New Methodology for Calculating Webometrics University Ranking: From Google Scholar to OpenAlex

Vladimír Moskovkin *

Posted Date: 18 August 2025

doi: 10.20944/preprints202508.1191.v1

Keywords: Webometrics University Ranking; Institutional Google Scholar Citation profiles; OpenAlex; ROR identifier; Figshare

 doi:10.20944/preprints202508.1191.v1

Article

New Methodology for Calculating Webometrics University Ranking: From Google Scholar to OpenAlex

Vladimir M. Moskovkin

Doctor of Geographical Sciences, Independent Researcher, Vimperk, Czech Republic; researchoa52@gmail.com; Tel.: +420721837839

Abstract

The Webometrics University Ranking website ceased to function in 2025 due to an inability to obtain citation data from Google Scholar. Since then, Webometrics University Ranking data has been published on the Figshare server, but the values of the three individual indicators have not been ranked. From July 2025 onwards, the Openness indicator values for citations have been calculated using OpenAlex via the ROR identifier. Data on the ranking of all three indicators will be provided twice a year in the form of an Excel file on a paid basis. Examples of universities included in the top 1,000 of the January 2024 and January 2025 Webometrics Rankings, which had missing legitimate Institutional Google Scholar Citation profiles, demonstrate a sharp increase in their rankings when switching to the new methodology for calculating Openness indicator values. The importance of creating and maintaining Institutional Google Scholar Citation profiles, despite changes in the Webometrics University Ranking calculation methodology, has been demonstrated.

Keywords: Webometrics University Ranking; Institutional Google Scholar Citation profiles; OpenAlex; ROR identifier; Figshare

Introduction

Since 2004, the Spanish Cybermetrics Laboratory has been calculating the Webometric Rankings of the World's Universities that have autonomous domains (URL addresses) (www.webometric.info). This global university ranking is on par with the THE-QS University Rankings, as well as with the Shanghai and Taiwanese University Rankings. The following indices are measured based on the responses received from four high capacity search engines (Google, Yahoo, Live Search, and Exalead): SIZE (the total number of pages obtained from the abovementioned engines for each university domain), VISIBILITY (the total number of unique external citations obtained with the help of the last three engines), RICH FILES (the number of pdf, ps, doc, and ppt file formats obtained using the first engine), SCHOLAR (the number of academic documents and their citation obtained with the help of the Google Scholar search engine); and subsequently, the integral webometric index, according to which the world's universities are ranked, was calculated using special mathematical procedures, including logarithmic normalization and weighing, which made it possible to build an integral indicator on an additive basis (Aguillo et al. 2006).

The launch of this ranking took a lot of attention from university management around the globe, since, unlike all other rankings, it made it possible to rank almost all universities in the world. Therefore, there naturally appeared a great interest towards analyzing this ranking in comparison with other rankings in scientific discourse, as well.

Since the launch of the Webometrics University Ranking (2004), its methodology has been constantly changing, but in July 2016, the most significant change happened. If originally the third indicator (Openness) concerned the number of PDF files affiliated with the University site found

 doi:10.20944/preprints202508.1191.v1

through the Google Scholar search engine, from the second half of 2016, this indicator shifted towards the citations found through the same search engine, with the Openness indicator weighing only 10%. Naturally, most of the world universities were not ready for this change. They should have had their scientists create personal Google Scholar Citation profiles in advance, and those profiles tied to the university domain would have automatically made-up institutional Google Scholar Citation profiles (IGSCPs). That very year also saw the introduction of the bibliometric Excellence indicator, weighing as much as 30%, to be calculated on the basis of the Institute Top-10% most cited Scopus papers profiles.

However, it should be noted that Google Scholar launched a new service called "Google Scholar Citations" on 16 November 2011 (Connor, 2011), i.e. five years before its integration into the Webometrics Ranking.

In July 2016, out of about 22, 000 universities of the world ranked with Webometrics, only 4,120 universities had IGSCPs. The further growth in the number of universities with such profiles had the following pattern: in January 2017 – 8, 634 (out of 26, 000 ranked universities), in July 2017 – 9,491, in January 2018 – 9,593, in July 2018 – 10,778 (Moskovkin, Yawei, Sadovski, 2019).

The illegitimacy of the IGSCP in the Webometrics Ranking (January 2018) of Boris Grinchenko Kyiv University was pointed out by Buinytska, Hrytseliak, & Smirnova (2018). They noted that the list of Personal GSCPs of this university included profiles of departments.

A comprehensive quantitative and qualitative analysis of the legitimacy of IGSCPs has been conducted for the Asian University Webometrics Ranking by V. M. Moskovkin and O. V. Serkina (2024).

I have not encountered other works devoted to a large-scale analysis of the legitimacy of IGSCPs, which is largely related to errors in the functioning of Google Scholar, as well as errors made by the creators of these profiles. In the early years of this search engine's operation, its errors were tracked on an annual basis by P. Jasco (2005, 2012).

The most comprehensive list of such errors immediately after the launch of the Google Scholar Citation service in calculating the Openness indicator in the Webometrics University Ranking was identified by A. Martin-Martin et al. (2016). They noted that errors that can compromise the metric portrait of an author offered by Google Scholar can be grouped into two main sections. First, the errors Google Scholar sometimes makes when it indexes a document or when it assigns citations to it. Second, the specific errors that are sometimes made during the creation of a Google Scholar Citations profile. Their list is given below:

a) Incorrect identification of the title of the document;
b) Ghost authors;
c) Book reviews indexed as books;
d) Incorrect attribution of documents to authors;
e) Failing to merge all versions of a same document into one record;
f) Grouping different editions of the same book in a single record;
g) Improper attribution of citations to a document;
h) Duplicate citations;
i) Missing citations;
j) Duplicate profile;
k) Variety of document types (including non-academic documents);
l) Inclusion of missattributed documents in the profile;
m) Deliberate manipulation of documents and citations in Google Scholar;
n) Duplicate documents in profiles;
o) Incorrectly merged documents;
p) Unclean document titles;
q) Missing or uncommon areas of interest.

 doi:10.20944/preprints202508.1191.v1

Thus, I have shown that after the introduction of of the Openness Indicator in the calculation of the Webometrics University Ranking, university management and researchers have been faced with the problem of the correct creation of personal GSCPs, which still receives very little attention. However, at the beginning of 2025, another, but more radical, change in the Webometrics University Ranking calculation methodology took place. IGSCPs were abandoned in the calculation of the Openness indicator.

Let us briefly consider the reasons for this abandonment and the new methodology for calculating this ranking.

New Methodology for Calculating the Webometrics University Ranking and Its Resource Support

In 2025, the Webometrics Ranking website ceased to function. As Isidro Aguillo, the head of the Webometrics project, informed me: "*Unfortunately, Google Scholar no longer allows data extraction,*" which makes it impossible to calculate the Openness indicator based on Institutional Google Scholar Citation profiles (IGSCPs).

The January 2025 Webometrics University Ranking, calculated using the old methodology, are published on the Figshare server: https://doi.org/10.6084/m9.figshare.28284617.v2

It ranks about 32,000 HEIs from over 200 countries. Data on the ranks of the Visibility, Excellence and Openness indicators are not provided. The first indicates the ROR identifier for universities. The Research Organisation Registry (ROR) is a global, community-led registry of open, persistent identifiers for research organisations. ROR makes it easy for anyone or any system to disambiguate institution names and connect research organisations to researchers and their outputs.

The July 2025 Webometrics University Ranking, calculated using a new methodology, has been published on the Figshare server: https://doi.org/10.6084/m9.figshare.29588921.v1

In this methodology, the values of the Visibility and Excellence indicators are calculated as before, while the values of the Openness indicator are calculated using the number of citations for the period 2020-2024 of institutions with the same weight (10%). These calculations are performed using OpenAlex via the ROR identifier. Approximately the first 8,000 universities have an ROR identifier of the form https://ror.org/03vek6s52 for Harvard University.

After the 8,000th place in the ranking, we see universities with missing ROR identifiers, and the closer to the end of the ranking, the more such universities there are. For universities with missing ROR identifiers, Total Citation calculations are not performed.

University managers and researchers can check the total number of citations over the last five years using the ROR identifier. For example, Leiden University's ROR identifier is https://ror.org/027bh9e22 528. To view its publication activity and citation rate, the following query should be made: https://api.openalex.org/institutions/ror:https://ror.org/027bh9e22

As Isidro Aguillo informed me, "*a full Excel file with a breakdown of indicators will be available at a price of 200 euros*."

In July 2025, all leading universities with missing legitimate IGSCPs sharply improved their positions in the ranking.

Earlier, in the TOP 1,000 universities of the July 2023 Webometrics Ranking, I observed 17 universities with missing legitimate IGSCPs. In January 2024, there were 10 such universities, which are shown in the Table 1.

Table 1. World Ranks of universities with missing legitimate IGSCPs included in the TOP-1,000 Webometrics University Ranking in January 2024 and January 2025, которые были улучшены в июле 2025 года при переходе на новую методологию расчета индикатора Openness.

 doi:10.20944/preprints202508.1191.v1

University	Jan. 2024	Jan. 2025	July 2025
Johns Hopkins University, Sch..of Med. (USA)	338	267	45
University of Birmingham (UK)	515	485	116
Leiden University (Netherland)	532	528	146
University of Minnesota Twin Cities (USA)	553	724	264
Hong Kong University of Sci.& Tech. (Hong Kong)	769	812	280
Weill Medical College, Cornell University (USA)	802	1148	555
Universidade Nova de Lisboa (Portugal)	839	881	372
University of Hawai at Manoa (USA)	897	1099	492
Universite du Quebec (Canada)	962	1036	601
Swinburne University of Technology (Australia)	975	963	433

In January 2025, their World Ranks did not change significantly, as the methodology for calculating them remained unchanged, but in July 2025, these rankings improved dramatically due to a change in methodology. In January 2025, the number of citations Top 310 authors (excluding the Top 20 outliers) according to Google Scholar for the universities under consideration were not taken into account, as there were no IGSCPs, and in July 2025, their absence was compensated for by total citations according to OpenAlex.

Considering the data for the first university to be a statistical outlier, we will calculate the average change in university rankings for the remaining nine universities. It is equal to 2.67, meaning that the universities with missing legitimate IGSCPs improved their positions in the Webometrics University Ranking by an average of 2.67 times when the calculation of total citations was switched from Google Scholar to Open Alex.

Let us now examine how the positions of universities have improved following the restoration of their illegitimate IGSCPs under the old calculation methodology. In July 2023, there were five universities with illegitimate IGSCPs in the TOP-1,000 Webometrics University Ranking, which restored their profiles in July 2024 (Table 2).

Table 2. World Ranks of universities with missing legitimate IGSCPs included in the TOP-1,000 Webometrics University Ranking in July 2023, которые восстановили свои профили в июле 2024 года.

University	July 2023	July 2024
Universidad de Sevilla (Spain)	748	285
University of Warsaw (Poland)	772	353
Universitat Pompeu Fabra (Spain)	772	266
Universidade Federal do Rio Grande do Sul (Brazil)	844	340
Universidade de Coimbra (Portugal)	851	315

From this table, we can calculate the average change in university rankings, which is 2.58. Consequently, the universities with missing legitimate IGSCPs improved their positions in the Webometrics University Ranking by an average of 2.58 times when their illegitimate IGSCPs were restored after a year.

 doi:10.20944/preprints202508.1191.v1

As we can see, similar average values for changes in world rankings were obtained in both cases. This suggests that the transition to the new calculation methodology will have little effect on university rankings. When creating personal GSCPs, researchers try to include their best publications from the Scopus and WoS databases, as well as all their Open Access publications with a DOI that are included in DOAJ. Virtually all of these publications are indexed by OpenAlex. The work of D. Chavarro, J.P. Alperin and J. Willinsky (2025) supports this. They made the following conclusions:

"To summarize, items with a Crossref DOI have a 96% chance of appearing in OpenAlex. Journals included in the Directory of Open Access Journals (DOAJ) without a DataCite DOI (and without a Crossref DOI) have an 81% chance of inclusion in Open Alex. Journals in Scopus, while lacking a Crossref DOI and absent from DOAJ also have an 81% chance of appearing in OpenAlex."

At the same time, the rejection of IGSCPs cuts off citations of old publications from closed sources and grey literature. Nevertheless, I believe that the correlation coefficient between the total citation values calculated by IGSCPs and OpenAlex will be close to unity. The same can be assumed for Spearman's rank correlation coefficient between World Ranks calculated using the two methodologies.

The advantage of switching to citation counting using OpenAlex is related to the synchronisation of the five-year time interval when calculating top cited papers (Excellence indicator) with the calculation of total citations (Openness indicator).

The Importance of Establishing and Maintaining Institutional and Personal GSCPs

Transition to a new methodology for calculating Webometrics University Ranking, however, this does not mean that university managers should ignore the creation of legitimate IGSCPs.

Despite the fact that Webometrics University Ranking will switch to OpenAlex for calculating the Openness indicator from July 2025, and therefore IGSCPs will no longer directly influence this specific indicator in Webometrics, it is still extremely important for universities to create and maintain these profiles, as well as encourage their staff and students to create and maintain individual GSC profiles for several reasons:

1. **Visibility and reach**: Google Scholar remains one of the most popular and accessible search engines for scientific literature worldwide. Many researchers, students, and even ordinary users start their search there. A well-maintained GSC profile for a university or individual researcher significantly increases their visibility and chances of being found, and improving the visibility of publications increases the likelihood of them being cited.

2. **Popularity and habit**: For many researchers, GSC is the de facto tool for tracking their own citations and searching for literature. It would be unwise to ignore such a large audience and user habit.

3. **Alternative assessment**: Although Webometrics is changing its methodology, GSC provides its own citation metric, which differs from OpenAlex (due to its broader coverage of sources) but indirectly influences it. This gives the university another perspective for assessing its scientific influence.

4. **Attracting talent and partnerships**: Researchers, students, and partners often use GSC to assess the scientific activity of the university and its staff. Maintaining up-to-date profiles helps demonstrate the strength of the university's scientific community.

5. **Internal monitoring and self-assessment**: Universities can use GSC data for internal monitoring of the publication activity and citation rates of their faculties and researchers, supplementing information from other databases.

6. **Reputation management**: An active and well-structured GSC profile helps the university manage its online reputation and present its scientific activities in the most favourable light.

 doi:10.20944/preprints202508.1191.v1

Thus, the change in the Webometrics University Ranking methodology does not negate the importance of Google Scholar Citation profiles. They remain a valuable tool for promoting, monitoring and evaluating scientific activity, although their role in the Webometrics ranking has changed.

Despite the fact that the webometrics.info website has ceased to function, the significance of the Webometrics Ranking has not declined, as evidenced by the high number of views and downloads on the Figshare server. According to data from 13 August 2025, the number of views and downloads for the January 2025 ranking was 126,091 and 34,429, respectively, and for the July ranking, it was 20,862 and 5,461.

Conclusions

Google Scholar's refusal to provide Cybermetrics Lab with data for calculating the Webometrics University Ranking led to the need to find a replacement for calculating the Openness indicator. OpenAlex proved to be a good replacement.

Difficulties in maintaining the informetrics.info website led to its closure and to the transition of publishing this ranking twice a year on the Figshare server with an additional fee for Excel files containing complete data on three individual indicators. I hope that this website will be restored in the future.

Experiments on the webometric ranking of universities with missing IGSCPs included in the TOP 1,000, when restored, showed identical average changes in world rankings, both in experiments using the old methodology and when transitioning from the old methodology to the new one. This circumstance, as well as a comparison of the content of the Google Scholar and OpenAlex databases, led to the conclusion that the transition to the new methodology will have virtually no effect on university rankings, although this conclusion requires experimental confirmation.

The necessity of continuing to support Institutional and Personal GSCPs is justified, as they will, among other things, have a positive impact on the growth of total citations in OpenAlex calculations.

It has been shown that the advantage of switching to citation counting using OpenAlex is related to the synchronisation of the five-year time interval when calculating the top cited papers (Excellence indicator) with the calculation of total citations (Openness indicator).

In conclusion, it should be noted that during its twenty-year evolution, Webometrics University Ranking has become a full-fledged scientometric ranking, similar to other prestigious rankings, whose only drawback is that, after the introduction of the Excellence indicator, it ceased to correspond to its webometric concept, since the closed statistical data on top cited papers for this indicator are taken from the Scimago laboratory, rather than from the Internet.

References

Aguillo, I. F., Granadino, B., Ortega, J. L., & Prieto, J. A. F. (2006). Scientific research activity and communication measured with cybermetrics indicators. *Journal of the American Society for information science and technology*, *57* (10), 1296-1302. https://isidroaguillo.webometrics.info/sites/default/files/publicaciones/Aguillo2006Scientific_research_activity_and_communication_measured_with_cybermetric_indicators..pdf

Buinytska, O., Hrytseliak, B., & Smirnova, V. (2018). Rating as assessment tool of quality and competitiveness of university. *Open educational e-environment of modern University*, *4*, 16-32. https://elibrary.kubg.edu.ua/id/eprint/24061/

Chavarro, D., Alperin, J. P., & Willinsky, J. (2025). On the Open Road to Universal Indexing: OpenAlex and Open Journal Systems. *Quantitative Science Studies*, 1-28. https://doi.org/10.1162/qss.a.17

Connor, J. (2011). Google Scholar citations open to all. http://googlescholar.blogspot.com/2011/11/google-scholar-citations-open-to-all.html

Jacso, P. (2005). 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*, *89*(9), 1537-1547.

 doi:10.20944/preprints202508.1191.v1

https://www.jstor.org/stable/24110924

Jasko, P. (2012). Google Scholar Author Citation Tracker: is it too little, too late? *Online Information Review*, *36* (1), 126-141. https://doi.org/10.1108/14684521211209581

Martín-Martín, A., Orduna-Malea, E., Ayllón, J. M., & Delgado López-Cózar, E. The counting house: measuring those who count. Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in the Google Scholar Citations, ResearcherID, ResearchGate, Mendeley & Twitter. *EC3 Working Papers*, 21. 19th of January 2016. 60 pages, 12 tables, 35 figures. https://arxiv.org/ftp/arxiv/papers/1602/1602.02412.pdf

Moskovkin, V., Yawei, L., & Sadovski, M. (2019). Identification of leading Russian universities without profiles in Google Scholar citation. *Alma Mater*, *1*, 10-15. [in Russian]

https://www.researchgate.net/publication/331150996_Moskovkin_VM_Identifikacia_vedusih_rossijskih_universitetov_s_otsutstvuusimi_v_Google_Scholar_Citation_profilami_VM_Moskovkin_Lu_Avej_MV_Sadovski_Alma_mater_-_2019_-_No1-_S_10-15_Identification_of_lea#fullTextFileContent

Moskovkin, V. M., & Serkina, O. V. (2024). Quantitative Frequency Analysis of Google Scholar Citation and Top 10% Most Cited Scopus Papers Profiles for Asian University Webometrics Ranking. *Informology*, *3*(2), 11-32. https://www.informology.org/2024/v3n2/a36.pdf