

Google Scholar Metrics revisado: Ahora empieza a ir en serio

Emilio Delgado-López-Cózar y Álvaro Cabezas-Clavijo

edelgado@ugr.es, acabezasclavijo@gmail.com

EC3: Evaluación de la Ciencia y de la Comunicación Científica,
Universidad de Granada

Delgado López-Cózar, Emilio; Cabezas-Clavijo, Álvaro (2012).
Google Scholar Metrics revisado: Ahora empieza a ir en serio.
EC3 Working Papers, 8: 16 de noviembre de 2012.

RESUMEN

En esta nota urgente se realiza un primer análisis de la nueva versión de Google Scholar Metrics (noviembre de 2012). Se detectan mejoras muy significativas como la posibilidad de consultar las publicaciones más destacadas según áreas y disciplinas, aunque únicamente para las editadas en inglés. Asimismo se ha actualizado la recopilación de citas hasta el 15 de noviembre de 2012, lo que genera la modificación de los indicadores de impacto de las revistas, posibilitando cambios en los rankings así como la entrada de nuevas publicaciones a los mismos. Igualmente se detecta un cambio de criterio en el tratamiento de los documentos de repositorios. Pese a las mejoras realizadas, Google Scholar Metrics continúa siendo, debido a las limitaciones metodológicas, un producto inadecuado de cara a la evaluación de revistas científicas.

PALABRAS CLAVE: Google Scholar, Google Scholar Metrics, Índice H, Repositorios, Revistas Científicas, Evaluación Científica.



INTRODUCCIÓN

En lo que empieza a ser ya la tónica de esta empresa, Google ha decidido seguir sorprendiéndonos con el lanzamiento el 15 de noviembre de la actualización de Google Scholar Metrics. Parece que Google se ha empeñado en seguir siendo diferente, y si las empresas dedicadas a la evaluación bibliométrica de revistas (Thomson Reuters y Scopus) actualizan sus productos anualmente, Google lo ha hecho apenas 7 meses y medio después del lanzamiento de su producto.

Hemos de confesar que esperábamos el movimiento de la compañía californiana. Para solaz de bibliómetras e intranquilidad de la competencia, la actualización y revisión del producto de Google es un paso natural dentro de la expansión de la familia Scholar. Ahora parece que deja de ser un mero divertimento [1], para empezar a ir en serio. Decimos empezar, porque son algunas las mejoras introducidas pero muchas aún las limitaciones de que adolece. Las críticas que formulamos en nuestra serie de documentos parecen que han hecho algo de mella [1-6] ya que se han corregido una buena parte de los problemas mencionados en los diversos análisis, siendo los más llamativo la inclusión de rankings por áreas y disciplinas, y un importante cambio de criterio respecto al tratamiento de los documentos de repositorios.

En esta nota *urgente* repasamos las características más significativas de la versión revisada (noviembre 2012) de Google Scholar Metrics, subrayando las novedades y las debilidades detectadas en un primer análisis.

DESCRIPCIÓN

Scholar Metrics sigue ofreciendo en esta versión la clasificación de las primeras 100 revistas del mundo por idioma de publicación, ordenadas según el índice h y la mediana del número de citas obtenida por los artículos que contribuyen al índice h. El período sobre el que se computa el indicador se mantiene en los últimos cinco años completos (2007-2011), solo que ahora el cálculo de las citas se extiende hasta la fecha del 15 de noviembre, cuando en la anterior versión se quedaba en el 1 de abril.

¿Qué se mantiene?

- Solo se incluyen revistas que hayan publicado al menos 100 artículos en el periodo 2007-2011 y que hayan recibido alguna cita (esto es, se excluyen las revistas con índice $h=0$).
- Los rankings de revistas se presentan por idiomas (siguen siendo diez: inglés, chino, portugués, alemán, español, francés, coreano, japonés, holandés e italiano). Por cada idioma se presenta un listado de las 100 revistas con mayor índice h.

- De cada revista solo se visualizan los artículos que contribuyen al índice h (pudiéndose consultar los documentos citantes (pinchando en *cited by*)
- Asimismo se pueden buscar revistas por palabras incluidas en el título usando el buscador. En este caso, solo se ofrecen 20 resultados.

¿Cuáles son las novedades?

La gran novedad es sin duda la presentación de las revistas por ocho áreas temáticas (*Humanities, Literature & Arts, Social Sciences,...*) y 313 disciplinas (*Religion, Language & Linguistics, History, Algebra,...*). Lamentablemente esta opción solo existe para las revistas en inglés y sólo se muestran las 20 revistas con mayor índice h en cada una de las categorizaciones realizadas.

Las revistas pueden estar clasificadas en varias áreas o disciplinas temáticas, si bien lo habitual es que cada revista se clasifique únicamente en un área y en una disciplina. Cuando se pincha en su índice h puede saberse en qué disciplinas figura (siempre y cuando esté entre las veinte primeras), y cuál es su posición.

En la página de ayuda, se indica con mayor nitidez cuales son las fuentes cubiertas por Google Scholar Metrics (artículos de revistas procedentes de las webs que cumplen los criterios de inclusión de Google, algunos artículos de congresos en Informática e Ingeniería Eléctrica, así como los preprints alojados en repositorios) y las no cubiertas (documentos judiciales, tesis, libros y patentes). Ahora bien la novedad la gran novedad a este respecto es que el índice h de los repositorios se calcula ahora no sobre el total de documentos almacenados (que implicaba una enorme distorsión [6], sino por series o colecciones. Así, dejan de aparecer *Arxiv, RePec* y *SSRN* en su conjunto, y en su lugar se pueden encontrar las colecciones de *Arxiv* (tabla 1) o series de working papers alojados en *SSRN* o *RePec* (por ejemplo *World Bank Policy Research Working Paper Series*, o *OECD Economics Department Working Papers*).

Tabla 1: Índice h y mediana de citas de las 20 colecciones con mayor impacto del repositorio Arxiv.

COLECCIÓN	ÍNDICE H	MEDIANA DE CITAS
arXiv Astrophysics (astro-ph)	176	245
arXiv High Energy Physics - Theory (hep-th)	137	180
arXiv High Energy Physics - Phenomenology (hep-ph)	135	182
arXiv Mesoscale and Nanoscale Physics (cond-mat.mes-hall)	132	193
arXiv Materials Science (cond-mat.mtrl-sci)	132	191
arXiv Quantum Physics (quant-ph)	126	181
arXiv Superconductivity (cond-mat.supr-con)	125	202
arXiv Cosmology and Extragalactic Astrophysics (astro-ph.CO)	107	140
arXiv Information Theory (cs.IT)	88	140
arXiv Other Condensed Matter (cond-mat.other)	85	127
arXiv Strongly Correlated Electrons (cond-mat.str-el)	84	108
arXiv General Relativity and Quantum Cosmology (gr-qc)	80	109

arXiv High Energy Physics - Experiment (hep-ex)	78	113
arXiv Optics (physics.optics)	77	99
arXiv Physics and Society (physics.soc-ph)	73	117
arXiv Nuclear Theory (nucl-th)	72	109
arXiv Statistical Mechanics (cond-mat.stat-mech)	72	107
arXiv High Energy Astrophysical Phenomena (astro-ph.HE)	70	91
arXiv Galaxy Astrophysics (astro-ph.GA)	67	91
arXiv Solar and Stellar Astrophysics (astro-ph.SR)	67	86

A modo de inventario las novedades introducidas son las siguientes:

- Rankings de las 20 publicaciones en inglés con mayor índice H consultable según 8 áreas del conocimiento y 313 disciplinas
- Actualización de los indicadores de impacto para publicaciones en todos los idiomas, alcanzando el marco temporal de las citas hasta el 15 de noviembre de 2012 (marco temporal anterior: hasta el 1 de abril de 2012)
- Cambios en los diferentes rankings por idiomas, y entrada de nuevas revistas a las distintas clasificaciones, dada la ampliación de la ventana de citación.
- Cambio de criterio respecto a los repositorios, que ahora se incluyen segmentados por colecciones o series.
- Mejora de la página que lista los trabajos incluidos en el índice h de cada revista. Así, junto al índice h y la mediana de citas, aparece la posición de la revista dentro de cada área y/o disciplina, siempre y cuando ésta aparezca entre las 20 primeras de su categoría.

LIMITACIONES

Las principales limitaciones en esta nueva versión de Google Scholar Metrics se refieren a la categorización realizada por la empresa, ya que no se explicita ni el criterio seguido para conformar las áreas y disciplinas ni el criterio para la clasificación de las revistas en un área o disciplina. Del mismo modo, es también criticable el hecho de que no se haya realizado esta división para las revistas en idiomas diferentes del inglés.

Como ejemplo, la publicación emblemática de la Biblioteconomía y Documentación, *Journal of the American Society for Information Science and Technology*, no aparece en la disciplina *Library & Information Science* como cabría esperar, sino en *Databases & Information Systems*. Lo mismo ocurre con otra revista tradicional de la disciplina como *Information Processing & Management*.

Tampoco se explicita en cuántas áreas o disciplinas se incluye cada revista. Así, aunque la mayoría de publicaciones parece estar incluida únicamente en un área y en una disciplina, revistas como *PLoS One* están categorizadas en dos áreas diferentes, mientras otras como *Nature Materials* figura en un área temática (*Engineering & Computer Science*) y en tres disciplinas (*Materials Engineering*, *Nanotechnology*, y *Chemical & Material Sciences*).

Asimismo, las revistas de Economía (no así las de Finanzas) están clasificadas tanto en el área *Business, Economics & Management* como en el área *Social Sciences*.

Igualmente, se constata la imposibilidad de acceder a la versión previa de la clasificación, por lo que no se pueden realizar análisis cronológicos sobre la evolución en el impacto de las publicaciones. Esta circunstancia, unida al desconocimiento sobre las fechas de actualización del producto (¿una vez al año, dos veces al año, cada siete meses y medio?) y al cambio en la ventana de citación respecto a la primera versión (que no parece que se vaya a mantener estable en el tiempo) provoca inconsistencia en los indicadores generados y confusión a la hora de interpretar los resultados obtenidos.

Por último, en cuanto al control bibliográfico de la información en Scholar Metrics, parece haberse avanzado en algo: se ha mejorado la normalización en los títulos de revistas (eliminación de títulos abreviados) pero no se han eliminado totalmente los títulos duplicados y, por otra parte, se mantienen muchos de los errores denunciados en la búsqueda y localización de revistas

De manera sumaria señalamos las principales limitaciones.

- Imposibilidad de acceder a la versión anterior de Google Scholar Metrics, con lo que no se puede analizar la evolución de la revista por marcos temporales
- Poca consistencia en los marcos temporales escogidos para el cálculo de los indicadores. No se justifica por qué se amplía el marco temporal para la citación mientras que los artículos analizados continúan ciñéndose al periodo 2007-2011.
- Cambio de criterio respecto a los repositorios, que altera de manera importante los resultados presentados en la primera versión del producto.
- No se especifica en cuántas series o colecciones se han dividido los distintos repositorios.
- Desglose por áreas y disciplinas únicamente para publicaciones en inglés. Para el resto de idiomas se han actualizado los indicadores pero sigue siendo imposible acceder a las revistas más destacadas por categorías.
- Visualización únicamente de las 20 primeras revistas de cada área y disciplina. No es posible conocer pues la posición el resto de publicaciones.
- Inexistencia de información acerca de los criterios tenidos en cuenta para la categorización efectuada de las revistas en las diferentes áreas y disciplinas.
- Inexistencia de información acerca del número de disciplinas en que se ha clasificado cada una de las revistas.
- Inexactitud en la clasificación de revistas dentro de áreas y disciplinas.

CONCLUSIONES

Como reza en el título de esta apresurada revisión de la nueva versión de Google Scholar Metrics, parece que Google empieza a tomarse en serio la evaluación de revistas científicas a través de los recuentos de citas, y son notorias algunas mejoras que se están realizando. Además de la corrección de algunos errores de bulto denunciados en anteriores trabajos, ya es

apreciable el concurso humano (no sólo de robots) en la definición de áreas, disciplinas y en la categorización de revistas. Sin embargo, y reconociendo los progresos efectuados, y la buena línea que apunta en el futuro, Scholar Metrics sigue distando de ser una herramienta fiable y válida debido a las múltiples limitaciones de índole metodológico señaladas. A nuestro pesar tenemos que reproducir lo que ya habíamos escrito anteriormente [5]: estamos ante “un producto inmaduro, que consta de múltiples limitaciones en su configuración actual para la evaluación de las revistas científicas, por lo que no se aconseja por el momento su uso de en ningún proceso de evaluación”.

Éstas son unas primeras valoraciones de urgencia de la actualización de Google Scholar Metrics. En próximos días ofreceremos un análisis más detallado del producto y de las consecuencias que se derivan para la evaluación investigadora.

FINANCIACIÓN

Este trabajo ha sido financiado con cargo al proyecto HAR2011-30383-C02-02 de la Dirección General de Investigación y Gestión del Plan Nacional de I+D+I, Ministerio de Economía y Competitividad.

MATERIAL SUPLEMENTARIO

Índice de áreas y disciplinas incluidas en Google Scholar Metrics

REFERENCIAS BIBLIOGRÁFICAS

1. Cabezas-Clavijo, Á., y Delgado-López-Cózar, E. (2012). Scholar Metrics: el impacto de las revistas según Google, ¿un divertimento o un producto científico aceptable? *EC3 Working Papers*, (1).
2. Cabezas-Clavijo, Á., y Delgado-López-Cózar, E. (2012). Las revistas españolas de Ciencias Sociales y Jurídicas en Google Scholar Metrics, ¿están todas las que son? *EC3 Working Papers*, (2).
3. Cabezas-Clavijo, Á., y Delgado-López-Cózar, E. (2012). ¿Es posible usar Google Scholar para evaluar a las revistas científicas nacionales en los ámbitos de Ciencias Sociales y Jurídicas? El caso de las revistas españolas. *EC3 Working Papers*, (3).
4. Delgado-López-Cózar, E., Robinson-García, N., y Torres-Salinas, D. (2012). Manipular Google Scholar Citations y Google Scholar Metrics: simple, sencillo y tentador. *EC3 Working Papers*, (6).
5. Delgado-López-Cózar, E., y Cabezas-Clavijo, Á. (2012). Google Scholar Metrics: an unreliable tool for assessing scientific journals. *El Profesional de la Información*, 21(4), 419–427. doi:10.3145/epi.2012.jul.15

6. Delgado-López-Cózar, E., y Robinson-García, N. (2012). Repositories in Google Scholar Metrics or what is this document type doing in a place as such? *Cybermetrics*, 16(1), paper 4.