

INFORMETRICS FOR LIBRARIANS: DESCRIBING THEIR IMPORTANT ROLE IN THE EVALUATION PROCESS

Informetría para bibliotecarios: descripción de su
papel clave en los procesos de evaluación

Isidro F. Aguillo

Nota: Este artículo puede leerse traducido al español en:
http://www.profesionaldelainformacion.com/contenidos/2016/ene/01_esp.pdf



Isidro F. Aguillo is head of the *Cybermetrics Laboratory* at the *Instituto de Políticas y Bienes Públicos (IPP)* from the *Spanish National Research Council (CSIC)*. He is editor of the journal *Cybermetrics*, first born-digital journal from CSIC; and of the *Ranking Web of Universities, Research Centers, Hospitals, and Repositories*. He has published more than 200 papers on cybermetrics, research evaluation, web indicators, and electronic journals. He has a bachelor degree on biology from the *Universidad Complutense de Madrid* and a masters on information science from *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*

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.

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, 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.

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>

Introduction

It is impossible to write a short review of the current situation, major problems, and future developments of the discipline known as Informetrics (Abrizah *et al.*, 2014; Bar-Ilan, 2008). The *El profesional de la información* journal's audience is comprised mainly of library professionals; therefore, a practical approach targeting the academic librarians' role in the evaluation process, is preferred. The aim of this article is not to merely describe a series of steps and procedures, but instead to reinforce the role of academic librarians in universities and research centers as key, neutral, objective actors in providing reliable and useful metrics information about the scientific performance of individuals and groups (Iribarren-Maestro *et al.*, 2015; González-Fernández-Villavicencio *et al.*, 2015). Additionally, taking into account the subjective nature and unwanted biases commonly associated with the decision committees (at least in Spain), the author proposes a protocol intended to decrease the abuse, misunderstandings, and incorrect interpretations of certain indicators; a problem common with many "experts" trained with 4-hours-only *WoS/Scopus* seminar. It is important to remember that all this was denounced by the global scientific community in the *San Francisco Declaration on Research Assessment (DORA declaration)*:

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

“The data provider (librarian) and the decision committee should be independent entities”

This article was inspired by the comments and discussion following a presentation given at the *ISSI2015* conference in Istanbul which was subsequently turned into a published paper (Gorraiz; Gumpenberger, 2015); however, please note that this article is not a summary or a review of that presentation or paper, but instead an independent contribution.

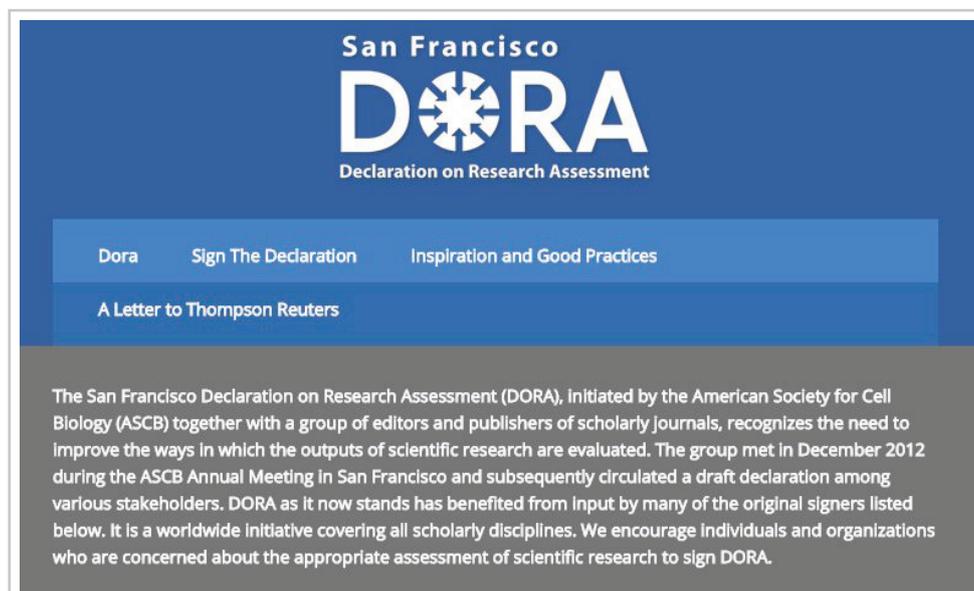


Figure 1. <http://www.ascb.org/dora>

Proposing a scenario

The data provider (librarian) and the decision committee should be independent entities. Candidates should submit their CVs and supporting information directly to the librarian that should be in charge of collecting the metrics from reliable sources, organizing the data into meaningful updated indicators, and producing the quantitative report that will serve as a basis for the discussion of the committee. The librarian would have direct contact with candidates, if needed, for clarification purposes regarding the CV contents.

“It is better to use the academic age: the number of years since the PhD or the first academic publication”

Anonymizing

The identity of the candidates should not be disclosed in the final metrics' report, and an identification number should be assigned to each scientist. However, relevant information should still be included; for example, years in the academic field, years since obtaining a PhD, or years since first publication; the years should be subtracted if the candidate gave birth (one year per child), suffered a serious illness (over 6 months), or he/she had (full-time) technical or managerial duties.

Collecting information

The data collection should be as comprehensive and current as possible regarding both the full list of items (as supplied in the candidates' CV) and the number of recognized sources, promoting diversity according to the committee needs. Although Informetrics is a global term, the collector should avoid combining all the results into one table, given the different nature of the data provided by each one of the quantitative sub-disciplines. Because there is not universal agreement about the design and correct usage of composite indicators, a combination of different indicators is discouraged. Moreover, although raw numbers are needed for checking purposes, this information should be elaborated and expanded upon to produce a more complete picture by providing supporting evidence of the methods, or tools used in each case. Dating the collection should be done at the level of the calendar day and, when appropriate, by identifying the data-collection location (web engines frequently geo-locate their results).

For practical purposes we can group the indicators into four larger categories or sub-disciplines:

- **Bibliometrics:** indicators related to formal publications in journals and books (including chapters, contributions in proceedings, theses, or similar).
- **Webometrics.** Indicators derived from web presence including personal or group pages, web portals, (full-text) documents in repositories, and other computer files (software, audio, video, etc.)
- **Altmetrics.** Indicators collected from academic or research-related information and distributed by social web tools including blogs (and micro-blogs), wikis, and academic (full or part) exchange networks.
- **Usagemetrics.** Still in their early phase these novel indicators are derived from visit log files generated from repositories, academic portals, and scientific projects' websites.

Note: A fifth category related to innovation, mostly consisting of information collected from patent databases, may be added to the scheme as needed, but an analysis is not included in this article.

When the evaluators instructions are vague or incomplete, the experience and criteria of librarians should be taken into account

neral (*WoS, Scopus, and Google Scholar*), specialized (like *PubMed, Chemical Abstracts, CiteSeer*), or regional (*Scielo, Dialnet*) others should be evaluated for a specific call or selection process. General search engines (*Google, Bing*) and regional search engines (*Baidu, Yandex*) are also relevant sources, especially when metrics from social tools cannot be reliably extracted. Among these "alternative" sources, the academic networks (*ResearchGate, Academia.edu, Mendeley*), blogs and microblogs (*Twitter*) and files' collections (*YouTube, Slideshare, GitHub*) are useful. In these cases using integration tools like *ImpactStory, Altmetric, or Plum Analytics* may be a practical alternative, but should be used with extreme care when composite indicators are involved.

When the evaluators instructions are vague or incomplete, the experience and criteria of librarians should be taken into account.

The librarian should choose carefully the rank variable (main ordering criteria) and, if possible, incorporate additional objective information

Choosing the variables

Basically we can identify two large groups of variables. The first set consists of those variables describing the activities to be measured —including the candidates or applicants circumstances, the classification and volume of the resources involved, and the number of results. The second set is intended to describe the audience that uses the search results.

Choosing sources

The selection of sources should be as comprehensive as possible. In addition to bibliometric databases that are ge-

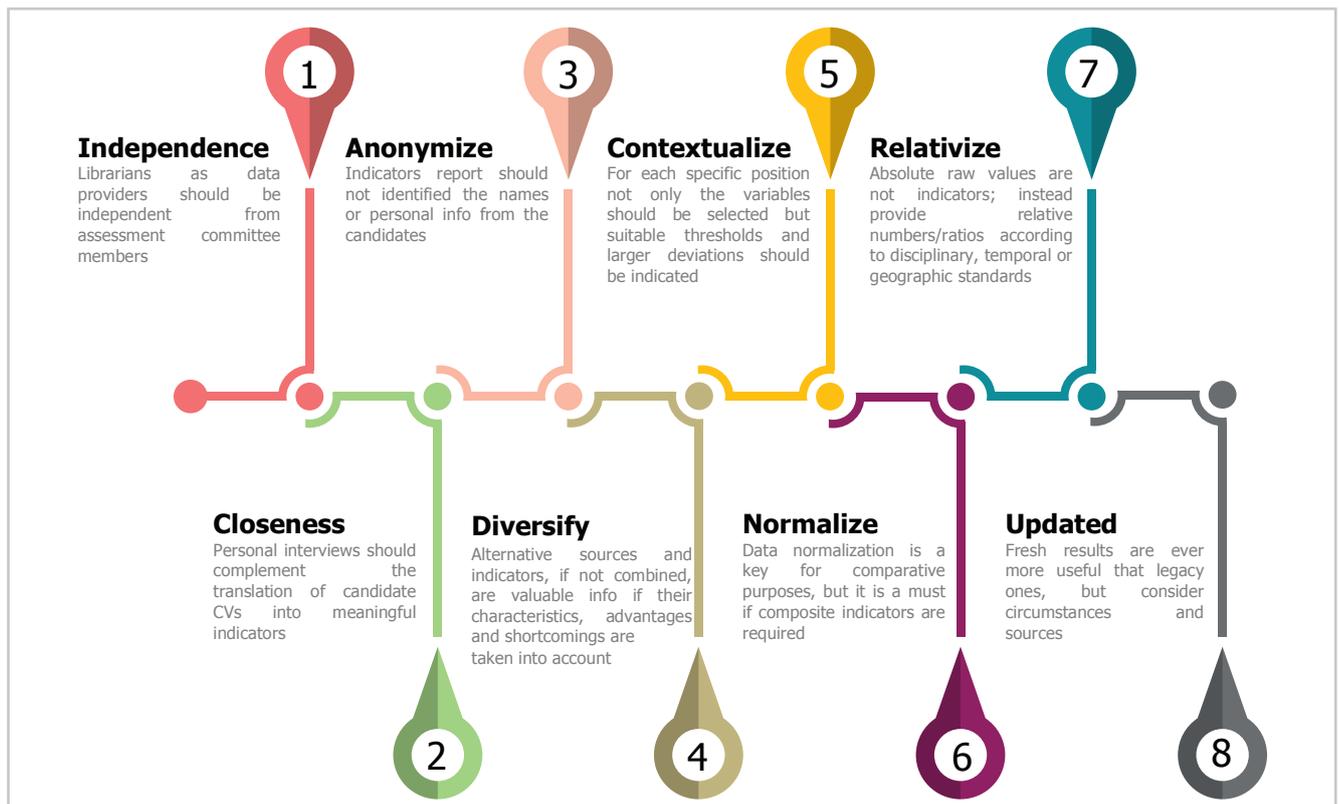


Figure 2. Model of the protocol described in the paper

The overall design of the reports, tables with variables as columns and candidates as entries, should be used for comparative purposes. As a general rule it is best to calculate indexes from data obtained from both groups of variables (ex. total number of citations by the number of papers in a given time period), remembering to never mix sources. If possible, include information about the global coverage of every database that is being used, preferably in exact quantitative terms. From this information relative indicators may be derived (productivity, efficacy, or efficiency), and are important for comparative purposes. It should be recalled, for example, that *Google Scholar* coverage is several times higher than *WoS* or *Scopus*.



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

Normalizing

Even for small populations the distribution of raw data can be highly skewed, thus making unfair comparisons between different aspects or sources. There are two possible solutions for normalizing the data:

1. When the total size of the population is known, to calculate the ratio between the individual value and such total, that can be expressed as percentage or proportion (parts per one), taking care not to include too many decimals.
2. When the total size is not known, the solution consists of calculating the ratio regarding the maximum value observed in the results' list.

If the distribution is really skewed it may be necessary to transform the data using logarithms (the strongest option). The use of z-scores is advisable when the data include combinations of multiple heterogeneous information sources.

“ With hundreds of papers being signed by one thousand or more authors, a new approach is needed ”

Thresholds

A neutral way of presenting the lists of values in tables is arranging them by the candidates' identifier, but this can be misleading and hide relevant information from the evaluator. The librarian should choose carefully the rank variable (main ordering criteria) and, if possible, incorporate additional objective information. For example, statistically it is easy to point out values in the top 5% or 10% (can be marked with one or two asterisks), although if the distribution is (close to) normal, an acceptable alternative is to mark those values when they are outside the two standard deviations.

Authorship

A difficult problem that has plagued bibliometrics for decades, but that is becoming a true nightmare with the pheno-

mena of hyper-authorship, is the attribution of authorship both to authors and institutions. Traditionally, in order to encourage scientific collaboration, authorship was equally shared between all signatories of the paper (and their institutions), as well as the total number of citations the paper receives. However, with hundreds of papers being signed by one thousand or more authors, a new approach is needed.

“ The impact factor is based on citations received by a journal in the previous 2 years, and paradoxically it is taken as the value of any of the articles published in it in the current year ”

In this context, the answer is to provide additional information about the co-authorship patterns of the candidates. The first step is to identify the number of publications in which the scientist is the main author. This number can be useful because the role of each individual in the publication is not usually declared. Another idea is to include the identity of the corresponding author (**De-Moya-Aneón, 2012**). All these values can then be combined in a single figure. However, in most cases, this value is not enough and additional information might include a discussion of the distribution of the number of authors in the set of papers. A centrality measure is probably enough, not the average (mean) but the median and even, for large sets, the mode. Keep in mind that it is now possible to compare the values with reference data available by source and discipline:

<http://www.coauthorindex.info>

As commented earlier, abnormal values should be marked, or if possible, use a box-and-whisker plot graph.

Bibliometric indicators

It is best to keep this section simple because bibliometric indicators are well known and, as a result, easily unders-

tood. The number of publications and citations are standard measures of output and visibility, respectively, although a few manipulations are valid for increasing their descriptive value, such as the h-index that –we must remember- it is an output indicator, not an impact indicator. The greatest danger in presenting a simple table of data regarding scientific output and impact of the candidates is that (academic active) age cannot be taken into account. One solution is providing indexes that divide values by the number of years –the index m is defined as h-index divided by academic age in years. But this is not necessarily the best option, as for example, when filling a new post then the recent performance is specially sought. In this case, papers published in the last 5 years (perhaps 3 years for youngsters) or h-index for that period could be fine enough.

Counting the output in quality controlled databases like WoS or Scopus is an accepted standard, but this is not the case with Google Scholar (GS), which results require a different process. GS collects informally communicated items, usually excluded from the other databases, that may eventually be cited –even highly. Due to the questionable quality criteria, the citations to these informal publications, or those from local journals, are generally excluded; however, considering that many papers published in indexed journals go uncited, there is no reason for such exclusion. Even Google Scholar Citations, which is derived from GS, proposes a threshold for general situations, as it supplies the number of items cited at least 10 times (i10). Beyond that, for certain disciplines or shorter periods, it will be necessary to use the number of papers cited (>0 times).

Impact or visibility in bibliometrics refers to observed (actual) citation counts, and although this method is popular it is also strongly criticized. This is the case with the infamous

impact factor that provides the previous two years’ mean of citations per paper of the journal as a value for any of the papers published in it in the current year. A popular compromise is the use of quartiles, which is not a true impact indicator, but instead, as in the case of the h-index, an output measure. Arranged by decreasing impact factor values, the journals of certain disciplines are divided into four groups, with the top one as the so-called first quartile. Usually only the papers published in a journal ranked in the first quartile are explicitly highlighted.

“ The main reason for considering some of the altmetrics indicators is their currency and relevancy (fresh and updated) ”

Self-citations are also a source of misunderstandings and are frequently excluded because they can be manipula-

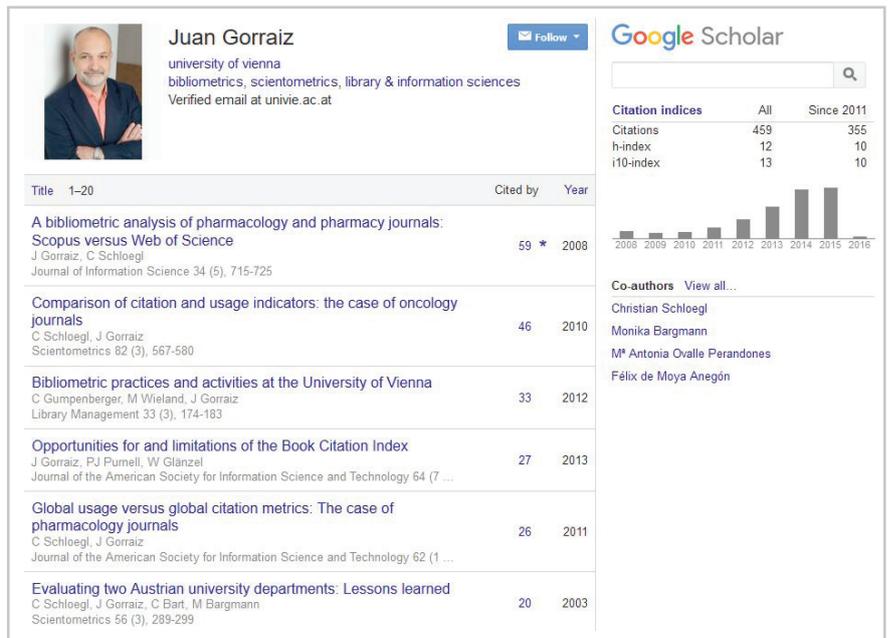


Figure 4. Example of personal profile in Google Scholar Citations <https://scholar.google.com/citations>



Figure 5. Example of Altmetric donut <http://www.springersource.com/an-introduction-to-altmetric-data-what-can-you-see>

ted. But a full exclusion is not warranted in most cases because they show previous developments by the same author/s and are relevant for understanding the current work. A logical proposal is to set a fixed percent (about 30%) that when surpassed would be marked. However, self-citations are obviously more frequent in senior authors, so perhaps it would be advisable to fix a lower value (20%) for authors with less than 10 years of publishing activity. One or more asterisks can be used to indicate excessive self-citation.

Altmetrics

The biggest issue with the current definition of this sub-discipline is that it consists of heterogeneous sources, variables, and indicators. It is too messy to be treated in a uniform way (the company *Altmetric* provides a composite index that should be avoided in this context) and even the unique providers (*Mendeley*, *ResearchGate*) are not close to the definition of “alternative” (**Torres-Salinas; Milanés-Guisado**, 2014). However, there are certain circumstances where a few altmetrics indicators may be useful and in these cases they should be provided in an isolated table and not mixed with other variables (**Adie**, 2014; **Borrego**, 2014; **Robinson-García et al.**, 2014).

The main reason for considering some of the altmetrics indicators is their currency and relevancy (fresh and updated). While citations are generated over time, years usually, evaluators often request data from the current year. The use of altmetrics indicators is then a better option when compared to applying impact factors as estimators of future citations. Given the different nature of each variable and the importance of the active actions of the authors to promote their work in social networks, the composite indexes (*ResearchGate Score*, *Altmetric donut*) and the self-promoted variables (number of items published by the authors: tweets, slides collections, papers deposited) should be discarded in favor of the impact-related ones: readers, *ResearchGate* citations, downloads, retweets, mentions, and similar.

For an extensive review of webometric and altmetrics indicators we strongly recommend papers by the *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-Julián, 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>