

MODELOS DE EVALUACIÓN DE PRODUCCIÓN CIENTÍFICA PARA EL ÁREA DE CIENCIAS DE LA COMPUTACIÓN

M. Rey¹, H. Kuna¹, E. Martini¹, A. Rambo¹, C. Biale¹, E. Zamudio¹, A. Canteros¹, A. Cantero¹

1. Depto. de Informática, Facultad de Ciencias Exactas Quím. y Naturales Universidad Nacional de Misiones.

{m.rey00, hdkuna}@gmail.com

RESUMEN

La evaluación de la producción científica de un autor o de un grupo de investigación, así como la determinación del impacto generado por una publicación científica han sido temas de interés en las últimas décadas. En la actualidad, se dispone de una gran cantidad de métricas para la valoración del impacto de documentos científicos, sin embargo, los modelos de evaluación utilizados en la generación de tales indicadores presentan diversas limitaciones que afectan su implementación. Entre estos inconvenientes se destaca la falta de adaptabilidad de los modelos para considerar las diferentes características inherentes a un área de conocimiento específica y el carácter unidimensional del análisis llevado a cabo. En este marco, la presente línea de investigación pretende desarrollar un modelo de evaluación de la producción científica que integre características a fin de mitigar los problemas de los métodos actuales y constituir una herramienta de calidad para la determinación del impacto de la producción científica, específicamente en el ámbito de las ciencias de la computación.

Palabras clave: indicadores bibliométricos, producción científica, citas, factores de impacto, evaluación de investigación.

CONTEXTO

Está línea de investigación articula el “Programa de Investigación en Computación” de la Facultad de Ciencias Exactas Químicas y Naturales (FCEQyN) de la Universidad Nacional de Misiones (UNaM) con el Grupo de Investigación Soft Management of Internet and Learning

(SMILE) de la Universidad de Castilla-La Mancha, España.

1 INTRODUCCION

1.1 Modelos de evaluación de producción científica

Los medios a través de los cuales un investigador puede presentar avances de su trabajo en un área de conocimiento abarcan, entre otros: comunicaciones a congresos, artículos en revistas científicas, libros, documentos técnicos [1]. Con fines prácticos en el presente documento se hará referencia a todos ellos como documentos, artículos o producción científica en general. La evaluación de la producción científica se realiza desde sus inicios en base a la cantidad de citas que recibe un artículo [2], [3], a partir de este enfoque de análisis se puede decir que la evaluación de los autores de tales documentos, podría valorar más a un autor cuyos artículos registran más citas que a uno cuyos resultados implican un avance mayor en una disciplina determinada, no siendo ésta una relación directamente proporcional [1]. De igual manera, la evaluación de publicaciones tales como revistas o actas de congresos se clasifican a partir de la cantidad de citas que reciben, en conjunto, los artículos que conforman la publicación [4]. Sin embargo esta metodología, presenta diferentes inconvenientes derivados de los métodos particularmente utilizados para determinar el valor que una métrica o indicador otorga para un objeto de análisis, indiferentemente si se trata de un autor, una revista científica o un documento individual [1], [5]–[7]. En general, las mayores críticas que se realizan sobre los métodos de evaluación se basan en que las citas no necesariamente implican

uso del conocimiento o avances que son presentados en el documento citado [8], [9], por lo tanto no pueden ser consideradas en forma taxativa como una métrica del impacto de la publicación, sino que se podría decir que están más cercanas a representar la visibilidad que ha tenido el documento [10].

Por otra parte, los modelos utilizados en la actualidad más allá de su objetivo de evaluación, tienden a limitar el análisis a una única dimensión [11], [1], minimizando la variedad y calidad de datos utilizados para la determinación del impacto de un artículo científico. Del mismo modo, las diferencias existentes en los esquemas y ciclos de publicación entre determinadas áreas de conocimiento [7], [12], plantea la necesidad de un modelo adaptable y capaz de considerar estas cuestiones al momento de la evaluación. En especial para las características propias del área de ciencias de la computación, ámbito de aplicación específica del modelo objetivo de la presente investigación.

1.2 Métricas disponibles en la actualidad

La diversidad de métricas accesibles ha aumentado considerablemente en los últimos años [13], pasando a constituir un recurso de mayor acceso y difusión con respecto al momento de aparición de las primeras métricas en los años 60 [4].

Esta proliferación ha ocasionado el reversionado de algunos indicadores con el objetivo de solucionar parte de los problemas atribuidos a sus versiones originales. Es así como, por ejemplo, a partir del índice H [2], métrica definida para valorar la producción de un autor determinado, se han definido los índices E [14], G [15] y W [16], que constituyen variantes de la primera. La variación que se presenta en las métricas nuevas se centra en cómo se procesan las citas obtenidas por los documentos publicados de un autor y en el tamaño de la muestra de los mismos que se emplea para calcular el valor final de la métrica para el investigador.

En forma similar, han surgido métricas para la evaluación de la calidad de revistas científicas, en un segmento en el cual el IF [3] se presentaba como principal alternativa, actualmente conviven los indicadores SJR [17], eigenfactor [18], SNIP y RIP [19], entre otros. En este caso las nuevas métricas proponen un cambio de paradigma al trabajar sobre bases de datos de revistas científicas y sus registros de citas, de mayor tamaño con respecto al IF [20]–[22], además de modificar los métodos para el cálculo de sus valores tomando como referencia los empleados en otras áreas, como por ejemplo el algoritmo PageRank de Goole [23] para la evaluación de sitios web, y buscando considerar no solo cantidad sino también calidad de las citas que recibe una determinada revista [10], [18]. Una tendencia incorporada en algunas librerías digitales es la de mantener datos sobre las descargas y/o compras de un determinado documento¹, saliendo de la exclusividad de la cantidad de citas recibidas por el mismo para presentar al investigador una estimación de la calidad del mismo, sumando opciones a métricas anteriormente definidas para este objetivo como el índice AR [24].

La generación de nuevas métricas se ha complementado con el surgimiento de nuevas fuentes de datos², algunas de ellas desarrolladas y mantenidas en forma independiente a una editorial o empresa, además del constante incremento de datos en las existentes previamente.

Finalmente, también se han implementado métricas preexistentes para la evaluación de objetos para los que no habían sido concebidas originalmente, por ejemplo: siendo el índice H un indicador originalmente planteado para la evaluación de la producción de un autor-investigador, con el paso del tiempo se ha utilizado para determinar la calidad de revistas científicas a través del cambio de la base de datos

¹ Por ejemplo: ACM Digital Library e IEEE Xplore Digital Library

² Por ejemplo: MicroSoft Academic Research, Arnetminer.org y CiteSeerX, entre otros.

sobre la cual es calculado el valor del indicador [25].

1.3 Inconvenientes detectados en los modelos disponibles

El crecimiento en lo que respecta a la cantidad de indicadores en el último tiempo no ha sido necesariamente acompañado por una mejora en la calidad de los mismos [13]. Algunos problemas atribuibles a las métricas con mayor “antigüedad” se mantienen vigentes en las de más reciente desarrollo [4], [26], [27]. Entre otros se reconocen: la ausencia de enfoques de análisis que involucren más de una dimensión, como podrían ser los planteados en base al algoritmo PageRank [18], sin limitar su ámbito a la evaluación de una revista o un autor, sino constituyendo herramientas integradoras que consideren al documento desde la unión de las dimensiones de análisis disponibles; por otra parte, las fuentes de datos a partir de las que son calculados los valores de las métricas no se encuentran libres de cuestionamientos en lo que respecta a la inclusión de grupos de publicaciones, tipos de documentos a considerar como “eítibles” y, en caso de que además de los datos de citas provean métricas, los métodos utilizados en el cálculo de las mismas también pueden ser cuestionados [9], [20], [28]; por último y como se mencionó anteriormente, las áreas de conocimiento presentan diferentes métodos, costumbres y ciclos en lo que respecta a sus esquemas de publicación, por lo que los métodos a utilizar para la evaluación de la producción científica deben ser adaptables con el objetivo de no beneficiar o perjudicar a un área en particular, incluyendo en tal adaptación los datos que son tomados en cuenta, los periodos de análisis y los tipos de publicación que son analizados [1], [11].

2 LÍNEAS DE INVESTIGACION, DESARROLLO E INNOVACIÓN

La evaluación de la producción científica es una actividad que ha cobrado gran relevancia con el paso del tiempo, distintos

organismos de financiamiento basan sus decisiones en lo que respecta al otorgamiento de fondos en el análisis bibliométrico de la producción científica. Tanto al evaluar a un investigador en forma individual, como para evaluar centros o institutos de investigación y hasta universidades, estos organismos utilizan conjuntos de indicadores y fuentes de datos sobre citas de los documentos publicados por quien se encuentre en evaluación.

Los modelos actuales para la evaluación de la producción científica, como se ha mencionado, realizan el análisis de los documentos, los autores y los medios de publicación en forma aislada, en ocasiones utilizando métricas con determinadas características que pueden parcializar la evaluación. Es por esto que la generación de un modelo de evaluación que plantee un análisis multidimensional y que sea adaptativo se considera como un objetivo más que válido para constituir una alternativa en la evaluación de producción científica en sus diferentes ámbitos y en especial para el área de ciencias de la computación.

3 RESULTADOS Y OBJETIVOS

3.1 Objetivos de la investigación

Dados los problemas detectados en los modelos de evaluación de la producción científica actuales, la presente línea de investigación se propone la generación de un modelo de evaluación integrador, en el cual sean consideradas las diferentes dimensiones de análisis aplicables a una unidad de producción científica, y se genere a partir de la aplicación del mismo un indicador del impacto generado por la misma. Además se propone que el mencionado modelo sea adaptable, con el fin constituir una herramienta aplicable en las ciencias de la computación, en función de sus características propias. Para ello se evaluará la factibilidad de incorporar diferentes técnicas propias del área de tratamiento de datos masivos con la

premisa de que el indicador generado no se vea sesgado por los datos de una única fuente de información sobre citas.

3.2 Actividades en curso

La presente línea de investigación se encuentra en una fase de inicio de relevamiento bibliográfico orientado a la identificación de un mayor número de opciones en lo que a modelos de evaluación de producción científica se refiere, incluyendo una mayor cantidad de métricas que permitan una mejor comprensión del estado del arte del tema de investigación planteado. Por otra parte, la búsqueda bibliográfica se plantea como una actividad a partir de la que se pasará a contar con mayor información sobre las desventajas planteadas en la sección 1.3 sobre los métodos y modelos actuales de evaluación científica. Especialmente aquellos que refieren a la ausencia de un modelo multidimensional de evaluación que sea parametrizable para su aplicación en distintas áreas de conocimiento.

3.3 Objetivos a corto plazo

A corto plazo los objetivos de la presente investigación son:

- Establecer una taxonomía de las métricas de evaluación de la producción científica relevadas que facilite su caracterización y utilización en distintos escenarios en base a las necesidades que plantee el contexto.
- Determinar las dimensiones de análisis que serán incluidas en el modelo a generar, caracterizarlas y determinar la aplicabilidad de las métricas existentes para la valoración de las mismas.
- Definir los métodos a partir de los cuales será calculado el impacto relativo de cada dimensión de análisis integrada en el modelo de evaluación.

4 FORMACION DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del “Programa de Investigación en Computación” de la FCEQyN de la UNaM, con ocho integrantes (todos ellos alumnos, docentes y egresados de la carrera de Licenciatura en Sistemas de Información de la FCEQyN – UNaM) de los cuales tres están realizando su tesis de grado, dos se encuentran realizando una maestría y dos se encuentran realizando un doctorado. La línea y el equipo de investigación se vinculan con el Grupo de Investigación Soft Management of Internet and Learning (SMILe) de la Universidad de Castilla-La Mancha, España.

5 BIBLIOGRAFIA

- [1] D. A. Pendlebury, «The use and misuse of journal metrics and other citation indicators», *Arch. Immunol. Ther. Exp.*, vol. 57, n.º 1, pp. 1-11, feb. 2009.
- [2] J. E. Hirsch, «An index to quantify an individual's scientific research output», *Proc Natl Acad Sci U S A*, vol. 102, n.º 46, pp. 16569-16572, nov. 2005.
- [3] Garfield E, «The history and meaning of the journal impact factor», *JAMA*, vol. 295, n.º 1, pp. 90-93, ene. 2006.
- [4] H. F. Moed, *Citation Analysis in Research Evaluation*. Springer Science & Business Media, 2006.
- [5] M. Amin y M. Mabe, «Impact factors: use and abuse», *Perspectives in publishing*, vol. 1, n.º 2, pp. 1-6, 2000.
- [6] J. Bollen, H. Van de Sompel, A. Hagberg, y R. Chute, «A Principal Component Analysis of 39 Scientific Impact Measures», *PLoS ONE*, vol. 4, n.º 6, p. e6022, jun. 2009.
- [7] P. O. Seglen, «Why the impact factor of journals should not be used for evaluating research», *Bmj*, vol. 314, n.º 7079, p. 497, 1997.

- [8] D. Torres-Salinas y E. Jiménez-Contreras, «Introducción y estudio comparativo de los nuevos indicadores de citación sobre revistas científicas en Journal Citation Reports y Scopus», *El profesional de la información*, vol. 19, n.º 2, pp. 201–208, 2010.
- [9] R. Adler, J. Ewing, y P. Taylor, «Joint committee on quantitative assessment of research: citation statistics», *Australian Mathematical Society Gazette*, vol. 35, n.º 3, pp. 166–88, 2008.
- [10] J. Bollen, M. A. Rodríguez, y H. Van de Sompel, «Journal status», *Scientometrics*, vol. 69, n.º 3, pp. 669–687, dic. 2006.
- [11] A. Carey, M. G. Cowling, y P. G. Taylor, «Assessing research in the mathematical sciences», *Australian Mathematical Society Gazette*, vol. 34, n.º 2, p. 84, 2007.
- [12] P. Dong, M. Loh, y A. Mondry, «The—impact factor” revisited», *Biomedical digital libraries*, vol. 2, n.º 7, pp. 1–8, 2005.
- [13] J. Bollen, H. Van de Sompel, A. Hagberg, y R. Chute, «A Principal Component Analysis of 39 Scientific Impact Measures», *PLoS ONE*, vol. 4, n.º 6, p. e6022, jun. 2009.
- [14] C.-T. Zhang, «The e-Index, Complementing the h-Index for Excess Citations», *PLoS ONE*, vol. 4, n.º 5, p. e5429, may 2009.
- [15] L. Egghe, «An improvement of the H-index: the G-index», *ISSI newsletter*, vol. 2, n.º 1, pp. 8–9, 2006.
- [16] Q. Wu, «The w-index: A significant improvement of the h-index», *Journal of the American Society for Information Science and Technology*, p. n/a-n/a, 2009.
- [17] B. Gonzalez-Pereira, V. Guerrero-Bote, y F. Moya-Aregon, «The SJR indicator: A new indicator of journals’ scientific prestige», *arXiv:0912.4141*, dic. 2009.
- [18] C. Bergstrom, «Measuring the value and prestige of scholarly journals», *Coll Res Libr News*, vol. 68, n.º 5, p. 3146, 2007.
- [19] H. F. Moed, «Measuring contextual citation impact of scientific journals», *Journal of Informetrics*, vol. 4, n.º 3, pp. 265–277, 2010.
- [20] J. Bar-Ilan, «Which h-index? — A comparison of WoS, Scopus and Google Scholar», *Scientometrics*, vol. 74, n.º 2, pp. 257–271, nov. 2007.
- [21] M. E. Falagas, V. D. Kouranos, R. Arencibia-Jorge, y D. E. Karageorgopoulos, «Comparison of SCImago journal rank indicator with journal impact factor», *FASEB J*, vol. 22, n.º 8, pp. 2623–2628, ene. 2008.
- [22] M. E. Falagas, E. I. Pitsouni, G. A. Malietzis, y G. Pappas, «Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses», *FASEB J*, vol. 22, n.º 2, pp. 338–342, ene. 2008.
- [23] S. Brin y L. Page, «The anatomy of a large-scale hypertextual Web search engine», *Computer Networks and ISDN Systems*, vol. 30, n.º 1–7, pp. 107–117, abr. 1998.
- [24] B. Jin, «The AR-index: complementing the h-index», *ISSI newsletter*, vol. 3, n.º 1, p. 6, 2007.
- [25] Microsoft Academic Search, *Help Center*. 2014.
- [26] R. Monastersky, «The number that’s devouring science», *Chronicle of Higher Education*, vol. 52, n.º 8, p. 14, 2005.
- [27] J. Ewing, «Measuring journals», *NOTICES-AMERICAN MATHEMATICAL SOCIETY*, vol. 53, n.º 9, p. 1049, 2006.
- [28] P. Jacso, «As we may search-Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases», *CURRENT SCIENCE-BANGALORE*, vol. 89, n.º 9, p. 1537, 2005.