Revista Facultad de Ingeniería -redin-, Universidad de Antioquia, No.107, pp. 7-8, Apr-Jun 2023

EDITORIAL

Correlation between article download and citation of scientific publications

The ease of access to information resources is becoming increasingly important when consulting a source of information. When researchers weigh the factors that lead them to decide which articles to read, the second most relevant factor, following the topic of the article, is the ease of access online. According to a recent survey, the ease of access is the main reason to choose to read a publication, even more so for young researchers [1].

A recent analysis of article downloading developed by researchers from 4 universities in Spain (Universidad de Burgos, León, Salamanca, and Valladolid) of the Elsevier, Emerald Springer, and Wiley journals shows that the downloads found have increased in the years 2012-2018, with Elsevier and ScienceDirect being the most used provider by far compared to the rest 2. A priori, the demand for information is elastic, due to the variety of preferences of researchers, the search for information is concentrated in a limited number of reputable titles and linked to obtaining stability in academia [3]. Therefore, several journals in these databases have little or no use; however, the increase in downloads that has been observed leads to an increase in the reading and citation of articles from these journals, so it is needed to verify if the rise in the use of subscribed electronic journals is accompanied by an increase in their use in the publications of their researchers through the citations they make.

Citation analysis has been widely used to demonstrate the value of collections since the 1970s. In one study, they used a sample of journals and their references by combining the data obtained with the results of a survey on reading trends to evaluate the collections of a university library, using the Web of Knowledge to identify the journals where professors publish and the journals they cite in their publications [4]. The use of download statistics as the only measure to evaluate journals and/or collections is not recommended, as demonstrated in a paper. It is indicated that several factors make it difficult to assign value to journals: the design of the platforms, the variety of titles included in the packages, the amount of content in the journals, and the changes associated with them -changes in title, platform, publisher -, and the difficulty of assigning a value to the items. It rightly shows that statistics need close scrutiny for peculiar anomalies before they can be considered valid [5]. Other research shows that article downloads only imply readings sometimes, or a real use of the article. Collection evaluation should not be based solely on usage of statistics because these show more of the usefulness of the resources being evaluated than their value, that is, the purpose of a resource [**?** , 6] In addition, another difficulty in assessing the collections is that the use of the articles varies in purpose and intensity depending on the disciplines, so without knowing the data environment, it is not possible to attribute an adjusted value. Statistics without context have relative value "it is dangerous to assume that a popular title is more valuable because students used it, than a specialized journal article used by few researchers in a specific discipline" [7].

While many examples in the literature rely on usage data (article downloads) or citation analysis, there are few studies that assemble data sets that combine both evaluative methods [8]. A recent study considering downloads vs. citations has concluded that Journals are the only means of diffusion that consistently fulfills all the functions that they have traditionally attributed -registration, curation, evaluation, distribution, and archiving- and that allow to institutionalize and confidently add a contribution to the body of knowledge. Along with the metrics based on the positioning of the journals in the main bibliographic databases or on the citations received, other alternative metrics have emerged, including download data. These alternative metrics are in the process of consolidation, mainly because they need to be studied in relation to traditional metrics. It was confirmed that there is a significant correlation between citations and downloads. The downloads indicate the use of certain contents, and the citations indicate the usefulness of these same contents to build new research [9].

Social networks are other means of disseminating scientific information. Alternative metrics have been generated evaluating the impact of scientific publications on social networks. In a study, we evaluated whether there is a correlation between the activity of internal medicine journals on social networks and traditional journal activity on Facebook, Twitter, YouTube, and Instagram. According to Scimago, SJR index was higher in journals with social networks vs. journals without social networks, showing a strong correlation between social network activity metrics, compared to traditional metrics based on the number of citations in internal medicine journals [6].

Additionally, Counter usage statistics are not yet considered to be a strong enough foundation to build a new global measure like Usage Factor (UF), but confidence in them is growing, and they are seen as the only viable foundation for UF. Opinions vary widely on how to calculate a usage factor, and in particular on how to define the following terms: "total usage," "specified usage

DOI: 10.17533/udea.redin.20230214

period," and "total number of articles published online." To refine the definitions of these terms, it will be necessary to carry out tests with real use data. Counter, in general, is trusted by both librarians and publishers, and it is seen as having a role in the development and maintenance of UFs, possibly in collaboration with another information industry's organization. The organization performing this function must be trusted by both librarians and publishers, and must include representatives from both groups. However, several structural issues with online usage data would need to be accounted for in order for UFs to be credible. The perception that online usage data is much more manipulable than citation data is noticeable. A typical large academic library subscribes to more than 20,000 journals. Only about 8,000 of them are indexed in the Science Citation Index, which acts as a source to calculate the Impact Factors. One of the advantages of the UF is that, in theory, it could cover all online journals [10].

References

[1] D. Nicholas, H. R. Jamali, E. Herman, A. Watkinson, A. Abrizah, and *et al.*, "A global questionnaire survey of the scholarly communication attitudes and behaviours of early career researchers," *Learned publishing*, vol. 33, no. 3, Jun. 30, 2020. [Online]. Available: https://doi.org/10.1002/leap.1286

[2] F. Shu, P. Mongeon, S. Haustein, K. Siler, J. P. Alperin, and *et al.*, "Is it such a big deal? on the cost of journal use in the digital era," *Profesional de la Información*, vol. 79, no. 6, 2018. [Online]. Available: https://doi.org/10.5860/crl.79.6.785

[3] C. S. Wilson and C. Tenopir, "Local citation analysis, publishing and reading patterns: Using multiple methods to evaluate faculty use of an academic library's research collection," *Journal of the association for information science and technology*, vol. 59, no. 9, Apr. 28, 2008. [Online]. Available: https://doi.org/10.1002/asi.20812

[4] T. Bucknell, "Garbage in, gospel out: Twelve reasons why librarians should not accept cost-per-download figures at face value," *The Serials Librarian*, vol. 63, no. 2, Aug. 02, 2012. [Online]. Available: https://doi.org/10.1080/0361526X.2012.680687

[5] J. Luther, "White paper on electronic journal usage statistics," *The Serials Librarian*, vol. 41, no. 2, Oct. 24, 2008. [Online]. Available: https://doi.org/10.1300/J123v41n02_10

[6] V. Martin, T. Gray, M. Kilb, and T. Minchew, "Analyzing consortial "big deals" via a cost-per-cited-reference (cpcr) metric," *Serials Review*, vol. 42, no. 4, Dec. 07, 2016. [Online]. Available: https://doi.org/10.1080/00987913.2016.1248218

[7] B. Rodríguez-Bravo, A. Fernández-Ramos, and C. Travieso-Rodríguez, "Relación entre descargas y citas de revistas científicas en el ámbito de la documentación: el caso de las universidades públicas de castilla y león," *Revista Española de Documentación Científica*, vol. 44, no. 4, Oct. 04, 2021. [Online]. Available: https://doi.org/10.3989/redc.2021.3.1806

[8] M. M. Bustos-Claro, V. Ávila, D. G. Fernández-Ávila, . Muñoz-Velandia, and . A. García, "Relationship between internal medicine journals' activity on social networks and the citations they receive," *Revista Clínica Española*, vol. 222, no. 1, Jan. 2022. [Online]. Available: https://doi.org/10.1016/j.rce.2020.10.007

[9] T. Baiget. (2011, Nov. 16,) Juzgar la calidad de las revistas según las descargas de artículos no es fácil. [Online]. Available: https://tinyurl.com/msycnhb3

[10] A. Fernández-Ramos, B. Rodríguez-Bravo, M. L. Alvite-Díez, L. Santos-De-Paz, M. A. Morán-Suárez, and *et al.*, "Evolution of the big deals use in the public universities of the castile and leon region, spain," *Profesional de la Información*, vol. 28, no. 6, Jan. 09, 2020. [Online]. Available: https://doi.org/10.3145/epi.2019.nov.19

Maryory Astrid Gómez Botero
Editor-in-Chief
Revista Facultad de Ingeniería -redin-
Professor-Universidad de Antioquia
https://orcid.org/0000-0001-9685-3080
http://www.redalyc.org/autor.oa?id=8587
https://scholar.google.com/citations?user=U_2Xx_cAAAAJ&hl=es