

Ranking the Open Access-ibility of Universities

Stian Håklev, shaklev@gmail.com, University of Toronto

University rankings and league tables, although controversial, always attract much attention from the institutions themselves and those affiliated, as well as other stakeholders. Given the amount of attention and possible impact such rankings can create, there has been a number of attempts at creating rankings that promote certain causes, whether it is environmental friendliness, most wired campus or most vegan friendly.

This project aimed to explore whether it would be feasible to create a ranking of universities in terms of their Open Access-ibility, and how it would be constructed.

Through a survey of the existing literature on open access and university rankings, this paper proposes a simple model of constructing a ranking that will encourage universities to strive for an increase in their Open Access publications.

League tables

University rankings in the form of league tables is a relatively new phenomenon that has caught on explosively. The first was probably published by US News & World Report in 1992, and after that many countries began to publish their own national rankings, based on different criteria and indicators. Global rankings were pioneered by Shanghai Jiaotong University in 2003 with their Academic Ranking of World Universities (ARWU), and later THES joined the fray. Both rankings (as well as national and regional ones) have had a large impact, and been roundly discussed and criticized for choice of indicators, and even the futility of holistic rankings.

In addition to holistic rankings, there has been a large amount of specialized university rankings, either for novelty (Princeton Review ranks universities in 42 tables, based on indicators such as marijuana consumption and cafeteria food), or to promote a specific agenda. Even without a scientific methodology, such rankings can create significant publicity, as is the case for the PETA ranking which lists University of Toronto as the most vegan friendly campus in North America.

There are also examples of scientifically sound rankings that aim to both put more focus on aspects of academia that are perceived to be neglected, and also to collect best cases of institutional and academic leadership. The German Hochschulranking nach Gleichstellungsaspekte which ranks German universities from a gender perspective is a case in point. Another ranking produced by researchers at St. Gallen University in Switzerland which looks at the lack of sustainable development components of business schools, is based on the belief that "To reveal and overcome those shortcomings as well as to identify best practise, ranking schemes are considered to be effective means."

We believe that producing a simple and transparent metric of the "Open Access-ibility" of an institution, can provide useful data, but more important, work as a strong encouragement for institutions to promote OA.

Conclusion

We choose to measure **output**, because that way inputs work as independent variables, and we can test the efficacy of for example mandates. We choose a **simple minimum definition of OA** "the article is available online for free", because it is easier to measure, there is wide acceptance of this requirement, and it provides for an easy-to-understand number. We choose to **construct a number as a percentage of publications that is OA**, because it is easy to understand, and does not penalize smaller institutions. We suggest a way of datamining, which relies on publication lists produced by institutions. This has to be tested in practice to test for feasibility.

We believe that producing a simple and transparent metric of the "Open Access-ibility" of an institution, can provide useful data, but more important, work as a strong encouragement for institutions to promote OA.

Three kinds of indicators

Input-based

Input-based indicators measure "what goes in". This is commonly used in more general rankings which measure for example amount of dollars spent per student and the size of libraries and class rooms. An openness index equivalent might be to look at funding of staff assigned to promote and facilitate open access publications, percentage of library budget reserved to pay for author fees, or existence and funding of an institutional repository.

Process-based

Look at what goes on within the black box, rare in university rankings, but common in quality assurance programs such as ISO 9002. There are several possible process-based indicators for openness, with an example being whether the unit has policy of obligatory open access for publications.

Output-based indicators

This indicator is by far the most commonly used in rankings that study research production, such as the Shanghai Jiaotong, because there are good bibliometric measures and infrastructure to measure academic output.

"To reveal and overcome [...] shortcomings as well as to identify best practise, ranking schemes are considered to be effective means."

Delimiting Open Access

There is not a binary opposition, but a continuum, between the poles of open publications on one side and closed or proprietary publications on the other side. Peter Suber (2004a) differentiates between price barriers and permission barriers: all advocates of open access require that the material will be free to access, removing the price barrier. But many open access articles are still under normal copyright.

Questions:

What should be the minimal definition of Open Access
Should we use a gradation, or an either-or test

Data gathering

We want to ascertain the percentage of all journal articles published by a certain institution, that are open access. We do not concern ourselves with the impact factor of the journals the articles are published in. We are hoping to make the process scaleable, to enable the inclusion and comparison of many different institutions. If our ultimate outcome is a percentage, or a ratio, then we need a part and a whole.

The whole

The whole is the amount of peer-reviewed journal articles published at a certain institution, and the part is the amount of those articles that are open access. As mentioned above, the whole could possibly be obtained by using citation indexes and text mining, but in cases where it is possible, we will prefer obtaining publication reports from the institutions themselves - as this is data that is commonly kept and made available publicly.

The part

To be "de facto" OA, articles need to be "findable", thus a semi-automatic solution seems possible. It should be possible to construct a program that parses a publication list from a given institution, and searches for the articles listed therein, in a number of repositories. These repositories could include both Google Scholar, as well as Directory of Open Access Journals (DOAJ), and possibly others that index open access scholarship. This robot should be able to ascertain whether it can find full-text articles for a given publication.

Acknowledgments

This project was conducted as an independent study under Professor Leslie Chan at University of Toronto at Scarborough. The research is part of the project "Institutionalizing Openness" at Knowledge, Media and Design Institute, University of Toronto.

