

Web, webometrics and the ranking of universities

Isidro F. Aguillo

Head of the Cybermetrics Lab, National Research Council (CSIC), Madrid, Spain. Editor of the Ranking Web of World Universities

isidro.aguillo@cchs.csic.es

Feasibility, a major problem in rankings

The evaluation of complex institutions is really a difficult task. For describing adequately such heterogeneous and diverse organizations like universities worldwide you need to take into account a large number of variables. A good ranking design should consider not only all missions involved, but all the activities performed and the results of these actions and the impact internally and externally.

From a statistical point of view, you can design a combined indicator with a large number of variables and estimate the weighting of each variable from empirical data using sophisticated non parametric multivariate analysis.

The problem is this approach is not feasible for a global ranking. For most of the Higher Education Institutions (HEIs) there are no reliable sources of data, but even obtaining comparable figures is also difficult as there are no standards for even the most basic indicators. Some of the variables have a limited range, so its discriminative value is only valid for the top organizations (only a few organizations have 2 or more Nobel Prizes).

Web as a mirror

World Wide Web offers an interesting alternative. The relative contribution of each mission and activity of a university to its electronic presence is difficult to estimate, but such presence reflects the overall performance of the organization and can be measured.

The central hypothesis is that today the university websites are no longer a brochure for introducing the institution, but its main platform for describing and communicating the university's activities. Moreover the link analysis allows estimating the visibility and impact of these contents as they are measuring a large virtual referendum of the external websites that "cite" a university webdomain. The success of Google's Pagerank algorithm, a webometric indicator, suggests this is a good tool for achieving the goal of evaluating performance and productivity.

Besides, as search engines are global intermediaries, there is an actual possibility of ranking not only a fraction of the number of institutions but the whole population of the HEIs having their own webdomain. There are biases in the coverage of searchers' databases and some methodological problems but using several search engines has been proved an effective approach, so more than 17,000 domains are currently scanned for building the Ranking.

The Ranking Web

For testing the hypothesis, a large effort was done to obtain basic web indicators from the main commercial search engines for all the universities webdomains. Since 2004 the methodology has evolved in order to reinforce the weight of the academic contents but the global picture obtained basically has not changed. The Ranking Web of Universities (www.webometrics.info) classified the HEI as you can expect with the prestigious ones in the leader positions in both the world and regional rankings. The correlation with other global rankings (Shanghai ARWU or Quacquarelli THE-QS) is very high and many of the outliers can be easily explained.

The webometrics indicators considered are the total size of electronic publications in the university main webdomain and the visibility of such webpages according to the number of external inlinks they received from other websites. The key point of the model is both variables maintain a

ratio 1:1 (size:impact) and other aspects like design, accessibility or specially popularity (number of visits and visitors) are excluded. In order to increase the weighting of “quality” academic material, the size component is splitted in three variables: number of webpages (20%), number of documents in rich files formats: pdf, doc, ps & ppt (15%) and number of papers in the citation database Google Scholar (15%). Obviously the other 50% according to the model is assigned to the number of “sitations” (site link citations).

Bad web practices

In many cases the reason for delayed positions is bad web practices. There are examples of Universities changing their webdomains, but maintaining older ones or even organizations with two or more webdomains. These practices not only penalize their web ranks but most important decrease the visibility of their contents in the search engines.

This a leadership problem that also explain the fact that the web performance of an institution could be below the expected position according to their academic excellence, as probably university authorities are not enforcing an adequate web policy, or promoting the electronic publication.

Academic Digital Divide

The main advantage of the Ranking Web is the coverage of developing countries. It is almost the only source for rankings of universities beyond the 500th position. Currently a list of 6000 HEIs is published where the rank of many Latin American, African and South and South-East Asia are provided. Large public institutions in these countries, especially those heavily involved in distance learning initiatives, achieve good positions in the Ranking. For students moving regionally this is very valuable information and it explains that our website receives close to 4 million visitors per year.

A large number of small, usually private, colleges in these countries have no web presence at all or they have a very limited one, consisting only of a few dozens of pages. Although it is not surprising, the widespread availability of free or cheap web hosting services and the easiness of web edition and publication suggest that extra effort is needed for convince them of taking the Web seriously.

However the most surprising result refers to developed countries. Consistently and contrary to other rankings results, the webometrics one shows a striking academic digital divide between North American universities and their European counterparts. For the Top 200 there are twice more US & Canada universities than European ones, a situation that disappear in the Top 500 where there is a virtual tie.

In a globalised world, with a stronger competition for high level scholars, excellent students or big funding, this divide is very concerning as the web is an important information source for taking decisions and Europe is lagging behind.

Taking the Web seriously

The new point of view of the Ranking Web is already providing exciting results and it can be as a complement to other global initiatives. Probably national rankings could provide more detailed insights but webometrics is already offering some advantages: better recognition of the importance of distance learning, measuring the commitment to Open Access initiatives, improving the evaluation of technologies, social sciences and humanities...

The Ranking Web is published two times (January and July) per year and currently it is part of a series of similar rankings covering Hospitals, Research Centers, Business Schools and Repositories.