

Almudena Sanz Perla



Reseña literaria sobre...

«La revolución Google Scholar: destapando la caja de Pandora académica»

**E. Orduña-Malea, A. Martín-Martín,
J. M. Ayllón y E. Delgado López-Cózar**

Prologado por P. Jacsó y A.-W. Harzing

Granada: UNE, 2016, 268 págs. ISBN: 978-84-338-5941-9

La aparición de internet en los años noventa supuso una revolución en el acceso a la información. Dos décadas más tarde nace Google Scholar, lo que va a suponer la democratización del acceso a la información académica, de manera que dicha información podrá llegar hasta aquellos con menos recursos.

El presente libro está estructurado en dos bloques. En el primero de ellos se aborda la génesis de Google Scholar, analizándose sus características como motor de búsqueda y como herramienta de evaluación científica. En el segundo bloque los autores se centran en los productos derivados que han sido creados, propios y ajenos a la empresa Google, y que permiten extraer datos bibliométricos de la plataforma Google Scholar. Cabe destacar que la obra cuenta con un par de prólogos que han sido escritos por dos de las personas que más han estudiado la plataforma desde su creación en el año 2004.

En el primer capítulo se narran los orígenes de la creación de Google Scholar, proyecto al que

se dedicaron únicamente dos personas durante un año sabático. Una década más tarde solo conformaban la plantilla 9 trabajadores que se dedicaban a mantener la herramienta, debido a que no producía beneficios a la compañía. Su objetivo era mejorar la indización y el *ranking* de los documentos académicos empleando el algoritmo del motor de búsqueda Google para la asignación de citas («votos»). Un detalle que marcó el arranque del buscador académico fue que, apenas unas semanas antes de su lanzamiento, la empresa Elsevier presentó su plataforma Scopus.

El segundo capítulo analiza las dos caras de la herramienta: por un lado, su perfil de buscador, y por otro, su aplicación como herramienta de evaluación. Paralelamente se hace un repaso de las reacciones de los diferentes investigadores que han analizado Google Scholar, y su impacto en los principales sectores de aplicación. Así, en el ámbito de las bibliotecas se produjeron tres fases: curiosidad, estudio sistemático crítico y aceptación;

en el académico, las reacciones variaron según el tipo de usuario, y dentro de la bibliometría se atendió al estudio como herramienta de evaluación científica enfocada a la democratización.

A lo largo del libro se citan multitud de estudios que se han realizado sobre el recurso. Por ejemplo, atendiendo a los estudios de usuarios, se siguen dos líneas: por un lado, los centrados en la calidad y en las prestaciones de la plataforma y, por otro, los que realizan una comparativa con otros recursos del sector. De este modo, se observa que la gran mayoría de los estudiantes, en torno al 90 %, realizan primero sus consultas en Google antes que en la biblioteca, lo que podría conllevar un problema de superficialidad intelectual causado por la escasa capacidad para evaluar la información por parte del usuario. También, en el ámbito de los investigadores, los motores de búsqueda han comenzado a «sustituir» a las bases de datos tradicionales. Son los recursos más utilizados en todos los campos, salvo en el de la medicina.

En los inicios de Google Scholar los profesionales de la información tenían cierto miedo a que los usuarios se acercaran a la plataforma y dejaran de usar la biblioteca, pero se ha observado cómo, poco a poco, ha aumentado su presencia en las bibliotecas y cómo mejora el uso de los recursos de información de la misma. De todas formas, el colectivo lamenta la poca participación de profesionales de la información en el desarrollo del producto.

Respecto a la comparación con otros productos bibliotecarios de similares características se concluye que Google Scholar es un recurso que complementa las bases de datos bibliográficas de pago, pero que no puede sustituirlas por diversas razones, entre las que destacan el control de la calidad y el tratamiento de los resultados. Es, por tanto, una herramienta alternativa que cuenta con una tecnología muy sofisticada, pero que no permite conocer la amplitud del material que contiene.

El tercer capítulo se centra en su aplicación como herramienta de evaluación de la producción científica. En los inicios se detectaron muchos errores que impedían darle valor como herramienta bibliométrica, pero, con el paso de los años, el margen de error se ha considerado asumible y ha permitido conocer información bibliométrica en disciplinas en las que de otro modo no sería posible; es el caso, principalmente, de las ciencias sociales y de las humanidades.

En este ámbito, se han realizado muchos estudios comparativos centrados en las bases de datos Web of Science (WOF), Scopus y Google Scholar que, además del número de citas, analizan el grado de correlación entre los datos obtenidos de los tres recursos. Esta comparativa ha llevado a que los investigadores se planteen si se deben tener en la misma consideración las citas obtenidas con Google Scholar que las que se muestran en WOF y Scopus debido a que no cuentan con el mismo control de indexación.

A lo largo del cuarto capítulo se explica el proceso de captura de la información de la web académica, detallando los requisitos de indización exigidos por la plataforma para la inclusión del contenido. Se recalca que es imprescindible que los documentos cuenten con un resumen y se incluyen una serie de consejos que habría que tener en cuenta en las páginas de universidades y otros organismos de investigación para facilitar el acceso de los robots al contenido alojado en las mismas y así mejorar la visibilidad de los resultados de investigación a través de Google Scholar.

El quinto capítulo se centra en las características de cobertura, tamaño y evolución de la plataforma, destacando la política opaca de Google sobre la cobertura real de la herramienta. Si bien en los estudios empíricos realizados cabe apreciar, en un principio, algunas debilidades, con el paso de los años, estas se han reducido notablemente. En esta línea, se mencionan también las colecciones

independientes «Disposiciones legales» y «Patentes», disponibles solo en la versión en inglés.

En cuanto al tamaño, los creadores defienden que no debe ser influyente para los usuarios finales, pero los profesionales de la información, por el contrario, creen que debe conocerse, ya que es una característica importante de cualquier plataforma de búsqueda de información. Por los estudios realizados, se estima que a fecha de 2015 podía contener 99,8 millones de documentos, por lo que se habla de *big data* académico, debido a su cobertura de fuentes documentales, lugares geográficos e idiomas. Es la mayor base de datos académica del mundo, superando en crecimiento a la WOS y a Scopus.

Los dos siguientes capítulos se dedican al estudio del proceso de búsqueda de información y al análisis de las fortalezas y debilidades de Google Scholar. En este sentido, se detallan las funcionalidades de búsqueda que ofrece el sistema, tanto en la modalidad de búsqueda simple como en la avanzada, así como los servicios de valor añadido que nos ofrece, desde las alertas de novedades a la exportación de referencias bibliográficas.

El segundo bloque comprende los últimos capítulos del libro y se basa en el análisis de los productos derivados. En primer lugar, destaca el análisis sobre Google Scholar Metrics, creado en 2012 para medir el impacto de las revistas científicas. Al igual que sucede con Google Scholar, al principio los estudios demostraron grandes limitaciones que fueron resueltas con rapidez, convirtiéndose muy pronto en una herramienta de apoyo a la hora de evaluar el impacto.

Otro de los productos que hay que destacar es la aplicación que permite la creación de perfiles académicos, gracias a la cual los autores pueden detallar toda su producción académica y hacerla visible marcando como público su perfil. Sin embargo, aunque la manipulación de los perfiles y de los indi-

adores bibliométricos es relativamente sencilla, los creadores siguen defendiendo la permisividad, confiando en la honestidad de la comunidad científica.

Por último, se analizan aquellos productos derivados que han sido creados por personas independientes a Google. Estos productos nacen con un doble objetivo: estudiar la plataforma Google Scholar y diseñar productos de utilidad para la comunidad. Los primeros en aparecer han sido Publish or Perish y Scholarometer, que ofrecen servicios de valor añadido a las búsquedas de Google Scholar.

El grupo de investigación EC3 ha diseñado y elaborado una serie de herramientas bibliométricas que analizan el impacto de la producción científica de ciencias sociales y humanidades de las universidades públicas españolas. Es el caso de H Index Scholar, Publishers Scholar Metrics y Journal Scholar Metrics. También han creado Proceedings Scholar Metrics, que recopila todas las conferencias relativas a los campos de la informática, de la electricidad y electrónica y de las telecomunicaciones. Por otro lado, han realizado el estudio de la disciplina de la biblioteconomía y documentación española, representando la producción científica en un perfil público en el que se han creado cinco *rankings* de los más citados: autores, documentos, revistas, editoriales e instituciones.

Como conclusión, la obra nos representa el mundo de la producción científica como un iceberg en el que la punta visible representa los productos de pago y la parte oculta todo lo que está por descubrir. Se trata, por tanto, de un libro muy completo que permite al lector conocer en profundidad una herramienta de gran importancia en el mundo científico y académico.

Almudena Sanz Perla

*Técnica de biblioteca de la
Universidad a Distancia de Madrid (UDIMA)*