

Las universidades públicas españolas en Google Scholar: presencia y evolución de su publicación académica web

Por Enrique Orduña-Malea, Jorge Serrano-Cobos y Nuria Lloret-Romero

Resumen: Se estudia la validez de Google Scholar como servicio que refleja adecuadamente la producción científica de una universidad, comparando su cobertura de las universidades públicas españolas con la mostrada por Scopus. Se analiza la presencia y evolución de la documentación científica de las sedes web de estas universidades desde enero hasta julio (ambos meses incluidos) de 2009. Los resultados muestran que, pese a encontrar cierta interrelación en la productividad de las instituciones en Scopus y Scholar, entre ambos sistemas existen grandes diferencias en los resultados totales que invalidan a este último como un reflejo válido de la producción universitaria. Finalmente, la mayoría de las webs universitarias analizadas presentan un crecimiento positivo aunque moderado en Scholar, durante el período estudiado.

Palabras clave: Sedes web universitarias, Google Scholar, Producción científica, Universidades públicas españolas, Rankings de universidades.

Title: Spanish public universities in Google Scholar: presence, evolution and coverage of their scientific output

Abstract: The validity of Google Scholar as a service that reflects properly the scientific output of a university is analyzed, comparing its coverage of Spanish public universities with that of Scopus. The presence and evolution of scientific documentation on the websites of the Spanish public universities from January to July (both months inclusive) 2009 is also studied. The results show that, despite finding some interrelationship between Scholar and Scopus concerning the productivity of institutions, there are large differences in the total results that override the latter as a valid reflection of university productivity. Finally, Scholar shows positive, albeit modest, growth during the studied period for most university websites analyzed.

Keywords: Academic websites, Google Scholar, Scientific output, Spanish public universities, University rankings.

Orduña-Malea, Enrique; Serrano-Cobos, Jorge; Lloret-Romero, Nuria. "La universidades públicas españolas en Google Scholar: presencia y evolución de su publicación académica web". *El profesional de la información*, 2009, septiembre-octubre, v. 18, n. 5, pp. 493-500.

DOI: 10.3145/epi.2009.sep.02



Enrique Orduña-Malea es ingeniero técnico de telecomunicaciones, esp. sonido e imagen, y técnico especialista en imagen fílmica, licenciado en documentación y máster en contenidos y aspectos legales en la sociedad de la información (especialidad en contenidos multicanal), por la Universidad Politécnica de Valencia (UPV). Investigador predoctoral en la división de investigación en tecnología e información (Calsi) del Instituto de Diseño y Fabricación (IDF) de la UPV. Redactor jefe del Anuario ThinkEPI, webmaster y miembro del think-tank sobre estudios de estrategia y prospectiva de la información, y miembro del equipo de trabajo del gestor social de noticias DocuMenea.



Jorge Serrano-Cobos, licenciado en documentación por la Univ. de Granada, ha trabajado como arquitecto de información y consultor en diseño de sistemas de información e intranets y especialista en marketing online para compañías como Planeta DeAgostini, Serikat Consultoría Informática y Google Inc. Es director del Departamento de Contenidos de MASmedios.com. Sus intereses profesionales van desde la arquitectura de información, el diseño de interacción (incluyendo opacs, juegos multimedia, etc.) al estudio de information seeking behaviour, web mining, search analytics, e-commerce y e-marketing. Miembro del think-tank sobre estrategia y prospectiva de la información ThinkEPI.



Nuria Lloret-Romero es doctora por la UPV, master en desarrollo de sist. de información online, soportes magnéticos cd-rom (Masdoc) de la UB, y master en dirección financiera de empresas (Fund. Estudios Bursátiles de Valencia). Prof. titular y directora del Dpto. de Comunicación Audiovisual, Documentación e Historia del Arte (Dcadha) de la UPV. Subdirectora del Inst. de Investigación Univ. IDF y directora del grupo de investigación Calsi, de dicho Instituto. Evaluadora de la UE (programa IST). Socia fundadora de MASmedios, vicepres. de la Fed. Valenciana de empresas de Servicios Avanzados y asesora del Comité de Tecnologías de la Información de la Cámara de Comercio de Valencia.

1. Introducción

La cibermetría como disciplina da un salto cuantitativo en 2004 con la puesta en marcha del *Ranking web of world universities*¹, primer ranking de universidades de cobertura mundial basado en criterios meramente cibernéticos (Aguillo; Granadino; Ortega; Prieto, 2006; Aguillo; Ortega; Fernández, 2008).

Pese a la existencia de otros rankings^{2,3} y estudios acerca de la presencia web de las universidades mediante la información proporcionada por los buscadores (Bar-Ilan, 2004; Payne; Thelwall, 2004), esta iniciativa marca un hito debido a sus objetivos, cobertura, impacto mediático y orientación documental.

Tal como los responsables del proyecto expresan¹, “la posición en el ranking resume el rendimiento global de la universidad, aporta información para estudiantes o profesores y refleja el compromiso académico con la difusión del conocimiento científico”.

En este contexto, el concepto de “sede web académica” debe entenderse como un conjunto de webs ligadas jerárquicamente a una página principal, identificable por una url, y que forma una “unidad documental” reconocible e independiente de otras por su representatividad institucional (Aguillo, 1998).

Por tanto, asumiendo la web como un canal de comunicación entre la universidad y sus usuarios, la calidad y cantidad de documentación accesible desde su dominio web (que podrá alojar a una o más sedes web) serán indicadores de su grado de comunicación; siendo éste, a su vez, indicador parcial de su calidad como institución.

No obstante, la documentación publicada dentro de un dominio web académico es heterogénea, pudiendo ser clasificada fundamentalmente en administrativa, bibliográfica, científica, docente e informativa.

“Google Scholar destaca por la inclusión de citas y por la mayor estabilidad de los resultados frente a los motores de búsqueda generalistas”

La cantidad y calidad de cada uno de estos tipos de documentación sirven potencialmente para valorar, entre otros criterios, las diferentes misiones de la universidad (docencia, investigación y transferencia), o al menos la sub-misión de su puesta online.

Sin embargo, los motores de búsqueda no discriminan apropiadamente estas clases de documentación automáticamente ni proporcionan comandos adecua-

dos para filtrarla de forma conveniente a nuestros propósitos (es decir, los ítems se encuentran etiquetados, pero no es posible filtrar por ellos), siendo este hecho delicado desde el punto de vista cuantitativo.

En 2006 se lanza *Google Scholar* con la intención de resolver algunos de estos problemas, así como de indizar la llamada internet invisible. Sus resultados pasan supuestamente un filtro por el que el sistema considera que son científicos, o se tienen al menos ciertas garantías. Esto se logra mediante acuerdos comerciales con proveedores, lo que explica en parte su cobertura.

De ser cierto significa que *Scholar*, con los comandos de búsqueda apropiados, podría ser usado como instrumento para medir la cantidad de documentación científica que las universidades publican en la Web y/o dentro su dominio web.

Esta última posibilidad permitiría medir parte del compromiso de comunicación entre la universidad y sus usuarios comentado anteriormente, dejando la documentación docente y de transferencia aparte. Este aspecto llevó a los responsables del ranking web a adoptar *Google Scholar* como uno de sus indicadores, a partir de su edición de 2006.

Google Scholar destaca sobre todo por la inclusión de citas (es la única base de datos multidisciplinar que las suministra, junto a *Scopus* y *WoS*) y por la mayor estabilidad de los resultados frente a los motores de búsqueda generalistas.

Aun así, se detecta una serie de inconvenientes en el uso de *Scholar* como herramienta bibliométrica (Torres-Salinas; Ruiz-Pérez; Delgado-López-Cózar, 2010):

- No todos los documentos indizados se pueden considerar académicos.
- No proporciona ninguna información sobre las fuentes que cubre.
- Cobertura disciplinar desigual en ciencias sociales y humanidades.
- Carencia de comandos básicos de consulta para realizar estudios bibliométricos.
- Ausencia absoluta de normalización en los datos (autores y fuentes).

A esto se unen otros problemas en la realización de análisis de citas (Jacsó, 2008a; Harzing; Van der Wal, 2008).

Este trabajo, por su parte, plantea otra serie de cuestiones relacionadas con la cobertura, visibilidad y evolución de los contenidos alojados en los dominios web de las universidades españolas:

Universidades	Código	url	Universidades	Código	url
Universitat Autònoma de Barcelona	UAB	http://www.uab.es/	Universidad de Oviedo	UO	http://www.uniovi.es/
Universidad Autónoma de Madrid	UAM	http://www.uam.es/	Universidad de Salamanca	USAL	http://www.usal.es
Universidad Carlos III de Madrid	UC3M	http://www.uc3m.es	Universidad de Santiago de Compostela	USC	http://www.usc.es/
Universidad Complutense de Madrid	UCM	http://www.ucm.es/	Universidad de Sevilla	US	http://www.us.es
Universidad de Alcalá	UAH	http://www.uah.es/	Universitat de València	UV	http://www.uv.es
Universitat d'Alacant	UA	http://www.ua.es/	Universidad de Valladolid	UVA	http://www.uva.es/
Universidad de Almería	UAL	http://www.ual.es/	Universidad de Vigo	UVI	http://www.uvigo.es/
Universitat de Barcelona	UB	http://www.ub.edu	Universidad de Zaragoza	UZ	http://www.unizar.es/
Universidad de Burgos	UBU	http://www.ubu.es	Universidad del País Vasco	EHU	http://www.ehu.es
Universidad de Cádiz	UCA	http://www.uca.es/	Universidad Internacional de Andalucía	UIA	http://www.unia.es/
Universidad de Cantabria	UC	http://www.unican.es	Universitat Jaume I	UJI	http://www.uji.es/
Universidad de Castilla-La Mancha	UCLM	http://www.uclm.es/	Universidad Miguel Hernández de Elche	UMH	http://www.umh.es/
Universidad de Córdoba	UCO	http://www.uco.es/	Univ. Nacional de Educación a Distancia	UNED	http://www.uned.es
Universidad de Extremadura	UNEX	http://www.unex.es/	Universidad Pablo de Olavide	UPO	http://www.upo.es/
Universitat de Girona	UDG	http://www.udg.edu/	Universidad Politécnica de Cartagena	UPCT	http://www.upct.es/
Universidad de Granada	UGR	http://www.ugr.es	Universitat Politècnica de Catalunya	UPC	http://www.upc.es/
Universidad de Huelva	UHU	http://www.uhu.es/	Universidad Politécnica de Madrid	UPM	http://www2.upm.es
Universidad de Jaén	UJA	http://www.ujaen.es/	Universitat Politècnica de València	UPV	http://www.upv.es
Universidad de La Coruña	UDC	http://www.udc.es	Universitat Pompeu Fabra	UPF	http://www.upf.edu/
Universidad de La Laguna	ULL	http://www.ull.es	Universidad Pública de Navarra	UPNA	http://www.unavarra.es/
Universidad de La Rioja	UR	http://www.unirioja.es/	Universidad Rey Juan Carlos	URJC	http://www.urjc.es/
Universitat de les Illes Balears	UIB	http://www.uib.es	Universitat Rovira i Virgili	URV	http://www.urv.es/
Universidad de Las Palmas de Gran Canaria	ULPGC	http://www.ulpgc.es/			
Universidad de León	UL	http://www.unileon.es/			
Universitat de Lleida	UDL	http://www.udl.es/			
Universidad de Málaga	UMA	http://www.uma.es/			
Universidad de Murcia	UM	http://www.um.es/			

Tabla 1. Universidades públicas españolas (fuente: Crue)

– ¿Hay alguna correlación entre la presencia web de las universidades españolas en *Google Scholar* y su producción académica?

– ¿Qué volumen de documentación presentan las universidades españolas a partir de sus dominios web en *Google Scholar*?

De esta forma, los objetivos principales son los siguientes:

– comparar la cobertura que proporciona *Google Scholar* de las universidades españolas con la mostrada por *Scopus*, y

– analizar la presencia y evolución de la documentación alojada en los dominios web de las universidades españolas.

Este trabajo se centra únicamente en las universidades públicas españolas, excluyendo a las privadas

con el propósito de tratarlas de forma exclusiva en el futuro.

2. Metodología

2.1. Correlación entre *Scholar* y *Scopus*

La productividad académica de las universidades se extrae directamente de *Scopus* mediante los datos del ranking *Scimago*⁴, relativos a la producción total de las universidades públicas españolas (tabla 1), desde 2003 hasta 2007.

Para la publicación web, se consultan todos los dominios web de las universidades públicas españolas⁵ desde *Scholar* para el mismo período de tiempo. Este proceso se lleva a cabo en junio de 2008.

La consulta se realiza mediante el comando “site” al dominio principal de cada universidad. En el caso

de la *UB* y la *UDG*, se combinan los resultados de los dominios .es y .edu.

Finalmente, los datos de ambas series (*Scopus* y *Scholar*) se exportan a una hoja de cálculo, donde se obtiene la correlación entre ambas.

Se debe indicar que no se recuperan los artículos publicados por autores de una determinada universidad recogidos en *Scholar*, sino los artículos publicados en la parte pública del dominio web.

Esto significa que no se evalúa la producción científica de la universidad, sino más bien su compromiso de facilitar documentación científica a su comunidad mediante la Web y en concreto, de *Scholar*.

2.2. Presencia y evolución

El marco de trabajo se sitúa en el análisis de los dominios web de las universidades públicas españolas, desde enero hasta julio (ambos meses incluidos) de 2009.

Se recogen datos de todas las universidades públicas españolas, a excepción de la *UPF*, *URV* y *UIA*, de las que no se dispone de información debido a problemas técnicos⁶, por lo que no se consideran.

Los datos se obtienen de nuevo mediante la consulta directa a *Scholar*, pero sin ninguna restricción temporal. Posteriormente se exportan, durante la primera semana de julio, a una hoja de cálculo para ser tratados.

La toma de datos es mensual; se realizan varias catas por mes y se obtiene la media de dichas muestras. Esta operación se repite para cada dominio web de cada universidad, obteniendo un valor en bruto por mes. De éstos se obtiene la media mensual, denominada *Scholar(m)*, el recorrido y un valor llamado de representatividad relativa en *Scholar* (*RScholar*).

El recorrido se obtiene como la resta del valor último (junio) menos el primero (enero) e indica la ganancia o pérdida en los meses analizados.

RScholar se calcula normalizando de 0 a 100 los valores en bruto de cada dominio web para cada mes. Se otorga el valor 100 al dominio web de la universidad con valor máximo, y el resto se normalizan de forma proporcional. Finalmente, la media de los 6 valores normalizados (uno por mes) de cada dominio web proporciona el valor de *RScholar*.

Se recogen asimismo los datos relativos a la resta entre los valores totales en *Google Scholar* y los valores delimitados temporalmente a partir del año 2000 (en adelante *Scholar recent*, *Gr*), para cada mes y dominio web.

De esta forma se premia la publicación web reciente de las universidades y no el simple volcado retrospectivo

“Pese a que las posiciones no coinciden, 7 de las 10 primeras universidades en *Scopus* se encuentran igualmente entre las 10 primeras en *Scholar*”

vo de material académico, que favorece a las universidades más antiguas al disponer de más documentación acumulada.

A partir de *Google recent* se obtiene la media de los 6 valores mensuales obtenidos, *Gr(m)*, así como los correspondientes a su representatividad relativa (*RGr*).

Finalmente, se indica que los resultados obtenidos no deben ser entendidos como aproximaciones a la calidad de las universidades o clasificaciones ordenadas donde un primer puesto signifique ser “mejor”, sino simplemente como reflejo de lo que *Google Scholar* está mostrando de los dominios web académicos, síntoma de la buena o mala comunicación vía web que las universidades públicas españolas tienen con sus usuarios.

3. Resultados

3.1. *Scopus* vs *Scholar*

La tabla 2 muestra las diez primeras universidades públicas españolas en productividad en *Scopus* y *Scholar* para 2003-2007.

Pese a que las posiciones no coinciden, 7 de las 10 primeras universidades en *Scopus* se encuentran igualmente entre las 10 primeras en *Scholar*. Aun así, se observan ciertos comportamientos divergentes: mientras la *UAM*, *USC* y *US* pierden representatividad en *Scholar*, la *UDG* y *UM* la mejoran. La *UPCT*, *UHU* y *UCO* mejoran igualmente, aunque quedan fuera de las 10 primeras.

Los casos de la *UR* y *UCM* son especialmente reseñables pues la cantidad de publicación web recuperada en sus dominios web es mayor que la producción académica recogida por *Scopus* en el mismo período. Es más, la *UR* aparece en el último puesto en productividad científica en *Scopus* mientras que se encuentra en primera posición en *Scholar*. Estos resultados se atribuyen, entre otros factores, a la indización de la base de datos *Dialnet* de la *UR* por parte de *Scholar*.

El caso de la *UCM* es diferente pues aun siendo la tercera universidad en producción científica, sus resultados en *Scholar* son desproporcionados debido a la indización de su servicio de sumarios *Compludoc*.

Por su parte, la figura 1 muestra la correlación entre los datos de *Scholar* y *Scopus* (dejando fuera de los cálculos a la *UR* y *UCM* por sus elevados valores en *Scholar*).

Universidades	Scopus	Universidades	Scholar
UB	11.243	UR	30.700
UAB	9.364	UCM	16.300
UCM	8.911	UM	1.550
UAM	7.851	EHU	1.490
UV	7.621	UB	1.445
UPC	6.884	UGR	1.280
UGR	5.774	UAB	1.070
USC	5.097	UPC	778
US	4.927	UDG	671
EHU	4.675	UV	661

Tabla 2. Productividad (Scopus) y publicación web (Scholar) de las 10 primeras universidades públicas españolas (2003-2007)

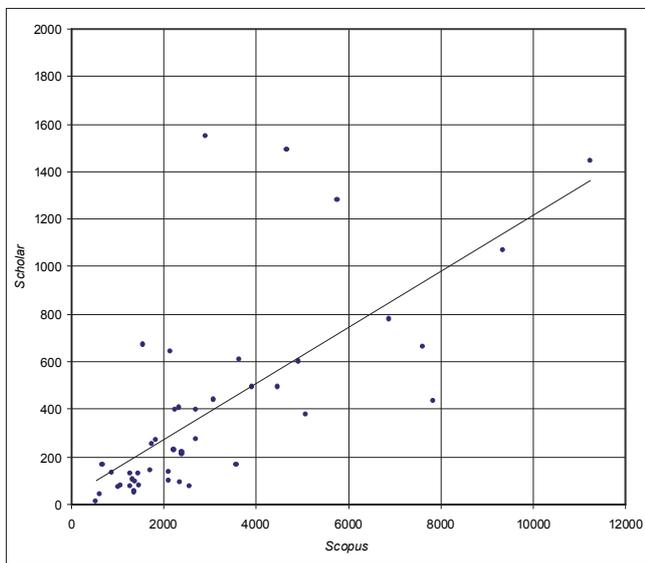


Figura 1. Correlación entre Scholar y Scopus (2003-2007)

Se obtiene una curva de regresión positiva, es decir, la mayoría de las universidades más productivas en Scopus lo son igualmente en Scholar, con un coeficiente de correlación de Pearson de 0,70 (si se tienen en cuenta UR y UCM resulta un coeficiente de sólo 0,08).

3.2. Producción y publicación web

La tabla 3 muestra los 46 dominios web estudiados, ordenados de forma descendente según su valor de RScholar. Así mismo se muestran los valores de representatividad extrayendo del cálculo a la UR y UCM (RScholar*) y la media de los valores brutos mensuales. Completan la tabla los valores de recorrido, RGr y Gr(m).

La UR y UCM aparecen las primeras en el ranking de RScholar, a una distancia considerable del resto. El dominio de la primera, de hecho, es abrumador incluso respecto a la UCM. Tras estas dos universidades, el resto aparece con valores más similares.

Respecto a RGr, los resultados muestran que 31 universidades (de 46) mejoran su representatividad al tener en cuenta únicamente la publicación web reciente, entre ellas las 17 primeras del ranking RScholar. Destaca la UCM (que incrementa cerca del 35% de representatividad total), así como la UM, EHU y UDG.

Si observamos el recorrido, vemos caídas de la UB (5.540 documentos menos que en enero), UM y UB. En cuanto a dominios web de universidades con recorrido positivo, destacan UGR y UJI. Así mismo, merecen atención los comportamientos de la UVA y UL (sin apenas cambios), y de la UDC (crecimiento 0).

Es igualmente de interés la relativamente buena posición de las universidades politécnicas, con dos entre las 10 primeras (UPC y UPM), y la UPV cerca, en la decimosegunda posición.

En la figura 2 se muestra la distribución –exponencial, o de larga cola-, de la media de registros por universidad.

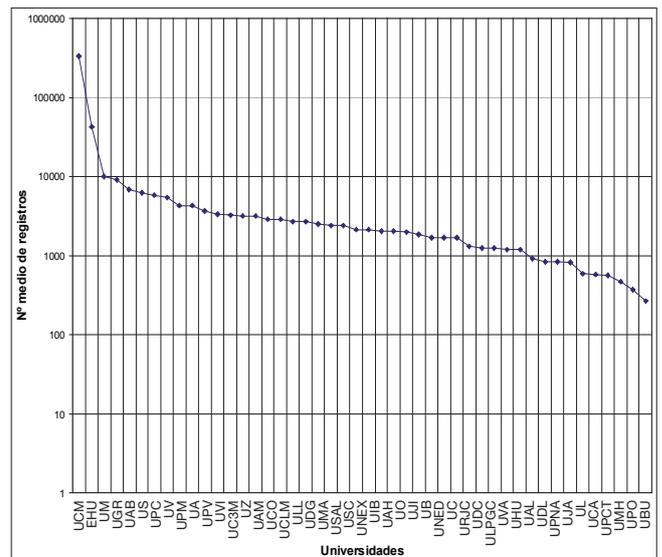


Figura 2. Distribución del número medio de registros por universidad (enero-junio 2009)

3.3. Evolución

En la figura 3 se detalla la evolución del tamaño de las 10 primeras universidades, a excepción de la UR y UCM.

En términos generales, UM, EHU y UGR sobresalen respecto al resto. Todos los dominios web, no obstante, presentan unas tendencias estables y ligeramente positivas durante todos los meses estudiados a excepción de marzo, donde se producen cambios importantes en prácticamente todos los dominios web, que determinan en buena medida los valores de recorrido obtenidos. Destaca el crecimiento de la UGR, US, UV y UA, y las caídas en la UM, UAB y UPC.

Universidades	RScholar	RScholar*	Scholar(m)	Recorrido	RGr	Gr(m)
UR	100,00	-	334.000	55.000	100,00	48.600,00
UCM	12,82	-	42.450	-2.300	47,78	19.450,00
EHU	3,03	98,58	10.087	380	6,00	2.561,70
UM	2,74	88,35	9.048	-1.900	9,37	4.000,00
UGR	2,07	67,68	6.922	1.140	4,02	1.795,00
UAB	1,88	60,75	6.218	-730	3,31	1.426,70
US	1,75	56,97	5.828	600	1,87	828,50
UPC	1,65	53,59	5.485	-370	2,47	1.050,00
UV	1,31	42,63	4.360	740	1,99	893,50
UPM	1,29	42,14	4.312	430	1,57	764,50
UA	1,11	36,17	3.698	830	1,15	528,83
UPV	0,99	32,26	3.300	830	1,64	720,83
UVI	0,97	31,77	3.250	410	1,27	536,00
UC3M	0,95	31,04	3.175	50	1,44	641,33
UZ	0,95	30,86	3.157	390	1,40	623,17
UAM	0,88	28,57	2.922	450	1,26	573,17
UCO	0,87	28,48	2.913	480	2,22	947,50
UCLM	0,81	26,44	2.705	210	0,66	299,33
ULL	0,80	26,22	2.682	480	0,78	336,83
UDG	0,75	24,51	2.505	480	2,66	1.208,20
UMA	0,73	23,60	2.415	-80	1,25	550,83
USAL	0,71	23,30	2.383	280	0,87	399,00
USC	0,64	21,00	2.148	250	1,32	580,00
UNEX	0,64	20,94	2.142	410	0,46	197,67
UIB	0,61	19,86	2.032	250	0,85	382,00
UAH	0,61	19,74	2.020	130	0,61	279,17
UO	0,60	19,66	2.012	200	0,60	267,50
UJI	0,55	18,22	1.863	1.200	0,98	519,17
UB	0,53	16,34	1.678	-5.540	0,98	316,50
UNED	0,51	16,69	1.707	180	1,01	467,67
UC	0,51	16,62	1.700	100	0,27	122,33
URJC	0,39	12,68	1.297	320	0,34	154,50
UDC	0,37	12,07	1.235	0	0,40	174,50
ULPGC	0,37	12,16	1.243	361	0,24	106,50
UVA	0,36	11,64	1.192	10	0,22	99,83
UHU	0,36	11,67	1.193	270	0,31	150,67
UAL	0,28	8,97	919	-190	0,23	111,00
UDL	0,26	8,25	845	-132	0,30	121,00
UPNA	0,25	8,17	836	65	0,16	67,50
UJA	0,25	7,99	817	42	0,33	146,67
UL	0,18	5,83	597	-13	0,21	95,33
UCA	0,17	5,64	577	98	0,10	50,00
UPCT	0,17	5,55	567	479	0,32	170,33
UMH	0,14	4,53	463	19	0,34	138,83
UPO	0,11	3,60	368	88	0,09	41,33
UBU	0,08	2,61	268	27	0,05	20,83

Tabla 3. Valores medios brutos y de representatividad de la producción (Scopus) y publicación web (Scholar) por universidad (enero-junio 2009). Algunas variables se han redondeado para mayor claridad.

Según la última edición del ranking web (enero-2008), la UM y EHU son las universidades con mayor tamaño en Scholar (sin contar a la UR y UCM), logrando las posiciones 71 y 94 a escala mundial respectivamente. Estos datos concuerdan con los resultados obtenidos, teniendo en cuenta la caída de la UM en marzo.

La figura 4 muestra por su parte la evolución de las 10 primeras universidades (excluyendo a la UR y UCM) en Scholar recent, destacando de nuevo el mes

de marzo como momento de cambios, con crecimientos muy notables para la UM, UGR, UDG y UPM.

De nuevo UM (en primera posición), EHU y UGR son las universidades con más registros. Aun así, sorprende la aparición en los primeros puestos de la UDG y UCO así como la pérdida de puestos de la UA y UVI, que abandonan el top ten en publicación web, al que llega la UPV gracias a su importante incremento en mayo.

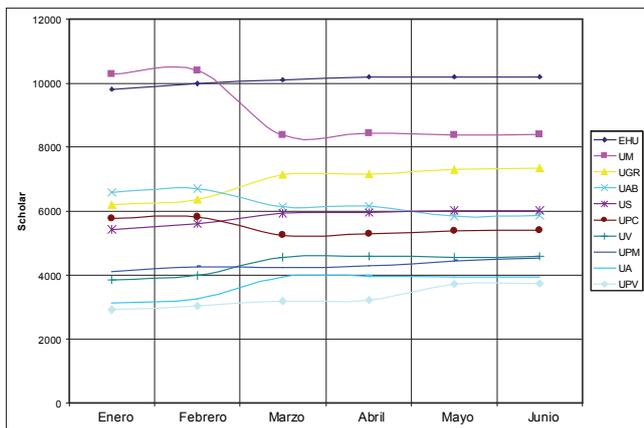


Figura 3. Evolución de los dominios web académicos españoles en Scholar (enero-junio 2009)

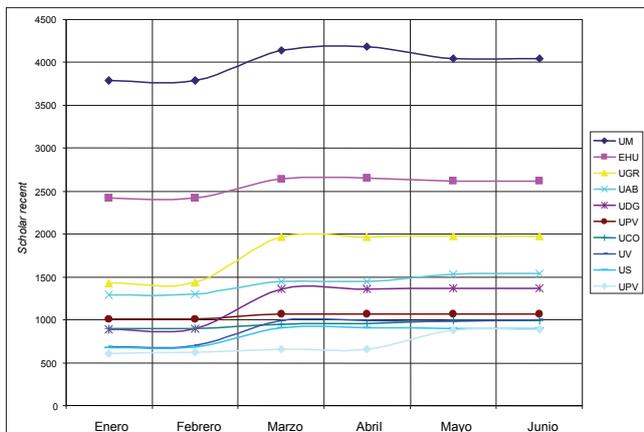


Figura 4. Evolución de los dominios web académicos españoles en Scholar recent (enero-junio 2009)

4. Discusión

Pese a obtener una correlación positiva (*coeficiente de Pearson* de 0,70) entre la productividad en *Scopus* y la publicación web en *Scholar* de las distintas instituciones (exceptuando a la *UR* y *UCM*), se asume la necesidad de una corrección futura por tamaño de universidad, aunque este parámetro no puede ser calculado de forma precisa pues no es proporcionado por todas las universidades, resultando además un término difuso y sin definición clara.

Existen grandes diferencias en los resultados brutos totales, que se asocian a:

- La recuperación tanto de artículos académicos no producidos por la universidad, como documentos no académicos, resúmenes, fichas catalográficas, etc. Muestra de ello es la indización de *Dialnet* o *Compludoc*.

- Diferentes políticas en cada dominio web de universidad en relación con el *Open Access*, que provoca que ciertas universidades queden mal representadas en *Scholar* aun disponiendo de suficiente material académico producido y publicado, como por ejemplo la *USC*⁷,

que pese a ser de las universidades con más producción, todavía no dispone de un repositorio institucional.

“La UR y UCM aparecen las primeras en el ranking de *RScholar*, a una distancia considerable del resto”

- Problemas de *Scholar* en la indización de algunos repositorios institucionales, donde se recopilan metadatos de los proveedores de servicios y no de los repositorios directamente, restando por tanto visibilidad a los dominios web de universidades.

Este es el caso –entre otros– de la *UAM*, que si bien dispone de un repositorio institucional⁸, su contenido no se encuentra disponible dentro del dominio web y, por tanto, no es recogido por la metodología seguida en este trabajo, lo que explica las diferencias entre producción científica y publicación web obtenidas para esta institución.

No es el propósito de este trabajo identificar todas las universidades que se encuentran en esta situación, aunque se asume que la posición de éstas mejoraría si solucionaran este problema, pues se considera una mala práctica⁹ la disposición de documentación académica fuera del dominio web institucional.

5. Conclusiones

El trabajo muestra una comparación entre la publicación web, recuperada por *Scholar* mediante los documentos alojados dentro de los dominios web de cada universidad pública española, y su productividad académica medida desde *Scopus*.

Los resultados demuestran objetivamente que, sin tener en cuenta los dominios web de universidades con grandes servicios de referencia (*Dialnet* y *Compludoc*), la producción académica en *Scopus* y la publicación web en *Scholar* están correlacionadas, aunque son conjuntos documentales de diferente naturaleza.

Pese a la curva positiva de regresión obtenida, la posible interrelación entre productividad en *Scopus* y publicación web en *Scholar* precisa de un análisis más profundo en el futuro, debido a la dependencia de variables externas no controlables como el filtrado de documentos y los métodos de indización de los repositorios institucionales (por parte de *Scholar*) o el tamaño y las políticas de acceso abierto (por parte de las universidades).

Pese a todo, las ventajas de *Scholar* como herramienta de búsqueda de información y como escaparate de atracción de usuarios para las universidades son innegables.

Por ello, tal vez las acciones emprendidas por la UR y UCM disfracen la verdadera actividad en publicación web de dichas universidades en *Scholar* (en todo caso, son carencias en el filtrado de esta herramienta), pero generan una presencia y visibilidad internacional que no logra ninguna otra universidad española.

Por otra parte, sólo 9 de los 46 dominios web universitarios analizados presentan en junio valores más bajos que en enero (recorrido negativo), lo que muestra un crecimiento generalizado, aunque moderado.

Los recorridos negativos (siendo destacable el caso de la UB) pueden deberse a la eliminación de cierta documentación temporal o cambios en las políticas web (como mayor uso de la intranet o de diferentes dominios web, consideradas ambas malas prácticas⁹).

Respecto a *Scholar recent*, los primeros puestos del ranking nacional no presentan muchas variaciones respecto a *Scholar*, aunque sí se observan diferencias en los valores de representatividad relativa. UCM es la que registra un mayor crecimiento y, en menor medida, EHU, UM, UGR, UDG y UCO.

En cuanto a trabajos futuros de investigación, se estima conveniente la comparación de los resultados obtenidos con los correspondientes a las universidades españolas privadas así como con algunas extranjeras, con el fin de contextualizarlos de forma más precisa.

Asimismo, también resultaría de interés el futuro análisis cuantitativo de los diferentes tipos documentales cubiertos en *Scholar* por dominio web de universidad.

6. Agradecimientos

Los autores agradecen enormemente el excelente trabajo realizado por los revisores, que con sus consideraciones y sugerencias han ayudado a mejorar notablemente el texto.

7. Notas

1. *Ranking web de universidades del mundo*
<http://www.webometrics.info/>
2. *Web popularity ranking*
<http://www.4icu.org/top200/>
3. *Ranking universitario de transparencia web*
<http://www.universidad.edu.co>
4. Ranking de universidades españolas por calidad investigadora, publicado por *Scimago* a partir de los datos extraídos de Scopus.
<http://www.magisnet.com/pdf/RankingEspanyolUniversidades.pdf>
5. *Conferencia de Rectores de las Universidades Españolas (CRUE)*
<http://www.crue.org/>
6. No se dispone de los datos relativos a enero debido a errores en las *queries* de consulta y en ningún caso hay problemas en dichos dominios para la extracción de información.
7. <http://www.accesoabierto.net/repositorios>
8. <http://digitool-uam.greendata.es/R>
9. http://www.webometrics.info/best_practices.html

8. Referencias

- Aguillo, Isidro F.** "Hacia un concepto documental de sede web". *El profesional de la información*, 1998, v. 7, n. 1-2, pp. 45-46.
- Aguillo, Isidro F.; Ortega, José-Luis; Fernández, Mario.** "Webometric ranking of world universities: introduction, methodology and future developments". *Higher education in Europe*, 2008, v. 33, n. 2/3, pp. 234-244.
- Aguillo, Isidro F.; Granadino, Begoña; Ortega, José-Luis; Prieto, José-Antonio.** "Scientific research activity and communication measured with cybermetric indicators". *Journal of the American Society for Information Science and Technology*, 2006, v. 57, n. 10, pp. 1296-1302.
- Bar-Ilan, Judit.** "A microscopic link analysis of academic institutions within a country – the case of Israel". *Scientometrics*, 2004, v. 59, n. 3, pp. 391-403.
- Harzing, Anne-Wil K.; Van der Wal, Ron.** "Google Scholar as a new source for citation analysis". *Ethics in science and environmental politics*, 2008, v. 8, n. 1, pp. 61-73.
- Jacsó, Péter.** "Google Scholar revisited". *Online information review*, 2008a, v. 32, n. 1, pp. 102-114.
<http://www.jacso.info/PDFs/jacso-GS-revisited-OIR-2008-32-1.pdf>
- Jacsó, Péter.** "The pros and cons of computing the h-index using Google Scholar". *Online information review*, 2008b, v. 32, n. 3, pp. 437-452.
<http://www.jacso.info/PDFs/jacso-pros-and-cons-of-computing-the-h-index.pdf>
- Kousha, Kayvan; Thelwall, Mike.** "Sources of Google Scholar citations outside the science citation index: a comparison between four science disciplines". *Scientometrics*, 2008, v. 74, n. 2, pp. 273-294.
- Mayr, Philipp; Walter, Anne-Kathrin.** "An exploratory study of Google Scholar". *Online information review*, 2007, v. 31, n. 6, pp. 814-830.
<http://www.ib.hu-berlin.de/~mayr/arbeiten/OIR31-6.pdf>
- Neuhaus, Chris; Neuhaus, Ellen; Asher, Alan; Wrede, Clint.** "The depth and breadth of Google Scholar: an empirical study". *Libraries and the academy*, 2006, v. 6, n. 2, pp. 127-141.
- Noruzi, Alireza.** "Google Scholar: the new generation of citation indexes". *Libri*, 2005, v. 55, n. 4, pp. 170-180.
<http://www.librijournal.org/pdf/2005-4pp170-180.pdf>
- Payne, Nigel; Thelwall, Mike.** "A statistical analysis of UK academic web links". *Cybermetrics*, 2004, v. 8, n. 1.
<http://www.cindoc.csic.es/cybermetrics/articles/v8i1p2.html>
- Smith, A.; Thelwall, Mike.** "Web impact factor for Australasian universities". *Scientometrics*, 2002, v. 54, n. 3, pp. 363-380.
- Torres-Salinas, Daniel; Ruiz-Pérez, Rafael; Delgado-López-Cózar, Emilio.** "Google Scholar: ¿una herramienta para la evaluación de la ciencia?". *Anuario ThinkEPI 2010* [en prensa].
<http://www.thinkepi.net/google-scholar-%C2%BFuna-herramienta-para-la-evaluacion-de-la-ciencia>

Enrique Orduña-Malea, Facultad de Informática, Universidad Politécnica de Valencia (UPV). Camino de Vera s/n, 46022 Valencia
enorma@fiv.upv.es

Jorge Serrano-Cobos, MASmedios, Álvaro de Bazán 10, 46010 Valencia
jorgemasmedios.com
jorgeserrano@gmail.com

Nuria Lloret-Romero, Universidad Politécnica de Valencia (UPV), Área de Biblioteca y Documentación Científica, Camino de Vera, s/n, 46022 Valencia
nlloret@upvnet.upv.es