

Agregación de índices bibliométricos para evaluar la producción científica de los investigadores

Por Francisco Herrera, Enrique Herrera-Viedma, Sergio Alonso y Francisco-Javier Cabrerizo



Francisco Herrera es catedrático del departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada. Sus líneas de investigación abarcan la computación con palabras y toma de decisiones, minería de datos, sistemas difusos, algoritmos evolutivos y bibliometría. Es miembro del comité editorial de numerosas revistas y ha publicado más de 150 artículos científicos. Presenta un índice $h = 38$ según WoS.



Enrique Herrera-Viedma es catedrático de Ciencias de la Computación e Inteligencia Artificial y Vicedecano de Investigación en la facultad de Comunicación y Documentación de la Univ. de Granada. Sus líneas de investigación incluyen la computación con palabras y toma de decisiones, recuperación de información web y bibliometría. Ha publicado más de 70 artículos científicos y presenta un índice $h = 23$ según WoS.



Sergio Alonso es profesor ayudante del departamento de Lenguajes y Sistemas Informáticos de la Universidad de Granada. Ha publicado diversos artículos científicos en revistas y en congresos internacionales y nacionales. Sus líneas de investigación incluyen la toma de decisión en grupo, modelos de consenso, modelado lingüístico difuso, bibliotecas digitales, evaluación de la calidad web y bibliometría.



Francisco-Javier Cabrerizo es profesor ayudante del departamento de Ingeniería del Software y Sistemas Informáticos de la Universidad Nacional de Educación a Distancia. Ha publicado diversos artículos científicos en revistas y en congresos. Sus líneas de investigación incluyen la toma de decisión en grupo, modelos de consenso, bibliotecas digitales, evaluación de la calidad web y bibliometría.

Resumen: Se muestra la importancia de la agregación de índices bibliométricos como método para evaluar la calidad científica de los investigadores, ya que permite crear nuevos índices mucho más significativos que combinan las ventajas de los índices agregados.

Palabras clave: Agregación de índices, Evaluación de investigadores, Índice h , Índice g , Índice hg .

Title: Aggregation of bibliometric indices to evaluate the scientific production of researchers

Abstract: In this paper we show that the aggregation of bibliometric indices is an appropriate way to evaluate scientific production of researchers because it allows us to develop much more precise new bibliometric indices that combine the advantages of each index that is aggregated.

Keywords: Aggregation of indices, Evaluation of researchers, h -index, g -index, hg -index.

Herrera, Francisco; Herrera-Viedma, Enrique; Alonso, Sergio; Cabrerizo, Francisco-Javier. "Agregación de índices bibliométricos para evaluar la producción científica de los investigadores". *El profesional de la información*, 2009, septiembre-octubre, v. 18, n. 5, pp. 559-561

DOI: 10.3145/epi.2009.sep.11

Introducción

LA EVALUACIÓN DE LA PRODUCCIÓN CIENTÍFICA de los investigadores es una tarea

fundamental para decidir sobre la aprobación de proyectos de investigación, contratación o promoción de investigadores, etc.

El uso de índices bibliométricos ha atraído una creciente atención, debido a los beneficios de obtener un criterio de evaluación que se re-

suma en un único valor numérico (Grupo SCImago, 2006).

Existen diferentes índices que permiten medir tanto la producción científica de los investigadores como el impacto de la misma. El índice h de Hirsch (2005) ha alcanzado una elevada popularidad y ha sido aplicado a gran variedad de áreas. Su idea es muy simple: un investigador tiene un índice h si h de sus N_p artículos tienen al menos h citas cada uno, y los otros ($N_p - h$) artículos no tienen más de h citas cada uno.

Para intentar superar sus limitaciones, numerosos autores han propuesto novedosos índices basados en éste. Uno de los más interesantes es el índice g (Egghe, 2006): un investigador tiene un índice g si g es el rango más alto tal que sus g trabajos más citados tienen juntos al menos g^2 citas. Esto también significa que sus $(g + 1)$ trabajos más citados tienen menos de $(g + 1)^2$ citas.

Una lista completa de publicaciones relacionadas con el índice h , sus variantes y material adicional sobre este tema se puede encontrar en Egghe (2009), Alonso; Cabrerizo; Herrera-Viedma; Herrera (2009b) y en este sitio web:

<http://sci2s.ugr.es/hindex>

Cada índice identifica aspectos diferentes de la producción científica de un investigador, con lo que la información que proporcionan se complementa muy bien (Borrmann; Mutz; Daniel, 2008). Siguiendo la idea de los metabuscadores de internet, que combinan los resultados de buscadores para enriquecer las respuestas a los usuarios, algunos autores están investigando la combinación de índices bibliométricos mediante operadores de agregación adecuados para crear metaíndices bibliométricos. Estos combinan las ventajas de los índices agregados permitiendo dar una visión más certera de la calidad científica de los investigadores. Al-

gunos ejemplos de metaíndices son: hg -index (Alonso; Cabrerizo; Herrera-Viedma; Herrera, 2009a), q^2 -index (Cabrerizo; Alonso; Herrera-Viedma; Herrera, 2009), y R -index (Jin; Liang; Rousseau; Egghe, 2007).

En este trabajo se muestra cómo la agregación de índices bibliométricos que miden distintos aspectos de la producción científica de un investigador puede proporcionar una evaluación más rica de la misma que si se usa exclusivamente un único índice. Para ello se presenta un ejemplo basado en el metaíndice hg -index.

El metaíndice hg

Permite combinar los diferentes aspectos evaluados por los índices h y g :

- el índice h representa el núcleo de publicaciones (h -core) más citadas del investigador;
- el índice g combina la cantidad de artículos más citados de un autor con la intensidad de citas de los mismos.

El hg -index de un investigador se calcula como la media geométrica de sus índices h y g :

$$hg = \sqrt{h \cdot g}$$

Atendiendo a las propiedades de la media geométrica y de los pro-

prios índices h y g , el índice hg nos proporciona generalmente un valor más cercano a h que a g , incrementa la granularidad en la clasificación, y nos da una perspectiva equilibrada entre el impacto de la mayoría de los mejores trabajos de un investigador y los más altamente citados.

Caso de estudio: bibliometría

En esta sección analizamos el comportamiento del índice hg con respecto a los índices h y g en un caso de estudio en el ámbito de la bibliometría. Para ello empleamos un ejemplo (Egghe, 2006) donde se comparan algunos investigadores del campo de la bibliometría usando los índices h y g y añadimos el índice hg (tabla 1).

En la tabla 1 podemos comprobar que cada uno de los índices mide aspectos diferentes de la producción científica de los investigadores. El índice h se centra en localizar el núcleo de artículos más citados. El índice g permite identificar y valorar los investigadores que, a pesar de tener un núcleo de publicaciones relativamente pequeño, son muy citadas. Por último, el índice hg pondera ambos aspectos de la producción de cada autor obteniendo una valoración más equilibrada de sus publicaciones, y proporcionando mayor granularidad en la clasificación de investigadores que los índices h y g por separado.

	Índice h		Índice g		Índice hg
Garfield	27	Garfield	59	Garfield	39,91
Narin	27	Narin	40	Narin	32,86
Braun	25	Small	39	Braun	30,82
Van Raan	19	Braun	38	Small	26,50
Glänzel	18	Schubert	30	Schubert	23,24
Moed	18	Glänzel	27	Van Raan	22,65
Schubert	18	Martin	27	Glänzel	22,05
Small	18	Moed	27	Moed	22,05
Martin	16	Van Raan	27	Martin	20,78
Egghe	13	Ingwersen	26	Ingwersen	18,38
Ingwersen	13	White	25	White	17,32
Leydesdorff	13	Egghe	19	Egghe	15,72
Rousseau	13	Leydesdorff	19	Leydesdorff	15,72
White	12	Rousseau	15	Rousseau	13,96

Tabla 1. Investigadores del área de bibliometría ordenados según sus índices h , g y hg

“La agregación de índices bibliométricos que miden distintos aspectos de la producción científica de un investigador proporciona una evaluación más rica que un único índice”

Si prestamos atención al índice *h*, comprobamos que **Van Raan** (índice *h* = 19) está por delante de **Small** (18) en la clasificación. Por el contrario, el índice *g* de este último (39) es bastante mayor que el del anterior (27). Sería más lógico que la evaluación de la producción científica de **Small** fuera mayor que la de **Van Raan**, ya que tienen un índice *h* similar, pero el impacto de los trabajos de **Small** es bastante mayor.

Si sólo tenemos en cuenta el índice *g*, observamos que **Small** (39) está por delante de **Braun** (38) en la clasificación. Sin embargo, observamos que el índice *h* de **Braun** (25) es bastante mayor que el de **Small** (18). Sería lógico que la evaluación de la producción científica de **Braun** fuera mayor que la de **Small**, ya que tienen un índice *g* similar, pero el índice *h* de **Braun** es bastante mayor que el de **Small**.

El problema es que los índices *h* y *g* evalúan dimensiones diferentes de la producción científica de los investigadores. Sin embargo, el índice *hg* distingue mejor entre los investigadores ya que tiene en cuenta la información proporcionada por ambos.

La afirmación de que los índices *h* y *g* evalúan dimensiones diferentes de la producción científica de un investigador se puede demostrar por la baja correlación existente entre ellos. Para medirla, como no está claro si los valores de ambos siguen una distribución normal, hemos calculado los coeficientes de correlación de *Spearman* ρ (tabla 2).

ρ	Índice <i>h</i>	Índice <i>g</i>	Índice <i>hg</i>
Índice <i>h</i>	1,000	0,881	0,922
Índice <i>g</i>	-	1,000	0,985
Índice <i>hg</i>	-	-	1,000

Tabla 2. Coeficientes de correlación de *Spearman*

La baja correlación entre los índices *h* y *g* se da porque el índice *h* puede ser bajo y el *g* muy alto debido a que unos pocos artículos hayan recibido muchas citas. Sin embargo, la correlación entre los índices *hg* y *h* y entre los índices *hg* y *g* es bastante alta.

“El índice *hg* es homogéneo y robusto, tiene un buen comportamiento y permite obtener una buena evaluación de la producción científica de los investigadores”

Teniendo en cuenta los resultados obtenidos en el ejemplo estudiado y la alta correlación con los índices *h* y *g*, podemos afirmar que el índice *hg* es homogéneo y robusto, que tiene un buen comportamiento y que permite obtener una buena evaluación de la producción científica de los investigadores. Comprobamos así cómo la agregación de diferentes fuentes de información proporciona mejores resultados que si las utilizamos por separado.

Bibliografía

- Alonso, Sergio; Cabrerizo, Francisco-Javier; Herrera-Viedma, Enrique; Herrera, Francisco. “*hg*-index: a new index to characterize the scientific output of researchers based on the *h*- and *g*-indices”. *Scientometrics*, 2009a, in press. DOI: 10.1007/s11192-009-0047-5.
- Alonso, Sergio; Cabrerizo, Francisco-Javier; Herrera-Viedma, Enrique; Herrera, Francisco. “*h*-index: A review focused in its variants, computation and standardization for different scientific fields”. *Journal of informetrics*, 2009b, in press. v. 3, n. 4, pp. 273-289. DOI: 10.1016/j.joi.2009.04.001.

Bornmann, Lutz; Mutz, Rüdiger; Daniel, Hans-Dieter. “Are there better indices for evaluation purposes than the *h*-index? A comparison of nine different variants of the *h*-index using data from biomedicine”. *Journal of the American Society for Information Science and Technology*, 2008, v. 59, n. 5, pp. 830-837.

Cabrerizo, Francisco-Javier; Alonso, Sergio; Herrera-Viedma, Enrique; Herrera, Francisco. “*q*²-Index: quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core”. *Journal of informetrics*, 2009, in press. DOI: 10.1016/j.joi.2009.06.005.

Egghe, Leo. “Theory and practice of the *g*-index”. *Scientometrics*, 2006, v. 69, n. 1, pp. 131-152.

Egghe, Leo. “The Hirsch-index and related impact measures”. *Annual review of information science and technology*, v. 4, 2009, in press.

Grupo SCImago. “El índice *h* de Hirsch: aportaciones a un debate”. *El profesional de la información*, 2006, v. 15, n. 4, pp. 304-306. <http://www.elprofesionaldeinformacion.com/contenidos/2006/julio/8.pdf>

Grupo SCImago. “El índice *h* de Hirsch: su aplicación a algunos de los científicos españoles más destacados”. *El profesional de la información*, 2007, v. 16, n. 1, pp. 47-49. <http://www.elprofesionaldeinformacion.com/contenidos/2007/enero/05.pdf>

Hirsch, Jorge. “An index to quantify an individual’s research output”. En: *Proceedings of the National Academy of Sciences of the United States of America*, 2005, pp. 16.569-16.572.

Jin, Bihui; Liang, Liming; Rousseau, Ronald; Egghe, Leo. “The *R*- and *AR*- indices: complementing the *h*-index”. *Chinese science bulletin*, 2007, v. 52, n. 6, pp. 855-863.

Rousseau, Ronald. “New developments related to the Hirsch index”. *Science focus*, 2006, v. 1, n. 4, pp. 23-25.

Francisco Herrera, Enrique Herrera-Viedma. Departamento de Ciencias de la Computación e Inteligencia Artificial. ETS de Ingenierías Informática y de Telecomunicación, Universidad de Granada. herrera@decsai.ugr.es viedma@decsai.ugr.es

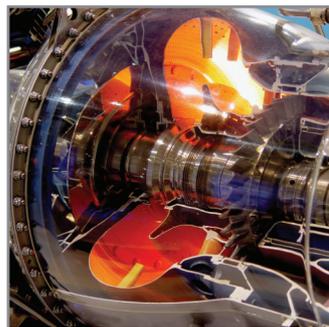
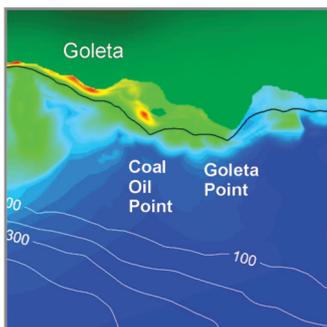
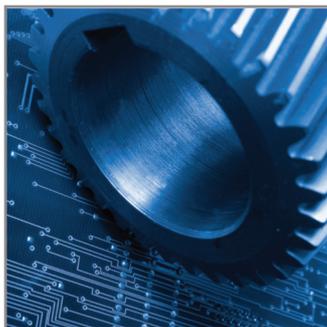
Sergio Alonso. Departamento de Lenguajes y Sistemas Informáticos. ETS de Ingeniería Informática, Universidad de Granada. zerjioi@ugr.es

Francisco-Javier Cabrerizo. Departamento de Ingeniería del Software y Sistemas Informáticos, Universidad Nacional de Educación a Distancia. cabrerizo@issi.uned.es



CSA Illustrata™

Para descubrir información oculta, **empiece aquí...**



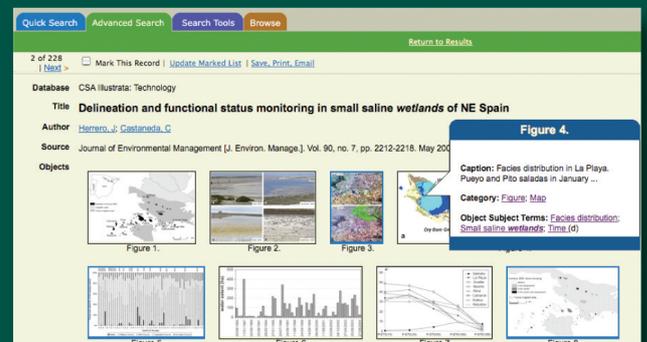
- Indexación completa de las tablas, ilustraciones, imágenes y todo tipo de objetos gráficos.
- Identifica fácilmente artículos difíciles de encontrar.
- Permite al investigador seleccionar los resultados más relevantes gracias a la indexación completa de tablas y gráficos (deep indexing).
- Más precisión, relevancia y eficacia en la recuperación de información.
- Ayuda a la biblioteca a rentabilizar sus inversiones en recursos con texto completo.

CSA Illustrata™: Natural Sciences

Más de 4.700.000 objetos indizados de más de 3.400 revistas revisadas en el área de las Ciencias Naturales.

CSA Illustrata™: Technology

Más de 4.400.000 objetos indizados de más de 2.700 revistas revisadas en el área de la tecnología.



Para más información contacte con la oficina de ProQuest en Madrid:
editor@spain.proquest.com con la referencia **AD 834 09**.

www.proquest.co.uk/go/csaiillustrata