *sets* recuperados a través de *WoS* era significativamente mayor, 481 en *Author Finder*, frente a los 2 que ofrecía *Author Search (Scopus)*.

Con respecto a la profesora Barré-Sinoussi la opción de búsqueda *Author Finder (WoS)* recuperó 135 *sets* de la autora frente a los 4 de *Author Search (Scopus)*. Sorprende el hecho de que en uno de los resultados de *Scopus* el trabajo es del autor Hung, Pham Pierre, error que posiblemente provenga de una defectuosa introducción de los datos.

Sobre los resultados en búsquedas en títulos

En relación al primer ejemplo se ha comprobado que ambas bases de datos utilizan la lematización de términos, que permite la recuperación de documentos utilizando formas alternativas para la búsqueda de un término concreto; por ejemplo, permite recuperar términos en singular y en plural indistintamente.

En el tercer ejemplo se recuperaron documentos con el término “AIDS” referidos a la tercera persona del singular del verbo inglés *to aid* (ayudar). Para paliar este problema y reducir el ruido de la tercera consulta, en el cuarto ejemplo se añadió el término “disease” con el fin de recuperar documentos referidos sólo a la enfermedad.

Por último, el ejemplo 5 sirvió para comprobar que son 7 los trabajos de Sanz Vergel incluidos en *WoS* y *Scopus*, pero parte de ellos habían sido asignados a la variante de nombre Vergel, A.I.S., dificultando su recuperación y enfatizando los problemas ya evidenciados en las búsquedas de autor de esta profesora.

CONCLUSIONES

Las consultas sobre autores han permitido constatar que la principal finalidad del asistente de la opción de búsqueda *Author Finder (WoS)* es la identificación de las posibles variantes del nombre de un autor, de forma similar a como lo hacía en versiones anteriores *Scopus* con su opción *Author Search*. Sin embargo, no es infrecuente que los resultados obtenidos agrupen variantes de nombres que corresponden a autores diferentes (como sucede en *Scopus* con Barbadic o en *WoS* con Demerouti), hecho que requiere de un esmerado

trabajo de desambiguación. Por otro lado, en demasiadas ocasiones *Author Finder* (*WoS*) obtiene un gran número de *author sets*, lo que dificulta la recuperación de todas las obras de un determinado autor debido a su excesiva dispersión. En estos casos es recomendable incluir información adicional, como la filiación del autor, para reducir al máximo estos inconvenientes.

Por otra parte, tanto en *WoS* como en *Scopus* es importante la inclusión del segundo apellido en las búsquedas de autores latinos o con apellidos compuestos. En las búsquedas de autores anglosajones la inicial del segundo nombre es de vital importancia para evitar ambigüedades indeseadas.

El guión que une los apellidos de los autores es tratado como un espacio en ambas bases de datos y, por tanto, no es un factor que actualmente influya en la recuperación de información aun cuando es necesario para la correcta identificación y adscripción de los trabajos y para evitar que el primer apellido de autores latinos se considere parte del nombre, algo típico en autores anglosajones (el mejor ejemplo es el caso Sanz-Vergel en *WoS*).

Gracias a las estrategias planteadas ha sido posible comprobar que ambas bases de datos presentan severos problemas de sinonimia y de descripción incorrecta de autores. En este sentido, herramientas como la lematización (recientemente incorporada a *WoS* en su última versión), que permite recuperar documentos ampliando los términos y conceptos, o iniciativas para la identificación inequívoca de autores como IraLis, Orcid o ResearchID tratan de paliar estos inconvenientes.

La imposibilidad de *WoS* para incluir tildes, eñes y otros signos diacríticos existentes en idiomas distintos del inglés constituye un inconveniente en la realización de las búsquedas y en la obtención de resultados precisos, además de ser una desventaja frente a *Scopus* en el mercado internacional no anglosajón. Esta circunstancia es de difícil solución, más allá de incluir información adicional como los datos institucionales en el caso de autores para reducir el exceso de documentos no relevantes.

Han sido numerosas las mejoras que se han producido en la versión 5.5 de *WoS* relacionadas tanto con el proceso de recuperación de información como con su interfaz gráfica. Sin embargo, son necesarios ajustes adicionales para garantizar la recuperación de las obras de todos los autores, ya que paradójicamente la versión actual impide la recuperación de obras de algunos autores, correctamente descritas y fácilmente recuperables en versiones anteriores.

Elsevier ha creado un producto con una gran interfaz gráfica que, teniendo como referente a Thomson Reuters, ha sabido conservar las virtudes de su predecesor y mejorar sus inconvenientes. *WoS*, por su parte, ha mejorado la interfaz en su última versión adoptando algunos de los aspectos más destacados de *Scopus*, como la consistencia visual y la posibilidad de recuperar

más de 100 000 documentos en una única búsqueda. Desde el punto de vista del contenido, Elsevier debe mejorar la cobertura cronológica en general y la temática en áreas como Ciencias Sociales y Artes y Humanidades de *Scopus* si aspira a competir en igualdad de condiciones con *WoS* en el mercado de las bases de datos de producción científica y tecnológica.

Finalmente, no debemos olvidar que todos los problemas mencionados, en especial los que atañen a la recuperación de registros adscritos erróneamente por falta de desambiguación o por dificultades para discernir términos con signos diacríticos distintivos, pueden afectar los frecuentes trabajos derivados de estas herramientas, por lo que su solución no sólo redundaría en el beneficio de ambas plataformas, sino también en la calidad, robustez y verosimilitud de las docenas de estudios anuales basados en sus datos que afectan a investigadores, instituciones, países y revistas.

BIBLIOGRAFÍA

- Bar-Ilan, J. (2010), "Citations to the 'Introduction to infometrics' indexed by *WoS*, *Scopus* and *Google Scholar*", en *Scientometrics*, 82 (3): 495-506.
- Deis, L. & Goodman, D. (2005), "*Web of Science* (2004 version) and *Scopus*", en *The Charleston Advisor*, 6 (3): 5-21.
- Escalona Fernández, M. I.; Lagar Barbosa, P. & Pulgarín Guerrero, A. (2010), "*Web of Science vs. SCOPUS*: un estudio cuantitativo en Ingeniería Química", en *Anales de Documentación*, 13: 159-175.
- FECYT (2009), *Portal de acceso a la Web of Knowledge*, Madrid: Fundación Española para la Ciencia y la Tecnología. Disponible en: http://www.accesowok.fecyt.es/?page_id=2329 [Fecha de consulta: 06/11/2012].
- Fingerman, S. (2005), "*SCOPUS*: Profusion and confusion", en *Online*, 29 (2): 36-38.
- (2006), "*Web of Science* and *Scopus*: current features and capabilities", en *Issues in Science and Technology Librarianship*, 48. Disponible en: <http://www.istl.org/06-fall/electronic2.html> [Fecha de consulta: 06/11/2012].
- Garfield, E. (1955), "Citation indexes for science: a new dimension in documentation through association of ideas", en *Science*, 122 (3159): 108-111.
- (1964), "Science Citation Index: a new dimension in indexing", en *Science*, 144 (361): 649-654.
- Jacso, P. (2005), "As we may search-Comparison of major features of *Web of Science*, *Scopus*, and *Google Scholar* citation-based and citation-enhanced databases", en *Current Science*, 89 (9): 1537-1547.

- LaGuardia, C. (2005), "E-Views and Reviews: *Scopus* vs. *Web of Science*", en *Library Journal*, January. Disponible en: <http://www.libraryjournal.com/article/CA491154.html> [Fecha de consulta: 06/11/2012].
- Leydesdorff, L. (2012), "World shares of publications of the USA, EU-27, and China compared and predicted using the new *Web of Science* interface versus *Scopus*", en *El Profesional de la Información*, 21: 43-49.
- Mingers, J. & Lipitakis, E. (2010), "Counting the citations; a comparison of *Web of Science* and *Google Scholar* in the field of business and management", en *Scientometrics*, 85 (2): 613-625.
- Santa, S. & Herrero-Solana, V. (2010), "Cobertura de la ciencia de América Latina y el Caribe en *Scopus* vs. *Web of Science*", en *Investigación Bibliotecológica*, 24 (52): 13-27.
- Scopus* (2010), *User guide*, Amsterdam: Elsevier.
- Vieira, E. S. & Gomes, J. (2009), "A comparison of *Scopus* and *Web of Science* for a typical university", en *Scientometrics*, 81 (2): 587-600.

