

Análisis de la cobertura de la base de datos Scopus

Por Grupo Scimago

A FINALES DE 2004, *ELSEVIER* lanzó al mercado un producto que sin lugar a dudas ha llamado la atención en el mundo de la información científica: *Scopus*¹. Se trata de una gran base de datos multidisciplinar, la mayor en la actualidad y muchos son sus puntos destacables. Un trabajo de Codina de principios del año pasado², nos adelantaba algunas de sus funcionalidades relativas a la recuperación de información. Sin embargo creemos que lo que está animando un cierto debate en los medios profesionales es que además es una base de datos de citas, por lo que el monopolio de facto del *ISI Web of Science (ISI-WoS)* de Thomson Scientific³, puede estar llegando a su fin.

Además del de Codina, y a pesar del poco tiempo que lleva este producto en el mercado, ya existen varios trabajos que intentan caracterizar y analizar *Scopus*. Las primeras noticias datan de marzo del año 2004^{4,5 y 6}, pero es a mediados de 2005 cuando aparece el primer trabajo comparativo con *ISI-WoS*⁷. En este estudio, de Deis y Goodman, *Scopus* (entonces con alrededor de 14.000 títulos de revistas) no queda tan bien parado frente a *ISI-WoS* (con casi 7.000), aunque hay que decir que el método de análisis, especialmente para la cobertura temática, no parece demasiado objetivo y sistemático, por lo que deviene en

polémica. Cheryl La Guardia realiza su propia revisión y termina llegando a conclusiones contrarias a las de Deis y Goodman⁸. Éstos, a su vez, contestan con una carta dirigida al editor de *The Charleston advisor* criticando la forma de realizar las consultas de La Guardia⁹. Por su parte, en una carta enviada por las mismas fechas al mismo editor por Sweeney y Singh, indican que el estudio de Deis y Goodman hace llegar al lector a conclusiones erróneas que no se desprenden de los datos¹⁰.

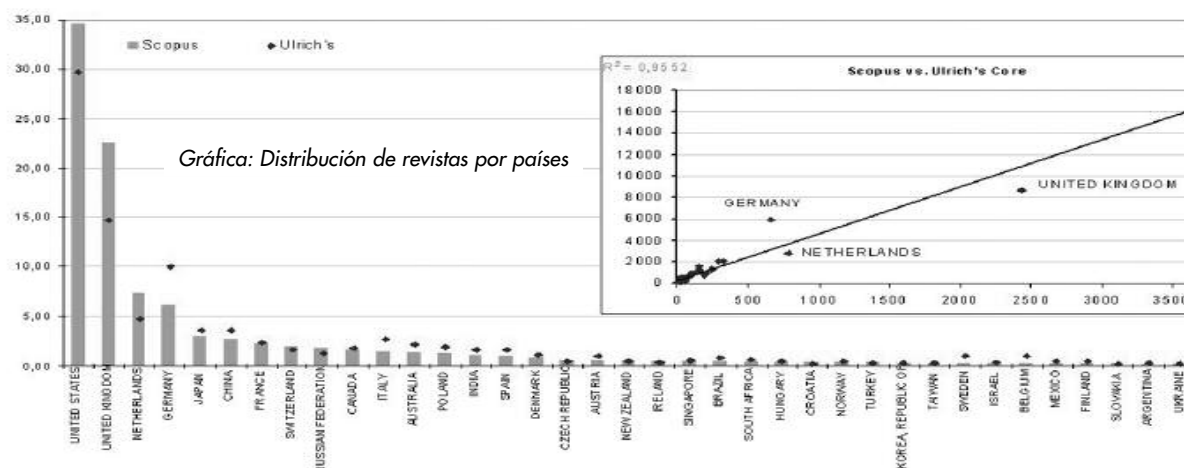
Al margen de esta polémica una cosa queda clara: *Scopus* e *ISI-WoS* han entrado en franca competencia y que cualquier comparación entre ambos productos debe ser realizada de forma reflexiva, cuidadosa y metodológicamente coherente. Es más, creemos que la comparación directa entre ambos productos es en sí misma problemática, por lo que para analizar la cobertura es necesario hacerlo con relación a algún modelo indiscutible y aceptado. Esto es lo que hicieron Braun, Glänzel y Schubert¹¹, comparando la cobertura del *ISI-WoS* con relación al *Ulrich's international periodicals directory*¹², universalmente aceptado como la base de datos que contiene la casi totalidad de las revistas editadas en el mundo. De todas formas esta herramienta es aún muy grande, debido a que incluye una gran cantidad de revistas que

no son estrictamente científicas. Por esta razón nosotros hemos optado por utilizar el núcleo de revistas científicas *peer-reviewed*: el *Ulrich's Core*, que agrupa un poco más de 28.000 títulos.

En este trabajo analizaremos la cobertura de *Scopus* en relación con *Ulrich's Core* tomando como modelo la metodología de Braun, Glänzel y Schubert, que tiene en cuenta cuatro ámbitos de cobertura: temas, países, idiomas y editores. Debido a las limitaciones de espacio no podemos ofrecer datos para los cuatro, por lo que presentaremos una gráfica de distribución por países y comentaremos brevemente el resto.

Resultados

En la parte principal de la gráfica se observan dos distribuciones: una de barras y otra de puntos, que corresponden a *Scopus* y *Ulrich's Core* respectivamente. Como se puede apreciar, existe una correlación bastante alta entre ambas que puede ser expresada como $R^2=0,95$ (ver a la derecha). En esta segunda gráfica se representan los países como nubes de puntos en función de la cantidad de revistas que tienen en cada base de datos (*Ulrich's Core* en el eje vertical y *Scopus* en el horizontal). Los países que aparecen por debajo de la línea de tendencia se encuentran sobre-representados en *Scopus* (United Kingdom y Net-



herlands), y por encima los que están sub-representados (Germany). El resto de países se ubican en las vecindades de la línea de tendencia. Por esta razón la correlación es alta y sus diferencias no son significativas.

En cuanto a la distribución por temas, la correlación entre *Scopus* y *Ulrich's Core* es también alta ($R^2=0,83$), aunque algo menor que en el caso de los países. Las diferencias en la distribución se concentran generalmente en los campos temáticos de humanidades y ciencias sociales (historia, educación, derecho, ciencias políticas, arqueología, etc.). Aquí el problema radica en la propia constitución de las disciplinas, que en muchos casos no se encuentran consolidadas desde un punto de vista del proceso formal de comunicación científica. Tan es así, que si las eliminamos y nos quedamos con las biomédicas, ciencias básicas y tecnologías, la correlación es $R^2=0,99$, más que satisfactoria.

La distribución por idiomas tiene poca discusión, ya que más del 80% de la producción científica está en inglés (en *Scopus* llega al 83%). Las diferencias en la distribución del resto de lenguas son importantes en términos relativos, pero bajas en términos absolutos ya que se reparten menos del 20% del total. La correlación total es de $R^2=0,90$, y las principales diferencias se centran en el francés, chino,

holandés y ruso (idiomas sobre-representados), y en el alemán, español e italiano (idiomas sub-representados).

Por último, y en lo referente a los editores, las distribuciones son bastantes similares ($R^2=0,96$). Aquí los problemas se presentaron en la recolección de los datos debido a que el mundo de las editoriales científicas está en constante movimiento. Las revistas y editores sufren de manera cotidiana cambios y fusiones; no obstante, los grandes nombres se encuentran consolidados y bien representados en ambas base de datos.

Podemos concluir que la cobertura de *Scopus* es equilibrada, estadísticamente hablando, en términos de temas, países, idiomas y editores, basándonos en el *Ulrich's Core*. Es importante indicar que **Braun**, **Glänzel** y **Schubert** llegaron a una conclusión similar en su estudio del *ISI-WoS*. Por ello, no sería descabellado indicar que estamos ante dos productos que representan a escala la composición de la ciencia mundial, particularmente *Scopus* cuyo tamaño es mayor que el del *ISI-WoS*. En números redondos *Scopus* representa el 50% del universo de revistas *peer-reviewed* (*Ulrich's Core*), mientras que *ISI-WoS* sólo el 25%.

Notas

1. <http://www.scopus.com>
2. **Codina, Lluís**. "Scopus: el mayor navegador científico de la web". En: *El profesional de la*

información, 2005, enero-febrero, v. 14, n. 1, pp. 44-49.

3. <http://www.isiknowledge.com>

4. "Scopus to challenge Web of science?". En: *Access: Asia's newspaper on electronic information products & services*, 2004, March, n. 48. <http://www.aardvarknet.info/access/number48/monthnews.cfm?monthnews=08>

5. **Hane, Paula**. "Elsevier announces Scopus service". En: *Information today*, 2004, March. <http://www.infotoday.com/newsbreaks/nb040315-1.shtml>

6. **Pickering, Bobby**. "Elsevier prepares Scopus to rival ISI Web of science". En: *Information world review*, 8 March 2004.

7. **Deis, Louise F.; Goodman, David**. "Web of science (2004 version) and Scopus". En: *The Charleston advisor*, 2005, January, v. 6, n. 3.

8. **La Guardia, Cheryl**. "E-views and reviews: Scopus vs. Web of science". En: *Library journal*. <http://www.libraryjournal.com/article/CA491154.html>

9. **Goodman, David; Deis, Louise F.** "Scopus vs WoS: additional review in LJ". En: *The Charleston advisor*, 2005. http://www.charlestonco.com/feedback/view_letter.cfm?id=35

10. **Sweeney, Richard T.; Singh, Haymwantee**. *Letter to Editor re: Web of science (2004 version) and Scopus*. En: *The Charleston advisor*, 2005. http://www.charlestonco.com/feedback/view_letter.cfm?id=34

11. **Braun, Tibor; Glänzel, Wolfgang; Schubert, Andras**. "How balanced is the Science Citation Index's journal coverage? A preliminary overview of macrolevel statistical data". En: **Cronin, Blaise; Barsky Atkins, Helen** (eds.). *The Web of knowledge, a festschrift in honor of Eugene Garfield*. Canada: American Society of Information Science, 2000, pp. 251-277.

12. <http://www.ulrichsweb.com/>

Informe enviado a EPI por el Grupo SCImago (Imago scientiae o visualización de la ciencia).

scimago@ugr.es

<http://www.atlasofscience.net>