



Hidden limitations of analyses via alternative bibliometric services

Libor Ansorge¹

Received: 16 November 2022 / Accepted: 20 December 2022 / Published online: 31 December 2022
© The Author(s) 2022

- Assigning a DOI to scientific publications increases the visibility of these works to the scientific community a bibliometric services.
- Alternative bibliometric services provide citation data usually based on CrossRef data.
- Despite its wide scope of scientific literature, CrossRef is still a selective database.

Discussion

Small professional and scientific journals publishing in national languages have a difficult position in international comparisons. This is due to developments in recent years, when English has become the dominant academic language (*Lingua Franca*) with both positive and negative impacts on the scientific community (Kamadjeu, 2019; Suzina, 2021; Tardy, 2004). Articles written in English receive higher citation rates (Di Bitetti & Ferreras, 2017) on the other hand, national research published in an English-language international journal may generally be inaccessible to experts in the country (Raitskaya & Tikhonova, 2020). As Kirchik et al. (2012) showed, articles published in local languages are usually cited again by articles published in the local language. As a manager of a scientific institution that publishes its own scientific journal, I continuously monitor the citation rate of our journal *Vodohospodářské technicko-ekonomické informace – VTEI* (in English: *Water Management Technical and Economic Information*) in the Scopus database (Ansorge, 2022).

At the same time as writing the above mentioned citation analysis, I did the same citation analysis for the "competing" journal *Vodní Hospodářství* (in English: *Water Management*). The two journals are similar in many ways. Both have been published regularly for more than 50 years. Both journals are aimed at the Czech and Slovak water management community and publish informative, technical and scientific articles mainly in Czech or Slovak. At the same time, they differ significantly in one area. This difference is the electronic presentation of the journal, although both editorial offices use WordPress as the default content management system for web presentation. *Vodní hospodářství* journal is still primarily published in print version and the electronic presentation of the journal at www.vodnihospodarstvi.cz is more of a supplement to the printed version of the journal than a full-fledged presentation of the content. Only selected articles are published in html, there is no DOI identifier in place, and the metadata for each article contains only basic

✉ Libor Ansorge
libor.ansorge@vuv.cz

¹ T. G. Masaryk Water Research Institute, Podbabská 2582/30, 160 00 Prague, Czech Republic

metadata provided by the Yoast SEO plugin. The PDF version of the full issues is published one time per year. In contrast, the *VTEI* editorial team introduced a brand new journal website in 2015, www.vtei.cz, with Dublin Core and Google Scholar metadata included for each article. A persistent Digital Object Identifier – DOI has been additionally assigned to all articles published on the new website.

The main problem of the citation analysis of both journals turned out to be the ambiguity of the Czech journal name entry in the Scopus bibliographic database. Together with the fact that both journals are not indexed in the Scopus database, verifying citations took the most time and manual work. Therefore, as a next step, I started to look for other bibliographic services that would provide information on the citation of articles in regional journals. I focused my attention beyond the "big three" sources of citation data (Scopus, Web of Science, and Google Scholar). I included Dimension.AI, Scilit, OpenAlex, Lens.org and Exaly in my research. An interesting finding of my research is that in all of these bibliographic services, articles published in the journal *VTEI* can be traced, but not articles published in the journal *Vodní hospodářství*. That is, to be precise, data for 226 articles that have been assigned a DOI can be traced in these services. The exception was the Exaly service, which "sees" only 166 articles in the *VTEI* journal. It was verified on another Czech journal, *Journal of Competitiveness*, that these bibliographic services do indeed limit their content through DOIs. The *Journal of Competitiveness* is included in the Directory of Open Access Journals (DOAJ) and has metadata filled in for articles since 2009 in this repository. The journal only introduced the DOI in 2012. However, the bibliographic services mentioned above only "see" articles published in *Journal of Competitiveness* from 2012 onwards, i.e. since the introduction of the DOI.

On the initiative of the publishing industry, the International DOI Foundation was established in 1998 to prepare the necessary technologies and standards for the implementation of the DOI system. The purpose of the DOI system is to maintain a permanent reference to an electronic resource (document) whose real web address may change quite frequently with the introduction of new technologies, e.g. upgrades to the content management systems used by individual journals. Due to characteristics of DOI such as uniqueness, actionability, interoperability, persistence, and granularity, scholarly works with DOIs could be identified, retrieved, cited, transferred, and preserved in an effective way (Liu, 2021).

DOI registration agencies such as CrossRef collect bibliographic metadata for digital content making these agencies well positioned to provide text and data mining services (Lammey, 2015). The availability of text and data mining services provided by CrossRef seems to be the pillar for the bibliographic services mentioned above. However, this means that bibliometric and scientometric analyses based on these bibliographic services do not include data from articles that do not have a DOI assigned. At present, Google Scholar is probably the most comprehensive source of bibliographic data. Web of Science and Scopus are selective databases that do not represent global scientific production (Tennant, 2020). And although the metadata behind DOI is a rich source of data for research in bibliometrics and scientometrics (Hendricks et al., 2020), the bibliographic services mentioned above again represent only a selection of scientific production. Further research will be needed to demonstrate the large volume of scholarly publications not covered by these bibliographic services. Initial studies identifying the number of articles with and without DOIs have already been conducted (Khurana et al., 2023).

At the same time, it is also a definite signal to managers of scientific journals that DOI is one way to increase the visibility of its journal.

Funding The author did not receive support from any organization for the submitted work.

Declarations

Conflict of interest The author have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ansoorge, L. (2022). Citační analýza časopisu VTEI. *Vodohospodářské Technicko-Ekonomické Informace*. <https://doi.org/10.46555/VTEI.2022.05.001>
- Di Bitetti, M. S., & Ferreras, J. A. (2017). Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. *Ambio*, 46(1), 121–127. <https://doi.org/10.1007/s13280-016-0820-7>
- Hendricks, G., Tkaczyk, D., Lin, J., & Feeney, P. (2020). Crossref: The sustainable source of community-owned scholarly metadata. *Quantitative Science Studies*, 1(1), 414–427. https://doi.org/10.1162/qss_a_00022
- Kamadjeu, R. (2019). English: The lingua franca of scientific research. *The Lancet Global Health*, 7(9), e1174. [https://doi.org/10.1016/S2214-109X\(19\)30258-X](https://doi.org/10.1016/S2214-109X(19)30258-X)
- Khurana, P., Ganesan, G., Kumar, G., & Sharma, K. (2023). A Bibliometric Analysis to Unveil the Impact of Digital Object Identifiers (DOI) on Bibliometric Indicators. In P. K. Singh, S. T. Wierzchoń, S. Tanwar, J. J. P. C. Rodrigues, & M. Ganzha (Eds.), *Proceedings of Third International Conference on Computing, Communications, and Cyber-Security* (pp. 859–869). Springer Nature. https://doi.org/10.1007/978-981-19-1142-2_67
- Kirchik, O., Gingras, Y., & Larivière, V. (2012). Changes in publication languages and citation practices and their effect on the scientific impact of Russian science (1993–2010). *Journal of the American Society for Information Science and Technology*, 63(7), 1411–1419. <https://doi.org/10.1002/asi.22642>
- Lammey, R. (2015). CrossRef text and data mining services. *Science Editing*, 2(1), 22–27. <https://doi.org/10.6087/kcse.32>
- Liu, J. (2021). Digital object identifier (DOI) and DOI services: An overview. *Libri*, 71(4), 349–360. <https://doi.org/10.1515/libri-2020-0018>
- Raitskaya, L., & Tikhonova, E. (2020). Pressure to publish internationally: scholarly writing coming to the fore. *Journal of Language and Education*, 6(1), 1. <https://doi.org/10.17323/jle.2020.10631>
- Suzina, A. C. (2021). English as lingua franca. Or the sterilisation of scientific work. *Media, Culture & Society*, 43(1), 171–179. <https://doi.org/10.1177/0163443720957906>
- Tardy, C. (2004). The role of English in scientific communication: Lingua franca or Tyrannosaurus rex? *Journal of English for Academic Purposes*, 3(3), 247–269. <https://doi.org/10.1016/j.jeap.2003.10.001>
- Tennant, J. P. (2020). Web of science and scopus are not global databases of knowledge. *European Science Editing*, 46, 51987. <https://doi.org/10.3897/ese.2020.e51987>