How well does OpenAlex cover the Flemish Social Sciences and Humanities?

Eline Vandewalle[1], Cristina Arhiliuc[2]

[1]eline.vandewalle@uantwerpen.be, [2]cristina.arhiliuc@uantwerpen.be
Centre for Research and Development Monitoring (ECOOM), University of Antwerp,
Middelheimlaan 1, 2020 Antwerp (Belgium)

Abstract

Since the launch of OpenAlex as a fully open and non-proprietary alternative to bibliographic indexing services, interest has risen in the extent to which OpenAlex covers the research landscape and in what areas it could increase coverage compared to the proprietary alternatives, particularly of the social sciences and humanities (SSH) and for publications in languages other than English. In this study, we have used the VABB-SHW database as a benchmark to compare OpenAlex with. VABB-SHW is a local comprehensive bibliographic database for the SSH. It includes many Dutch-language publications, and non-article publication types. We find that OpenAlex covers 50.46% of publications from the local bibliographic database (both peer-reviewed and non-peer-reviewed publications), with higher percentages for publications that are also indexed in the Web of Science (94.51%). Coverage is lower for non-English language publications and publication types other than articles. Additionally, we explore the metadata coverage in OpenAlex and find that 86 percent of the publications found in OpenAlex have reference data available and 91 percent of them have affiliation information. We also report on the strategy for matching records between the local VABB-SHW database and OpenAlex given the limited availability of DOIs in our local database.

Introduction

OpenAlex has come onto the stage of large indexing databases in late 2022, taking over the backlog of the discontinued Microsoft Academic, and promising an open and non-commercial alternative to indexing databases. Unlike the proprietary alternatives, OpenAlex data can be shared freely under a CC0 license, which enables bibliometricians to share data openly. In the context of initiatives such as the Barcelona Declaration for Open Research Information (Kramer et al. 2024), this is a promising development for the field of bibliometrics. So far, OpenAlex has been used by several major institutions. The French Sorbonne university announced in 2023 that they would unsubscribe from Clarivate-owned Web of Science and opt for a partnership with OpenAlex[1]. Notably the latest Leiden Ranking, published by CWTS has added an open version using OpenAlex as a data source[2]. Another source for enthusiasm regarding OpenAlex is its promise to be both open and comprehensive. The OpenAlex website states: "We strive to be as comprehensive and inclusive as possible, especially for works in other languages and the Global South"[3].

---

[1] https://www.sorbonne-universite.fr/en/news/sorbonne-university-unsubscribes-web-science
[2] https://open.leidenranking.com/
[3] https://help.openalex.org/hc/en-us/articles/24396686889751-About-us

Insufficient coverage of non-English publications and insufficient coverage of the social sciences and humanities (SSH) is a researched limitation of the big international indexing databases (Mongeon & Paul-Hus, 2016; Kulczycki et al. 2018). This has a particularly significant effect on the representation of the SSH since authors from the SSH still publish more frequently in local, non-English language publication channels and books (Kulczycki et al. 2020; Giménez-Toledo 2020). Our goal is to examine to what extent OpenAlex covers SSH publications, including non-English language publications by using the comprehensive bibliographic database VABB-SHW (henceforth VABB) which includes all publications (co-) authored by researchers associated with SSH departments of Flemish universities. So far, around half of peer-reviewed records in the VABB database are covered by the Web of Science. However, Dutch-language publications are less likely to be covered by the Web of Science (only 1.5 percent of WoS-covered publications are classified as Dutch-language in VABB).

OpenAlex can be used as an open bibliometric data source, but for the SSH it is particularly important to track its coverage of diverse publication types and languages other than English. Much of the research on the coverage and metadata of OpenAlex is quite new, and not all has appeared in journal publication form by the time of writing. Researchers have investigated the reference coverage of OpenAlex, Web of Science and Scopus (Culbert et al., 2024) and found that OpenAlex performs similarly to the Web of Science and Scopus in terms of source reference coverage (an important difference is that OpenAlex does not include references to non-source items). Delgado-Quirós and Ortega found that while OpenAlex coverage is high, the source has a low completeness for bibliographic information (pages, issue, volume) (Delgado-Quirós & Ortega, 2024). In recent conference contributions, the coverage and metadata of African publications in OpenAlex, Scopus and Web of Science was investigated (Alonso-Álvarez & van Eck, 2024). Results show that OpenAlex outperforms Scopus and Web of Science in terms of coverage, and some metadata fields (notably ORCID) while underperforming in others. Another contribution has matched OpenAlex with the Norwegian Cristin database and found that OpenAlex covers almost all of the publications that have a DOI in the Cristin database (Armitage and Seland 2024). Maddi et al. (2024) have investigated coverage of Open Access journals in OpenAlex, Scopus and Web of Science and found that OpenAlex offers a comparatively more inclusive coverage of world regions and more balanced coverage of disciplines, with in particular a better representation of the social sciences. Researchers have also looked into the suitability of OpenAlex for bibliometric studies through a comparison with Scopus and concluded that analyses based on the Scopus master list can reliably be repeated with OpenAlex data, but also pointed to some areas of concern, including the completeness and accuracy of metadata, such as the language field (Alperin et al. 2024). Additionally, for the records indexed in the database, several concerns regarding data quality have been discussed. Zhang et al. (2024) concludes that institutional information is missing more frequently than in Web of Science. As mentioned, Delgado-Quirós and Ortega (2024) find that bibliographic information is frequently missing for OpenAlex records. Céspedes et al. (2024) determine that for 14.7% of papers in OpenAlex the

language declared on the platform is incorrect. Jiao et al. (2023) have found inconsistencies in the reporting of document types, with OpenAlex reporting all data articles as regular research articles. As OpenAlex uses data from the previously discontinued Microsoft Academic Graph (MAG), it initially inherited some of the properties of this earlier bibliographic service (Scheidsteger & Haunschild, 2023). OpenAlex lists as its main data sources MAG and Crossref but also sources such as Pubmed and arXiv, and adding additional metadata from ORCID, Unpaywall, ROR and others[4]. However, it is important to keep in mind that OpenAlex is evolving and improving quickly, which means that some issues reported in earlier studies may already be fixed by now.

In this study, we use a comprehensive database as a benchmark, which allows us not only to analyse how coverage compares to the Web of Science, but also to get a full overview of which publications are well-covered by OpenAlex and which publications are missing. The process of matching the regional database with OpenAlex through DOI, ISSN, title and author names is also an important element in this effort. We hope that this analysis may prove useful to researchers planning to use OpenAlex for bibliometric research that includes the SSH and non-English language sources in particular. Additionally, we also hope that this might be of interest to the community around OpenAlex which is working towards improving the database.

In this study we will focus on two main aspects. One is the data matching between the local bibliographic database and OpenAlex. We report on the number of records we were able to match with a record from OpenAlex by three different matching methods. We then report on the number of publications from VABB we were able to find in OpenAlex, the characteristics of those publications and the metadata coverage for those publications. Specifically, we are interested in the following:

1. How many records from VABB (2013-2022) can we find in OpenAlex with different matching strategies?
2. What are the characteristics of the publications we could find/could not find in OpenAlex?
    a. In terms of indexation in the Web of Science and peer review status
    b. In terms of language
    c. In terms of publication type
3. What is the metadata coverage of VABB publications in OpenAlex?
    a. Inclusion of reference/citation information
   b. Completeness of affiliation information

Data

As mentioned, we use the Flemish bibliographic database VABB to compare coverage of publications in OpenAlex. The VABB database is created and

---

[4] https://help.openalex.org/hc/en-us/articles/24397285563671-About-the-data

maintained as part of the Flemish performance-based research funding system. Part of the publications that are in VABB are also indexed in the Web of Science (37.7 %). A second part is not indexed in the Web of Science but is published in publication channels approved by an Authoritative Panel (GP) (32.9%). These publications are considered to be peer-reviewed. A third group of publications consists of publications that were not approved for various reasons (29.3%). We split this group into: publications that were not approved because of formal criteria (missing ISSN/ISBN, missing page info or under 4 pages long) and publications that were not included because they are not considered peer-reviewed by the GP. Figure 1 gives an overview of the peer-reviewed and non-peer-reviewed parts of the dataset. For this study, we use VABB records published between 2013 and 2022, including non-peer-reviewed publications. In the data cleaning process, we removed publications that were not considered as part of the peer-reviewed publications in VABB because they were of the wrong discipline (non-SSH), these were 503 publications. In total, this leaves us with a dataset of 146,680 publications to be matched with OpenAlex. The latest version of the peer-reviewed records in VABB can be accessed online (Aspeslagh et al., 2024).
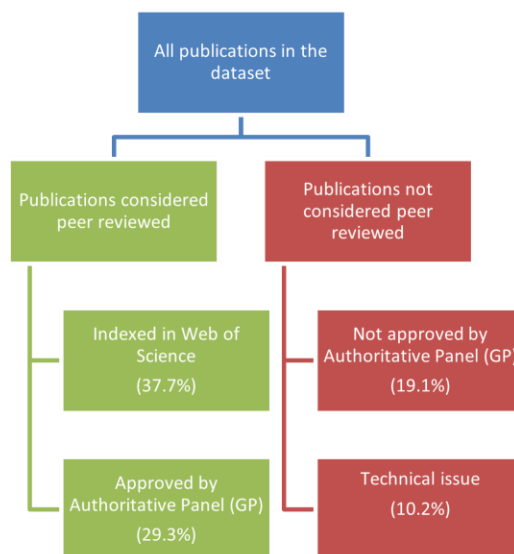
Figure 1. Schema of publications in VABB database.

VABB records belong to one of the following categories: journal article, conference proceedings paper, edited book, book chapter and monograph. Figure 2 gives an overview of the number of records in each of the publication types in VABB.
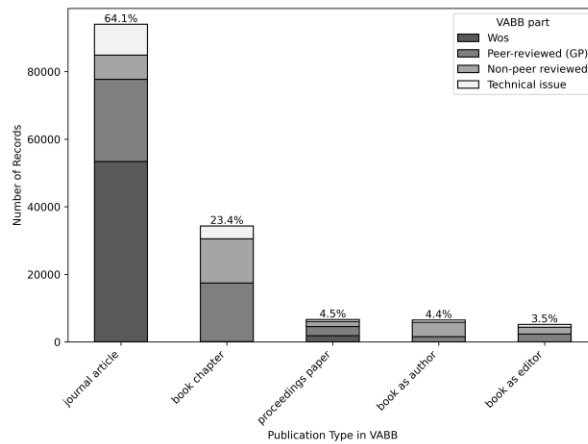
Figure 2. Number of VABB publications per publication type.

A majority (64%) of publications belong to the category journal article. The number of books has been decreasing over the years, but remains an important publication type for SSH although it is not covered well by the Web of Science.

We are using the OpenAlex snapshot of October 2024 (hosted by the Insyspo project). The records could be accessed through Google BigQuery.

Matching procedure

We have adopted a three-step search strategy for identifying VABB publications in OpenAlex (Figure 3). The first step is a matching based on DOI (digital object identifiers). The second step is a matching based on exact title, year (allowing for 1year difference) and at least one author. We chose to allow the publication year to be higher or lower to allow for variations related to preprints and online early access. A third step is a matching based on ISSN, year and author followed by a fuzzy title match. The fuzzy title matching uses the ratio Levenshtein distance. A ratio of above 0.80 is considered a match.
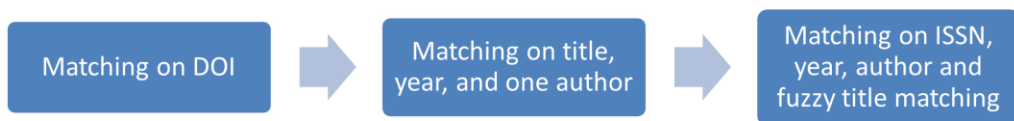
Figure 3. Overview of the matching procedure.

Results

Overview of the number of records matched with the three search steps

We were able to match 74,021 records from VABB to a record in OpenAlex, this is slightly over 50%. Most of the publications could be matched through DOI. Including

the other two search steps yields more incremental gains. Table 1 gives an overview of how many publications can be found with each of the steps.

Table 1. The number of records found with each of the matching steps and the percentage of total publications in VABB and the number of records added by including the step.

Search step	Number of records found in OpenAlex	Number of records added by including the step
Step 1: DOI	65,921 (44.94%)	65,921
Step 2: Exact title	34,636 (23.61%)	6,103
Step 3: Fuzzy title	53,028 (36.15%)	1,997

In total, 67,698 records in the VABB-SHW have a DOI identifier, which is 46.2 percent of the records. Matching on DOI yielded 66,014 matches in OpenAlex which means that 97.5 percent of records with a DOI could be found in OpenAlex. However, there are a few records for which the same DOI was associated with multiple records in VABB. This is the case in particular for book chapters where the DOI listed in VABB refers to the whole book rather than the individual chapter. We excluded these book chapters with the same DOI from the DOI results. In addition, there were 4 DOIs that yielded multiple work_ids. Upon reviewing, we found that one was a mistake in OpenAlex, on an erratum and two cases were preprints, these were excluded as well. With these cleaning steps we arrive at a final set of 65,921 records matched through DOI.

With 97.5 percent of DOI's matched in OpenAlex, matching with DOI has very good results. A recent conference contribution matching academic publications from the Norwegian Cristin database to OpenAlex yielded coverage of 99% for academic works and 97% for the other works (Armitage and Seland 2024). Figure 4 shows the annual number of records with and without DOI in VABB over the time period.
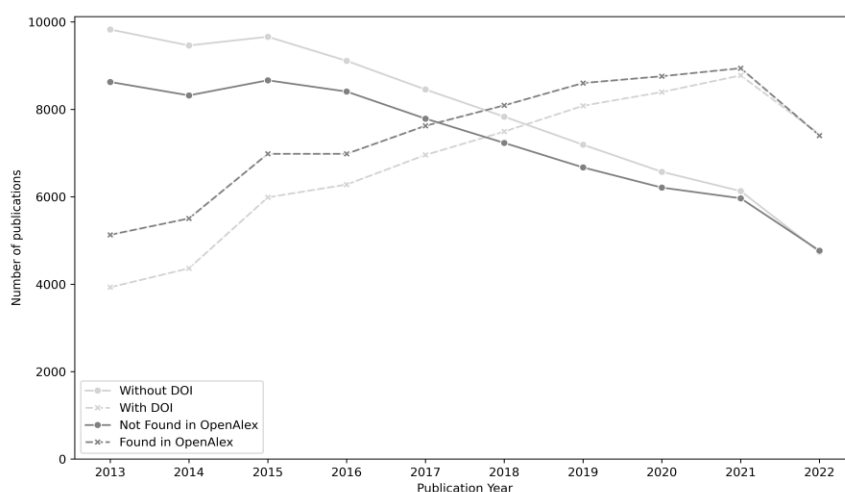

Figure 4. Evolution of the number of records with and without DOI in VABB (2012-2022).

The number of publications with DOI is increasing, while the number of publications without DOI is decreasing. Considering how well publications with DOI are covered in OpenAlex, we expect that the number of publications matched with OpenAlex will increase as more publications are issued a DOI. The increasing availability of DOIs for records in VABB largely tracks the increasing visibility of VABB records in OpenAlex.

The second step was a search for publications for which we did not find a matching DOI. This step consisted of a matching by title, publication year and at least one author. As mentioned above, we allowed the publication year to differ by one. We found that in some cases, the second step found multiple work-id's. In case of multiple work-id's we gave preference to the matching based on DOI for the final dataset as these are more likely to refer to the final publication. There is a significant overlap between publications found with DOI matching and the exact title matching. A third step included a matching by publication year (again allowing a one year difference) and one author as in the previous step. Additionally, we matched on ISSN followed by a fuzzy title matching (using Levenshtein distance ratio of above 0.80). Evidently this search strategy only yields results for records with an ISSN (typically journal publications). All publications that are found in step 2 and that have an ISSN can also be found with the fuzzy title matching. Fuzzy title matching is more computationally intensive and therefore only an option as a 'last resort'. The number of additional records found with the fuzzy matching is limited (1,997).

Figure 5 shows that there is significant overlap in the results obtained with the three search steps, with DOI-based matching yielding the largest number of unique matches.
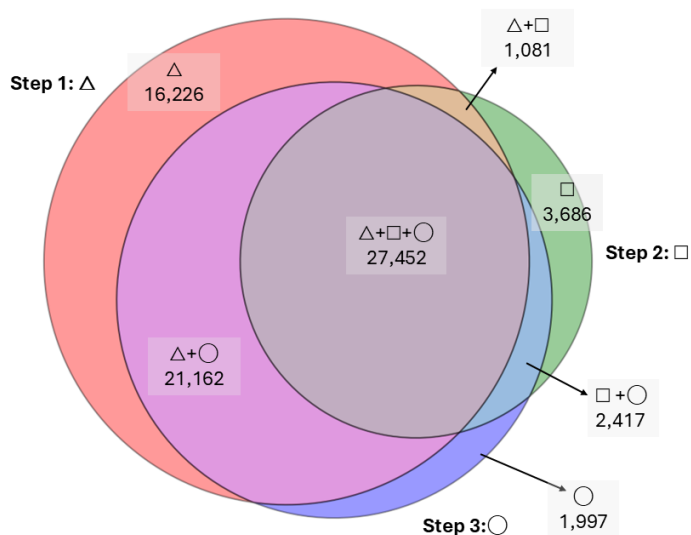
Figure 5. Overlap between records found in each of the search steps.

Characteristics of publications matched with OpenAlex

In table 2, we show the breakdown of coverage in OpenAlex for the publications that are also in the Web of Science, publications that are approved by the Authoritative Panel (considered peer-reviewed) and publications that are not peer-reviewed or do not count in the Flemish PRFS for technical reasons. We do not expect a high proportion of publications that are not considered peer-reviewed to be found in OpenAlex, as these may include grey literature and publications aimed at a broader audience, but we are including them for the sake of completeness. The breakdown shows that most of the publications from our database that are indexed in the Web of Science are also present in OpenAlex. This aligns with findings from previous studies on OpenAlex that have indicated that it provides good coverage for publications indexed in the Web of Science or Scopus (Alperin et al. 2024; Culbert et al. 2024). For publications that are not indexed in the Web of Science, the coverage is lower. Peer-reviewed publications that are not covered in the Web of Science, have a coverage of about 37 percent in OpenAlex.

Table 2. Number and percentage of publications found in OpenAlex according to the different parts of the VABB database (publications indexed in the Web of Science, other peer-reviewed publications (GP), non-peer-reviewed publications and publications not included for technical reasons).

Part of VABB	Found in OpenAlex
Indexed in WOS	52,315 (94.51%)
Other peer-reviewed (GP)	17,954 (37.15%)
Non-peer-reviewed	2,336 (8.35%)
Technical issue	1,416 (9.43%)
Total	74,021 (50.46%)

Considering the coverage per publication type (Table 3), we observe that journal articles are the most comprehensively represented, while only a small proportion of book publications are retrieved. This could be related to our methodology for the retrieval of the information from OpenAlex. It is possible that we are missing book publications because many book publications do not have DOIs and we were unable to conduct searches based on ISBN. Nevertheless, we can assume that coverage is better for journal articles, especially for journal articles in internationally visible English-language journals.

Table 3. Overview of publications found in OpenAlex. Breakdown by publication type in VABB.

Publication type	Found in OpenAlex
Journal article	63,560 (67.6%)
Book chapter	6,117 (17.83%)
Proceedings paper	3,124 (46.89%)

Book as editor	629 (12.12%)
Book as author (monograph)	591 (9.1%)
Total	74,021 (50.46%)

Coverage of non-English language sources is an ongoing concern for the social sciences and humanities. The multilingual nature of the VABB database allows us to investigate the coverage in OpenAlex for sources in languages other than English, which is of particular importance as it would constitute an advantage over other international data sources. Table 4 shows the coverage of sources in the most frequent publication languages in VABB. English publications are covered best, whereas Dutch sources are covered only about 8%. This suggests that OpenAlex does not cover Dutch language VABB publications very well.

Table 4. Overview of publications found in OpenAlex. Breakdown by language.

Language	Found in OpenAlex
English	68,819 (70.91%)
Dutch	3,087 (8.05%)
French	911 (15.94%)
other	549 (22.07%)
Spanish	356 (28.03%)
German	299 (16.35%)
Total	74,021 (50.46%)

This is of course partly related to the more limited DOI coverage for Dutch-language publications in general and the relatively higher share of book publications (book chapters, monographs and edited volumes) in Dutch language publications. Only 1,452 out of 38,334 Dutch language publications have a DOI associated with them in the VABB database.

The availability of references and affiliation information

Apart from coverage in OpenAlex, we are also interested in the availability of metadata. For bibliometric studies, the availability of metadata is of crucial importance. A quick note on the way in which OpenAlex deals with records is warranted here. OpenAlex is envisioned as a graph connecting different entities. Each of the different entities in the graph is accorded a unique identifier. There are works, authors, venues and institutions. These entities are connected to each other. OpenAlex does not record references to 'non-source' items. All references recorded also refer to a work entity in OpenAlex. In terms of metadata about institutions, OpenAlex assigns a ROR identifier to all institutions. This is a useful addition because it makes it easier to link the institutions to other datasets. OpenAlex also attaches considerable importance to ORCIDs. In previous studies it has been noted that OpenAlex makes more ORCIDs available than other bibliographic sources (Alonso-Alvarez and Van Eck 2024; Culbert et al. 2024). We study two aspects of metadata coverage: references and affiliation information. In terms of the coverage

of references we look at the number of publications that have at least one reference and at the median number of references per publication.

In total, 63,518 publications matched with OpenAlex include at least one reference, this is 86 percent of records. In table 5 we show the inclusion of references broken down by publications that are also covered in the Web of Science, publications that are peer-reviewed, publications that are non-peer-reviewed and publications that are not included because of technical issues. While reference coverage is high for publications that are also indexed in the Web of Science, there are a large number of publications with zero references for the other parts of the database. For the non-peer-reviewed publications and publications excluded for technical reasons, this may be a reflection of the nature of those publications, which may include grey literature, short reviews and editorial material. For the peer reviewed publications approved by the GP, the number of publications may be due in part to the publishers not providing access to the reference information, making it harder for references to be included in Open Alex. In terms of the median number of references per publication, we note the relatively high values for publications that are also indexed in the Web of Science.

Table 5. Number and percentage of publications in Open Alex that include at least one reference. Median number of references for records that have at least one reference.

VABB part	Includes references in OpenAlex	Median number of references (for records with references in OpenAlex)
Web of Science	50,169 (95.9%)	44
Other peer-reviewed (GP)	11,945 (66.53%)	27
Non-peer reviewed	847 (36.26%)	19
Technical issue	557 (39.34%)	12
Total	63,518 (85.81%)	41

The affiliation fields gathered from OpenAlex are the following: raw affiliation string, institution name, institution id, ROR identifier, country code (of the institution) and ORCID (of the author). As mentioned, OpenAlex assigns ROR identifiers to all affiliation instances. Affiliation information is completely missing for 6,432 publications (or 8.7 percent of records). For the other publications, there is at least some affiliation information present. We show the number of complete or missing fields per author in table 6.

Table 6. Number and percentage of publications in OpenAlex that have missing affiliation information.

Data field	# Publications missing entries (total)	Missing entries – WoS part	Missing entries – peer-reviewed (GP) part	Missing entries – non-peer-reviewed part	Missing entries – technical issue
ORCID	36,553 (49.4%)	25,386 (48.53%)	9,227 (51.39%)	1,200 (51.37%)	740 (52.26%)
Country	6,500 (8.8%)	1,973 (3.77%)	3,954 (22.02%)	320 (13.70%)	253 (17.87%)
Institution ID	6,432 (8.7%)	1,915 (3.66%)	3,946 (21.98%)	318 (13.61%)	253 (17.87%)
ROR ID	6,432 (8.7%)	1,915 (3.66%)	3,946 (21.98 %)	318 (13.61%)	253 (17.87%)
Institution name	6,432 (8.7%)	1,915 (3.66%)	3,946 (21.98%)	218 (13.61%)	253 (17.87%)

OpenAlex includes ORCIDs, although ORCID identifiers are not available for all authors. The reason for this is twofold. Not all researchers have ORCID profiles, and it is not always straightforward to link ORCIDs to researchers. Furthermore, OpenAlex links each institution to a ROR ID (which is why ROR IDs are available for most affiliation instances). However, it is not clear whether each of these links are accurate. Breaking down by peer-review status and indexation in the Web of Science, we can see that affiliation information is more available for publications that are also indexed in the Web of Science, and more likely to be missing for publications that are not. This is in line with other studies on metadata completeness in OpenAlex. Metadata is more available for journal articles and less for books and other publication types. These numbers do not give an indication of the quality of the metadata, which relies in large part on the performance of the disambiguation algorithms used by OpenAlex that connect authors to ORCID profiles and affiliation information to ROR identifiers.

Discussion

From the records found in OpenAlex we can gather that OpenAlex does include additional publications that are part of the VABB database but not covered in the Web of Science, but does not come close to covering all peer-reviewed publications in VABB. More specifically, Dutch language publications are not covered well and non-journal articles are also not covered well. There are some reasons for why this might be the case. The most successful way in which we were able to match publications across databases was through DOI. OpenAlex covers records with DOI quite well. This is probably due to the way in which records are added to the database. Crossref is one of the main sources of OpenAlex and is also one the main DOI

registration agencies[5]. Records with DOI are more easily traceable and identifiable online. However, many publications do not have a DOI. In particular, books are frequently not assigned a DOI and many (local) journals similarly do not regularly assign DOIs. This is due to several reasons, including the fact that registering a DOI is not free of charge. A recent conference contribution of the coverage of publications from the CRISTIN database came to similar conclusions with regards to the inclusion of books and publications without DOI (Armitage and Seland 2024). This is important to keep in mind as studies may rely solely on DOI to match with OpenAlex. In terms of the coverage of publications that are also covered in the Web of Science, OpenAlex covers a large majority of publications. The publications that were not found in this way could be due to several reasons, including incomplete or inaccurate data in one of the sources (missing DOI in VABB, title variations, etc). The coverage of records that are not considered peer-reviewed in VABB is lower, which is understandable considering OpenAlex's focus on research publications.

An overview of the metadata covered in OpenAlex gives us insight into its potential usefulness to enrich our local database and use for the purposes of bibliometric research. The VABB database does not include reference information, meaning that a citation analysis of the Flemish SSH needs to rely on additional data sources. While OpenAlex does not offer broad coverage of non-English language SSH literature, it offers more comprehensive coverage than the Web of Science.

Limitations

First, we have to note that our search strategies do not exhaust all of the possible ways in which records could be matched with OpenAlex. Alternative approaches could focus on ISSN coverage as a proxy, allow for errors in author names etc. We have tried here to use an approach that could potentially be replicated with other publication databases.

Second, it is possible that there are publications in VABB that have a DOI that is not in our database. The VABB records as many DOIs as possible, and DOIs are frequently added as part of the data enrichment process, but universities are not required to add DOIs to publications they submit to VABB, which means that DOI coverage in VABB is not complete.

Third, we should note that OpenAlex is changing rapidly. We have used a snapshot of October 2024, but it is possible that by the time of the conference, the results of this exercise may differ.

Fourth, our results with regards to metadata only include whether or not a field was available for a particular record. Our results do not provide evidence to the quality or accuracy of the metadata included. Additional research could look further into the quality of references and affiliation information.

---

[5] https://www.crossref.org/

Conclusion

We have matched records from the local bibliographic database VABB with OpenAlex and reported on the results of matching with several search strategies and the coverage of OpenAlex across language, publication type, indexation in Web of Science and peer review status. Our main conclusions are that OpenAlex provides good coverage of publications with DOI, which means that it covers the parts of the local database that have a DOI (mainly journal articles and publications in English). This also means that coverage for books, and publications in languages other than English is low. In terms of metadata, OpenAlex provides most metadata for records that are also found in the Web of Science, but also includes metadata for many of the records that are not included in the Web of Science.

From the perspective of open data, the high number of references available in OpenAlex is an exciting possibility to use open and non-proprietary data.

We think the results of this research could be of interest to the bibliometric community, the community around OpenAlex and also local publishers who would like to increase the international visibility of their scholarly publications in OpenAlex. International bibliographic databases usually perform worse for the SSH and for publications in languages other than English, which poses difficulties for bibliometricians interested in those fields. While we could retrieve more publications from the comprehensive regional database in OpenAlex than in Web of Science, there is still a large number of publications that were not found in OpenAlex. While some of these discrepancies could be explained by the obscurity of the material (publications that are not strictly scholarly), many of the publications are peer-reviewed scholarly materials. Improvements to OpenAlex could include making searches based on ISBN easier and attempting to include more book publications. For local publishers, we think these results show that registering DOIs increases visibility in OpenAlex. Coverage of non-English language sources will improve if more records are assigned a DOI. Alternatively, adding the records from VABB (and by extension other national bibliographic databases) to OpenAlex could be an interesting way forward. Adding VABB data to OpenAlex would increase the visibility of the Flemish SSH. For bibliometricians, our results indicate that caution is warranted when performing bibliometric studies focusing on the SSH with OpenAlex. Coverage of non-English language sources and book publications is still relatively low, even if it is higher than for alternative sources (notably the Web of Science). OpenAlex is, at this point, a valuable source to enrich the local database, but it is not at the level of replacing it.

Acknowledgements

References

Alonso-Álvarez, P., & van Eck, N. J. (2024). *Coverage and metadata availability of African publications in OpenAlex: A comparative analysis*. Zenodo. https://doi.org/10.5281/zenodo.14006425

Alperin, J. P., Jason Portenoy, Demes, K., Larivière, V., & Haustein, S. (2024). *An analysis of the suitability of OpenAlex for bibliometric analyses* (arXiv:2404.17663). arXiv, 26 april 2024. https://doi.org/10.48550/arXiv.2404.17663

Armitage, C., & Seland E. H. (2024). *NWB2024: Will OpenAlex Solve Our Problems? A Coverage and Metadata Comparison with Cristin* [Conference presentation], 16 december 2024. https://doi.org/10.6084/m9.figshare.28033397.v1.

Aspeslagh, P., Guns, R., & Engels, T. C. E. (2024). *VABB-SHW: Dataset of Flemish Academic Bibliography for the Social Sciences and Humanities (edition 14)* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.14214806

Céspedes, L., Kozlowski, D., Pradier, C., Holmberg Sainte-Marie, M. Shokida, N. S., Benz, P., et al. (2025). Evaluating the linguistic coverage of OpenAlex: An assessment of metadata accuracy and completeness'. *Journal of the Association for Information Science and Technology*, 1-12. https://doi.org/10.1002/asi.24979.

Culbert, J., Hobert, A., Jahn, N., Haupka, N., Schmidt, M., Donner, P., & Mayr, P. (2024). *Reference Coverage Analysis of OpenAlex compared to Web of Science and Scopus* (arXiv:2401.16359). arXiv. https://doi.org/10.48550/arXiv.2401.16359

Delgado-Quirós, L., & Ortega, J. L. (2024). Completeness degree of publication metadata in eight free-access scholarly databases. *Quantitative Science Studies*, 5(1), 31–49. https://doi.org/10.1162/qss_a_00286

Giménez-Toledo, E. (2020). Why books are important in the scholarly communication system in social sciences and humanities. *Scholarly Assessment Reports,* 2(1), 6. https://doi.org/10.29024/sar.14.

Jiao, C., Li, K., & Fang, Z. (2023). How are exclusively data journals indexed in major scholarly databases? An examination of four databases'. *Scientific Data*. 10(1), 737. https://doi.org/10.1038/s41597-023-02625-x.

Kramer, B. , Neylon, C and Waltman, L. (2024). *Barcelona Declaration on Open Research Information*, Zenodo. https://zenodo.org/records/10958522.

Kulczycki, E., Guns, R., Pölönen, J., Engels, T.C.E, Rozkosz, E., Zuccala, A., Bruun, K. et al. (2020) Multilingual Publishing in the Social Sciences and Humanities: A Seven-Country European Study. *Journal of the Association for Information Science and Technology* 71(11): 1371-1385. https://doi.org/10.1002/asi.24336.

Maddi, A., Maisonobe, M., & Zeghmouri, C.B. (2024) *Geographical and Disciplinary Coverage of Open Access Journals: OpenAlex, Scopus and WoS*. OSF. https://doi.org/10.31235/osf.io/8wa4q.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. https://doi.org/10.1007/s11192-015-1765-5

Scheidsteger, T., & Haunschild, R. (2023). Which of the metadata with relevance for bibliometrics are the same and which are different when switching from Microsoft Academic Graph to OpenAlex? *Profesional de La Información*, 32(2), Article 2. https://doi.org/10.3145/epi.2023.mar.09

Zhang, L., Cao, Z., Shang, Y., Sivertsen, G., & Huang, Y. (2024). Missing institutions in OpenAlex: Possible reasons, implications, and solutions. *Scientometrics* 129. 5869-5891 https://doi.org/10.1007/s11192-023-04923-y