Abstract for CRIS2018 conference

Comparing bibliometric tools for research assessment from a CRIS point of view

Ellen Fest, Hilde van Zeeland, Jorik Booij, Theo Jetten and Peter van der Togt (Wageningen University & Research Library)

Introduction

For many years already, the CRIS has played a central role in the research assessment at Wageningen University & Research (WUR) (Van Veller et al., 2010). However, the currently used methodology, based on the Essential Science Indicators (ESI) (Clarivate Analytics), has some major drawbacks. The main limitation is the normalisation of the indicators to 22 very broad categories from ESI. For this reason, an evaluation of alternative bibliometric tools was necessary. WUR Library maintains the CRIS, which is of high quality (Fonderman & Van der Togt, 2016) and therefore a comparison of systems could be done relatively easy. In this study, we look at the methodology of the current system (Staff Publications) and compare it to three available commercial systems (InCites, SciVal, and Dimensions) and what data and quality checks are needed in the CRIS to easily get citation data from these commercial systems.

Method

The four evaluated systems are Staff Publications, InCites, SciVal, and the recently launched Dimensions. We compare the coverage of the underlying databases, the total amount of citations, and the field normalised citation score. The latter varies slightly in the methodology of the various systems.

The study was performed on research group level and the selection of these groups is based on: 1) the type of research field in relation to the broad ESI-categories; 2) low coverage of scientific articles by Web of Science; and 3) groups with a high (>3) or a lower Relative Impact (≤1) (Van Veller et al., 2010).

The groups' publications are taken from WUR's CRIS, and matched in the various tools by respectively their WOS-identifier, Scopus EID or DOI.

Needs for a CRIS

Most bibliometric tools ensure a link with the available CRIS system. For example, publication sets can be exported from Pure to SciVal, and SciVal metrics can be added to Pure. Since WUR Library is calculating the bibliometric indicators manually, publications must be assigned a Web of Science identifier. We use the API to find the right identifier, which works best when a DOI is available.

However, the Web of Science data must be of the same quality and enriched with the right DOI. This makes a check of the metadata of the found identifiers inevitable.

The same holds for automatically assigned identifiers by CRIS systems. We found that strange errors occurred with assigning Scopus IDs by an algorithm in Pure. We developed checks to ensure that correct Scopus IDs were linked to the publications in our CRIS.

In Dimensions, the DOI is the main identifier. The quality and availability of DOIs in the CRIS-system is therefore a serious requirement.

Differences between the bibliometric tools

In the presentation, we will show the between the various systems in terms of coverage, the total amount of citations, and the field normalised citation score of refereed articles from Wageningen University & Research.

Although Scopus indexes more journals than Web of Science does, the coverage of SciVal is not necessarily the highest. This is due to a difference in the assigning of document types to scientific articles. Scopus classifies conference papers published in refereed journals as a separate category called 'conference papers', while Staff Publications (Pure) cannot differentiate between these two. With the new tool, Dimensions, we are curious what the coverage will be. Results from the relative new tool Dimensions are not yet available, as this is work in progress.

Groups from the social sciences and groups in applied sciences (WUR's research institutes) profit from a change from Web of Science as a citation database to Scopus, resulting in better coverage of the scientific articles. For some groups, Staff Publications results in higher Mean Normalized Citation Scores (Relative Impact) and for other groups, SciVal results in higher scores. Again, groups from the social sciences and groups in applied sciences benefit from a change in system from Staff Publications to SciVal. Groups in more fundamental sciences score lower in systems with more differentiated categories such as SciVal and InCites. This can be explained by the distribution of the publications over more and smaller categories with higher citation baselines compared to ESI-baselines.

It is interesting to compare our own method based on the Essential Science Indicators to SciVal and Dimensions, because the normalisation of the citations is done in three different ways. Our own method normalises the citations by very broad journal based research fields. SciVal uses smaller research fields, but fixes the value after three years. Dimensions uses a method based on article-level subject classification. The question to be answered is therefore: what happens with the performance of a group using these three systems?

Reference

Van Veller, M. G. P., Gerritsma, W., van der Togt, P. L., Leon, C. D., & van Zeist, C. M. (2010). Bibliometric analyses on repository contents for the evaluation of research at Wageningen UR. In A.

Katsirikou & C. H. Skiadas (Eds.), *Qualitative and Quantitative Methods in Libraries: Theory and Applications* (pp. 19-26). http://edepot.wur.nl/7266.

Fondermann, P., & van der Togt, P. L. (2017). How Wageningen University and Research Centre managed to influence researchers publishing behaviour towards more quality, impact and visibility. Procedia Computer Science, 106, 204-211.