

INFORMETRÍA PARA BIBLIOTECARIOS: DESCRIPCIÓN DE SU PAPEL CLAVE EN LOS PROCESOS DE EVALUACIÓN

Informetrics for librarians: Describing their important
role in the evaluation process

Isidro F. Aguillo

Nota: This article can be read in its original English version on:
<http://www.elprofesionaldelainformacion.com/contenidos/2016/ene/01.pdf>



Isidro F. Aguillo es jefe del *Laboratorio de Cibermetría* del *Instituto de Políticas y Bienes Públicos (IPP)* del *Consejo Superior de Investigaciones Científicas (CSIC)*. Es editor de la revista *Cybermetrics*, que fue la primera revista nacida digital del *CSIC*, y de los *Rankings Web de Universidades, Centros de Investigación, Hospitales, y Repositorios*. Ha publicado más de 200 artículos sobre cibermetría, evaluación de la investigación, indicadores web y revistas electrónicas. Es licenciado en biología por la *Universidad Complutense de Madrid* y master en ciencias de la información por la *Universidad Carlos III de Madrid*.

<http://orcid.org/0000-0001-8927-4873>

*Consejo Superior de Investigaciones Científicas, Instituto de Políticas y Bienes Públicos, Laboratorio de Cibermetría
Albasanz, 26-28. 28037 Madrid, España
isidro.aguillo@csic.es*

Resumen

Los bibliotecarios tienen un papel secundario en el proceso de evaluación de las actividades de investigación, por lo general como proveedores de datos en bruto extraídos de fuentes ya pre-seleccionadas. Dada la naturaleza subjetiva de las decisiones de los comités de evaluación, especialmente entre los encargados de seleccionar candidatos a un puesto, hay una necesidad real de que los procedimientos sean imparciales y objetivos, utilizando para ello un grupo de profesionales que pueden garantizar la independencia y la rigurosidad. Se propone un enfoque neutral, amplio, moderno y cuantitativo, guiado por bibliotecarios académicos, que ofrecen la ventaja de su cercanía a los candidatos, y que incluye utilizar el anonimato de las identidades en los informes, la contextualización de los datos (edad académica, género y disciplina), y el uso de valores no-centrales relativos, con indicación de los umbrales y la diversificación de las fuentes, incorporando bases de datos bibliométricas y no bibliométricas.

Palabras clave

Informetría; Bibliotecarios; Evaluación científica; Indicadores relativos.

Abstract

Librarians are playing a secondary role in the process of evaluating research activities, usually as auxiliary providers of raw data extracted from pre-selected sources. Given the subjective nature of the decision committees, there is a strong need for unbiased and objective procedures guaranteed by independent professionals. A neutral, comprehensive, modern, quantitative approach guided by academic librarians is proposed, including closeness to applicants, anonymity of their identities in the reports, contextualization (academic age, gender, and discipline) of data, usage of relative non-central values with indication of thresholds, and incorporation of new bibliometric and non-bibliometric sources.

Keywords

Informetrics; Librarians; Scientific evaluation; Relative indicators.

Aguillo, Isidro F. (2016). "Informetrics for librarians: Describing their important role in the evaluation process". *El profesional de la información*, v. 25, n. 1, pp. 5-10.

<http://dx.doi.org/10.3145/epi.2016.ene.01>

Introducción

Es imposible hacer una reseña breve de la situación actual, los principales problemas y la evolución futura de la disciplina paraguas ahora conocida como Informetría (Abri zah *et al.*, 2014; Bar-Ilan, 2008). Teniendo en cuenta el público profesional de la revista *EPI*, es preferible adoptar aquí un enfoque más práctico, enfocado especialmente al papel de los bibliotecarios académicos en los procesos de evaluación. El objetivo de esta contribución es no sólo describir una serie de pasos y procedimientos, sino intentar reforzar el papel de dichos bibliotecarios en las universidades y centros de investigación, reconociéndolos como actores clave, objetivos y neutrales en la provisión de información confiable, así como proveedores de métricas útiles sobre el rendimiento científico de individuos y grupos (Iribarren-Maestro *et al.*, 2015; González-Fernández-Villavicencio *et al.*, 2015). Además, teniendo en cuenta la naturaleza subjetiva y los indeseables sesgos asociados usualmente a los comités de evaluación (españoles), el autor propone un protocolo destinado a disminuir los abusos, los malos entendidos y las interpretaciones incorrectas de ciertos indicadores, un pecado común a muchos “expertos” que se han preparado con seminarios de 4 horas sobre *WoS / Scopus*. Es importante recordar que todo lo cual ha sido severamente denunciado por la comunidad científica mundial en la *Declaración de San Francisco sobre evaluación de la investigación (Declaración DORA)*: <http://www.ascb.org/dora>

“ El proveedor de datos (el bibliotecario) y el comité evaluador deben ser entidades independientes ”

Este artículo se inspira en una presentación realizada durante la conferencia *ISSI2015* en Estambul que ahora se ha publicado en forma de artículo (Gorraiz; Gumpenberger, 2015), aunque no se trata de un resumen o una revisión de la misma, sino una contribución independiente.

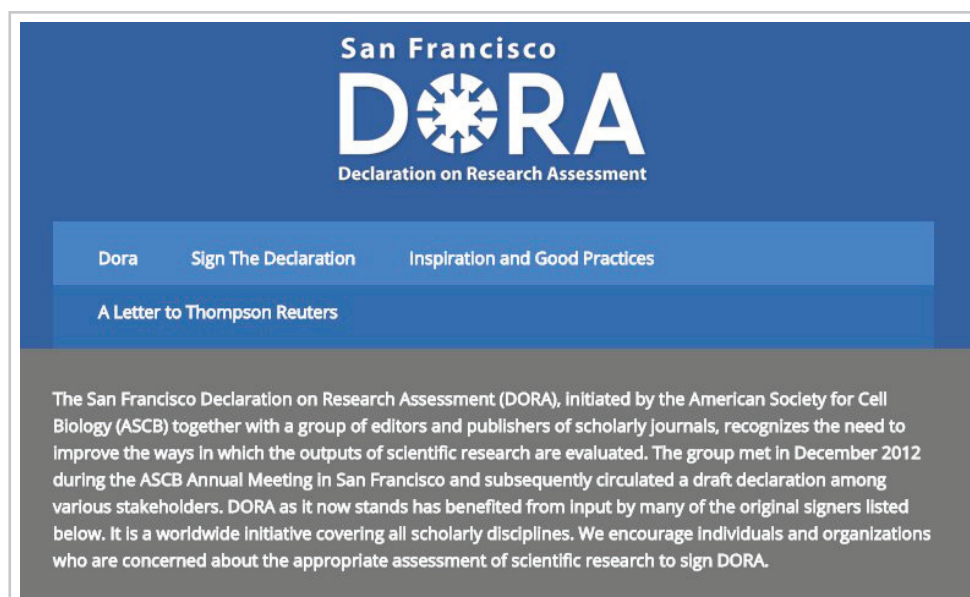


Figura 1. <http://www.ascb.org/dora>

Propuesta de escenario

El proveedor de datos (el bibliotecario) y el comité evaluador deben ser entidades independientes. Los candidatos deben enviar sus CVs y demás información directamente al bibliotecario encargado de recopilar las métricas de fuentes confiables, organizar los datos en indicadores claros y actualizados, y producir el informe cuantitativo que sirva de base para la discusión del comité. Si es necesario el bibliotecario podrá contactar directamente con los candidatos para pedir aclaraciones sobre sus CVs.

“ Es mejor utilizar la edad académica: el número de años desde la obtención del doctorado o de la primera publicación académica ”

Anonimización

La identidad de los candidatos no debe revelarse en el informe final con la métrica. A cada científico debe asignársele un número identificativo. Sin embargo debe incluirse la información personal relevante, sobre todo con respecto a la edad. Como regla general, es mejor utilizar la edad académica: el número de años desde la obtención del doctorado o de la primera publicación académica. La indicación del género no es recomendable, aunque la edad académica debe descontar si los candidatos dieron a luz (un año por niño), sufrieron una enfermedad grave durante más de 6 meses, o ejercieron funciones técnicas o de gestión a tiempo completo.

Recolección de información

La recopilación de datos debe ser lo más completa y actualizada posible en relación a la lista de trabajos incluidos (según conste en el CV de los candidatos) y con el mayor número posible de fuentes reconocidas, ajustando su diversidad según las necesidades del comité. Aunque la Informetría es una disciplina global, la persona que haga la recolección debe evitar combinar todos los resultados en una única tabla, dada

la distinta naturaleza de los datos proporcionados por cada una de las sub-disciplinas cuantitativas. Como no hay un acuerdo universal sobre el diseño y el uso correcto de indicadores compuestos, por ahora es muy desaconsejable el uso de cualquier combinación de indicadores diferentes. Por otra parte, aunque con fines de verificación se necesitan las cifras originales o en bruto, el recolector debe elaborar los datos para producir indicadores de confianza, y explicar los métodos o herramientas utilizados en cada caso. En la colección deben ponerse

las fechas a nivel de día del calendario y a veces también es relevante identificar la ubicación donde se hizo la recogida (los motores de búsqueda con frecuencia geolocalizan los resultados).

Para fines prácticos podemos agrupar los indicadores en cuatro grandes categorías o sub-disciplinas:

- **Bibliometría:** indicadores relacionados con las publicaciones formales en revistas o libros (incluyendo capítulos, ponencias en actas, tesis o similares).
- **Webmetría:** indicadores derivados de la presencia en la Web, como páginas personales o de grupo, portales web, documentos (a texto completo) en repositorios, y otros archivos informáticos (software, audio, vídeo, etc.)
- **Altmétricas:** indicadores recopilados a partir de información académica o relacionada con la investigación, distribuida por las redes sociales, incluyendo blogs (y micro-blogs), wikis, y redes de intercambio, total o parcialmente académicas.
- **Métricas de uso:** aún en una fase muy temprana, estos nuevos indicadores se derivan de los ficheros-log de visitas a los repositorios, portales académicos o sitios web de proyectos científicos.

« Cuando las instrucciones de los evaluadores sean vagas o incompletas deben tenerse en cuenta la experiencia y los criterios de los bibliotecarios »

Una quinta categoría relacionada con la innovación, consiste principalmente en información obtenida de las bases de

datos de patentes. Se puede añadir al informe cuando sea necesario, pero no se analiza en este trabajo.

« El bibliotecario debe elegir con cuidado la variable del ranking (el principal criterio de ordenación) y si es posible incorporar información objetiva adicional »

Elección de las fuentes

Dependiendo del tiempo y de los recursos disponibles, la selección de las fuentes debe ser lo más completa posible. Además de las bases de datos bibliométricas generales (*WoS*, *Scopus* y *Google Scholar*), también deben consultarse las especializadas (como *PubMed*, *Chemical Abstracts*, *CiteSeer...*), y regionales (*Scielo*, *Dialnet*) para los concursos sobre temas específicos. Cuando las instrucciones de los evaluadores sean vagas o incompletas deben tenerse en cuenta la experiencia y los criterios de los bibliotecarios. Son fuentes pertinentes los motores de búsqueda generales (*Google*, *Bing*), pero también los regionales (*Baidu* o *Yandex*) sobre todo cuando no se pueden extraer de forma fiable las métricas de las redes sociales. Entre las fuentes "alternativas", son útiles las redes académicas (*ResearchGate*, *Academia.edu*, *Mendeley*), blogs, microblogs (*Twitter*) y colecciones de archivos (*YouTube*, *Sli-deshare*, *GitHub*). En estos casos las herramientas de integración como *ImpactStory*, *Altmetric* o *Plum Analytics* son una alternativa práctica, pero hay que usarlas con mucho cuidado puesto que a menudo se trata de indicadores compuestos bastante opacos en su construcción.



Figura 2. Modelo del protocolo descrito en el artículo

Elección de las variables

Básicamente podemos identificar dos grandes grupos de variables:

- las que describen las actividades que se van a medir, incluyendo la caracterización de los actores, la clasificación y el volumen de recursos involucrados, y la cantidad de resultados, incluyendo tantos aspectos como sea posible;
- las que describen la audiencia que utiliza los resultados en términos de su visibilidad, la utilidad o el impacto, generalmente considerados como *proxis* o representativos de la calidad de la investigación.

El diseño general de los informes, tablas con variables en las columnas y candidatos en las filas, debería ser útil para fines comparativos. Como regla general, es muy deseable calcular índices a partir de datos obtenidos de ambos grupos de variables (como el número total de citas por el número de documentos en un período), pero nunca mezclar de diferentes fuentes. Si es posible, puede ser útil para el comité informarle de la cobertura mundial de cada base de datos que se está utilizando, preferentemente en términos cuantitativos exactos. Esto permitirá obtener indicadores relativos (productividad, eficacia o eficiencia), una información clave para poder comparar. Cabe recordar, por ejemplo, que la cobertura de *Google Scholar* es varias veces superior a la de *WoS* o *Scopus*.

“ Ahora que se publican cientos de artículos firmados por mil o más autores es muy necesario un nuevo enfoque ”

Normalización

Incluso en pequeñas poblaciones, la distribución de los datos de las variables en bruto puede ser muy sesgada, por lo que es inexacto hacer comparaciones directas entre aspectos o fuentes diferentes. Hay dos soluciones posibles para normalizar los datos:

1. Si se sabe el tamaño total de la población, calcular la relación entre el valor individual y dicho total, que se puede expresar como porcentaje o proporción (partes por uno), teniendo cuidado de no dejar demasiados decimales;
2. Cuando no se conoce el tamaño total, es útil calcular la proporción con respecto al valor máximo observado en la lista de los resultados.

Si la distribución es realmente muy sesgada puede ser necesario transformar los datos utilizando logaritmos (la opción más sólida). El uso de las puntuaciones *z* (*z-scores*) es aconsejable cuando los datos son combinaciones de fuentes de información múltiples y muy heterogéneas.

Umbrales

Una forma neutral de presentar listas de valores en tablas es organizarlas por el identificador de los candidatos, pero



Figura 3. <http://www.coauthorindex.info>

esto puede ser engañoso y oculta información relevante para el evaluador. El bibliotecario debe elegir con cuidado la variable del ranking (el principal criterio de ordenación) y si es posible incorporar información objetiva adicional. Por ejemplo, es sencillo señalar los valores del top 10% o top 5% (se pueden marcar con uno o dos asteriscos), aunque si la distribución es normal (o casi), una buena alternativa es marcar esos valores cuando caen fuera de las dos desviaciones estándar.

Autoría

Un problema muy difícil que ha plagado la bibliometría desde hace décadas, pero que ahora se está convirtiendo ya en una verdadera pesadilla con el fenómeno de las hiper-autorías, es la atribución de autoría, tanto a los autores como a las instituciones. Con el fin de no desalentar la colaboración científica, la solución estándar fue el conteo completo, es decir, asignar el 100% de la autoría de todos los firmantes del documento (y sus instituciones), así como el número total de citas que recibe. Aparte de las cuestiones técnicas, el recuento fraccional no era muy popular, ya que castigó a los equipos de investigación de tamaño mediano o grande. Sin embargo, ahora que se publican cientos de artículos firmados por mil o más autores es muy necesario un nuevo enfoque.

En nuestro contexto, la respuesta es proporcionar información adicional acerca de los patrones de co-autoría de los candidatos. El primer paso es identificar el número de publicaciones en las que el científico involucrado es el autor principal. Por lo general en los trabajos no se declara el papel de cada individuo, pero, dependiendo de cada disciplina, se puede utilizar el número de veces que aparece como primer (o último) autor. Otra posibilidad es comprobar la identidad del autor responsable de la correspondencia (**De-Moya-Anegón, 2012**). Todos estos valores se pueden combinar en un solo número. Sin embargo, este valor puede no ser suficiente en muchos casos, por lo que se sugiere analizar la distribución del número de autores en el conjunto de documentos. Una medida de centralidad es probablemente

suficiente, no el promedio (media), sino la mediana e incluso, para grandes conjuntos, la moda. Hay que tener en cuenta que ahora ya es posible comparar los valores con datos de referencia disponibles según fuente y disciplina:

<http://www.coauthorindex.info>

Como se ha comentado anteriormente, deben marcarse los valores anormales o inesperados o, si es posible, usar un gráfico tipo *box-and-whiskers*.

Indicadores bibliométricos

No vamos a extendernos en esta sección ya que los indicadores bibliométricos son bien conocidos desde hace varias décadas, y son fáciles de entender. El número de publicaciones y el número de citas son medidas estándar de resultados y de visibilidad, respectivamente, y con unas sencillas operaciones puede aumentarse su valor, obteniendo por ejemplo el índice h, aunque hay que insistir que éste es un indicador de resultados, no de impacto. Es muy peligroso presentar una simple tabla con la producción científica y el impacto de los candidatos, puesto que así no se tiene en cuenta la edad (en años de actividad académica). Una solución obvia es proporcionar índices que dividan los valores por el número de años. Por ejemplo, el índice m se define como el índice h dividido por la antigüedad académica en años. Pero esto no es necesariamente la mejor opción, como ocurre, por ejemplo, al solicitar un puesto de trabajo, pues en este caso es más importante conocer el desempeño reciente. Podría ser suficiente saber los documentos publicados en los últimos 5 años (tal vez 3 años para los más jóvenes) o el índice h para ese período.

Es una norma aceptada tener en cuenta los resultados de las bases de datos de calidad fuertemente controlada como *WoS* o *Scopus*, pero los resultados de *Google Scholar (GS)* requieren un tratamiento diferente. *GS* recoge una gran cantidad de trabajos informales, generalmente no presentes en *WoS* o *Scopus*, que con el tiempo se pueden citar (incluso muchas veces). Debido a criterios de calidad equivocados, en general se excluyen las citas de esas publicaciones informales o las de las revistas locales. Teniendo en cuenta que muchos artículos publicados en revistas indizadas no consiguen ser citados, no hay ninguna razón para tal exclusión. Un producto derivado de *GS*, *Google Scholar Citations*, utiliza un indicador donde importa no la revista, sino las citas recibidas por encima de cierto umbral: el número de artículos citados al menos 10 veces (i10). Se podría ir incluso más allá y en ciertas disciplinas o para períodos más cortos utilizar directamente el número de documentos citados al menos una vez.

Impacto o visibilidad en bibliometría se refieren al conteo de citas

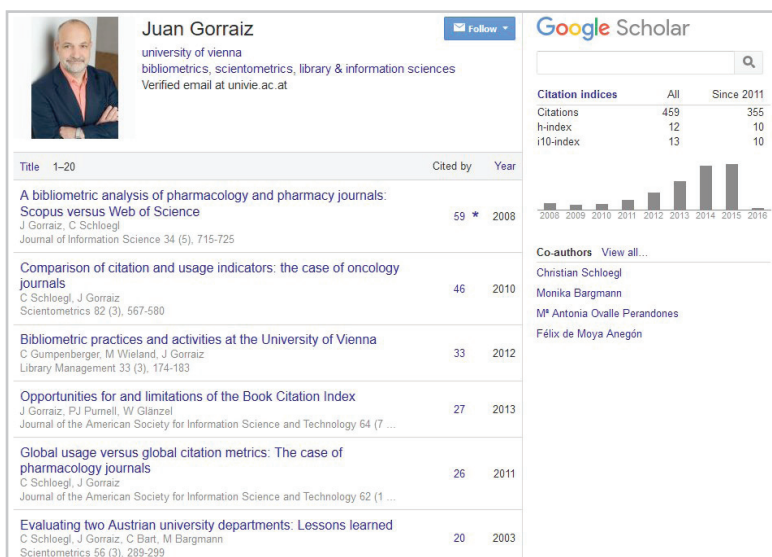


Figura 4. Ejemplo de perfil personal en *Google Scholar Citations* <https://scholar.google.com/citations>

observadas (reales) a un documento concreto, pero sigue siendo muy popular, aunque fuertemente criticado, el uso de las citas esperadas. Este es el caso del *infame* factor de impacto, que es el valor medio de citas recibidas por una revista en los 2 años anteriores y que se toma como el valor de cualquiera de los artículos publicados en ella en el año en curso. Un compromiso frecuente es el uso de cuartiles, que tampoco son indicadores de impacto cierto, pero, como en el caso del índice h, son una medida de resultados. Ordenadas por valores decrecientes del factor de impacto, las revistas de una disciplina se dividen en cuatro grupos, siendo llamado el de arriba primer cuartil. Por lo general, solamente se resaltan de forma explícita los artículos publicados en revistas del primer cuartil.

“ El factor de impacto se basa en la media de las citas recibidas por una revista en los años 2 anteriores, y en cambio se toma como el valor de cualquiera de los artículos publicados en ella en el año en curso ”

Las autocitas también son una fuente de malentendidos: con frecuencia se excluyen en base a que se pueden manipular. Sin embargo, no deberían excluirse de forma automática, pues en la mayoría de los casos muestran desarro-

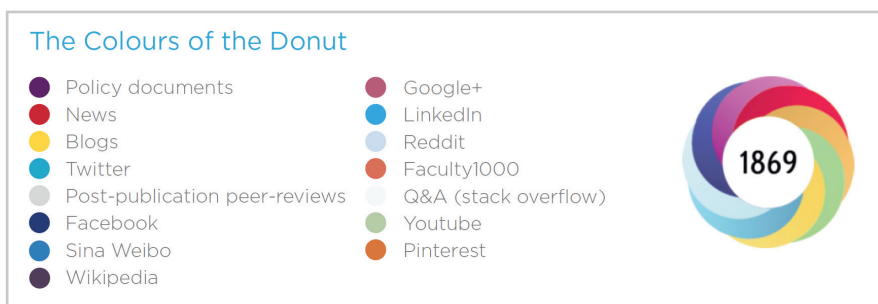


Figura 5. Ejemplo de *Altmetric donut* <http://www.springersource.com/an-introduction-to-altmetric-data-what-can-you-see>

llos anteriores del mismo autor que son relevantes para la comprensión de la actual obra citante. Una propuesta lógica es fijar un porcentaje que, si se supera se debe señalar en las tablas (alrededor del 30%). Evidentemente, las autocitas son más frecuentes en los autores senior, por lo que quizá sea aconsejable fijar un valor más bajo (20%) a los autores con menos de 10 años de actividad editorial. Para señalar un número excesivo de autocitas pueden utilizarse, como en casos anteriores, uno o más asteriscos.

Los indicadores altmétricos son los únicos datos disponibles realmente actuales (recientes y al día)

Altmétricas

Incluso peor que el inapropiado nombre dado a esta subdisciplina es el problema de que conste de muy diferentes y heterogéneas fuentes, variables e indicadores. Las altmétricas son demasiado complejas como para ser tratadas de una manera unitaria (la empresa *Altmetric* proporciona un índice compuesto que en este contexto debe evitarse). Ni los proveedores más interesantes (*Mendeley*, *ResearchGate*) están siquiera cerca de ser una “alternativa” (**Torres-Salinas; Milanés-Guisado**, 2014). Pero hay razones y circunstancias en las que algunos de los indicadores altmétricos pueden ser útiles si se les presenta en una tabla separada, sin mezclar con las otras variables (**Adie**, 2014; **Borrego**, 2014; **Robinson-García et al.**, 2014).

La principal razón para considerar algunos de los indicadores altmétricos es que son los únicos datos disponibles realmente actuales (recientes y al día). Las citas necesitan tiempo, generalmente años, para ser acumuladas (e indexadas) y sin embargo muchos evaluadores solicitan datos incluso del presente año. Los indicadores altmétricos son entonces una opción mejor que los factores de impacto como estimadores de citas futuras. Dada la diferente naturaleza de cada variable y la importancia de las acciones de los autores para promover su trabajo en las redes sociales, los índices compuestos (*ResearchGate score*, *Altmetric donut*) y las variables de auto-promoción (número de items publicados por los autores: tweets, presentaciones, documentos depositados) deben ser descartados en favor de las relacionadas con el impacto: lectores, citas en *ResearchGate*, descargas, retweets, menciones y similares.

Para una revisión extensa de los varios indicadores webmétricos y altmétricos recomendamos los artículos del *Statistical Cybermetrics Research Group* (**Thelwall; Kousha**, 2015a,b; **Kousha; Thelwall**, 2015).

Bibliografía

Abrizah, Abdullah; Erfanmanesh, Mohammadamin; Rohani, Vala-Ali; Thelwall, Mike; Levitt, Jonathan M.; Didegah, Fereshteh (2014). “Sixty-four years of informetrics research: productivity, impact and collaboration”. *Scientometrics*, v. 101, n. 1, pp. 569-585
<http://dx.doi.org/10.1007/s11192-014-1390-8>

Adie, Euan (2014). “Taking the alternative mainstream”. *El profesional de la información*, v. 23, n. 4, pp. 349-351.
<http://dx.doi.org/10.3145/epi.2014.jul.01>

Bar-Ilan, Judit (2008). “Informetrics at the beginning of the 21st century — A review”. *Journal of informetrics*, v. 2, n. 1, pp. 1-52.
<http://dx.doi.org/10.1016/j.joi.2007.11.001>

Borrego, Ángel (2014). “Altmétricas para la evaluación de la investigación y el análisis de necesidades de información”. *El profesional de la información*, v. 23, n. 4, pp. 352-357.
<http://dx.doi.org/10.3145/epi.2014.jul.02>

De-Moya-Anegón, Félix (2012). “Liderazgo y excelencia de la ciencia española”. *El profesional de la información*, v. 21, n. 2, pp. 125-128.
<http://dx.doi.org/10.3145/epi.2012.mar.01>

González-Fernández-Villavicencio, Nieves; Domínguez-Aroca, María-Isabel; Calderón-Rehecho, Antonio; García-Hernández, Pablo (2015). “¿Qué papel juegan los bibliotecarios en las altmetrics?”. *Anales de documentación*, v. 18, n. 2.
<http://dx.doi.org/10.6018/analesdoc.18.2.222641>

Gorraiz, Juan; Gumpenberger, Christian (2015). “A flexible bibliometric approach for the assessment of professorial appointments”. *Scientometrics*, v. 105, n. 3, pp. 1699-1719.
<http://dx.doi.org/10.1007/s11192-015-1703-6>

Iribarren-Maestro, Isabel; Grandal, Teresa; Alecha, María; Nieva, Ana; San-Julían, Teresa (2015). “Apoyando la investigación: nuevos roles en el servicio de bibliotecas de la Universidad de Navarra”. *El profesional de la información*, v. 24, n. 2, pp. 131-137.
<http://dx.doi.org/10.3145/epi.2015.mar.06>

Kousha, Kayvan; Thelwall, Mike (2015). “Web indicators for research evaluation. Part 3: books and non standard outputs”. *El profesional de la información*, v. 24, n. 6, pp. 724-736.
<http://dx.doi.org/10.3145/epi.2015.nov.04>

Robinson-García, Nicolás; Torres-Salinas, Daniel; Zahedi, Zohreh; Costas, Rodrigo (2014). “Nuevos datos, nuevas posibilidades: revelando el interior de *Altmetric.com*”. *El profesional de la información*, v. 23, n. 4, pp. 359-366.
<http://dx.doi.org/10.3145/epi.2014.jul.03>

Thelwall, Mike; Kousha, Kayvan (2015a). “Web indicators for research evaluation. Part 1: Citations and links to academic articles from the Web”. *El profesional de la información*, v. 24, n. 5, pp. 587-606.
<http://dx.doi.org/10.3145/epi.2015.sep.08>

Thelwall, Mike; Kousha, Kayvan (2015b). “Web indicators for research evaluation. Part 2: Social media metrics”. *El profesional de la información*, v. 24, n. 5, pp. 607-620.
<http://dx.doi.org/10.3145/epi.2015.sep.09>

Torres-Salinas, Daniel; Milanés-Guisado, Yusnelkis (2014). “Presencia en redes sociales y altmétricas de los principales autores de la revista *El profesional de la información*”. *El profesional de la información*, v. 23, n. 4, pp. 367-372.
<http://dx.doi.org/10.3145/epi.2014.jul.04>