

Web Indicators for the World Universities: A comparative analysis

Isidro F. Aguillo; Victor M. Pareja; José A. Prieto; Gloria Vegue; Natalia Arroyo
CINDOC-CSIC
Joaquin Costa, 22
Madrid. Spain
isidro@cindoc.csic.es
Ph: +34-91-5635482
Fx: +34-91-5642644

Introduction.

During last years a great effort has been made for providing both a theoretical basic for the cybermetric analysys as well as enough empirical data for providing a global escenario of R&D in the Web. New automatic agents are now available as more stable search engines so new techniques are faster, powerful and reliable. The aim of this paper is to provide for the first time a very large population of web indicators for the whole of world universities. The objective is to show some basic and descriptive analysis that can be use in future work in combination with other traditional bibliometric and scientometric data.

Methods.

A series of web indicators is proposed including definitions, the problems faced for obtaining them, several proposed applications and the shortcomings of their use and visualisation. We built the following 22 different indicators:

Number of hosts	By country/supranational
	By domain/subdomain
	By institutional domain
Websize (number of webpages)	By country/supranational
	By domain
	By (academic) subdomain
	By institutional domain
Websize (total number of objects)	By website
Websize (total number, size in Mb and type of mulytimedia objects)	By website
Websize (number and type of rich files)	By (academic) subdomain
	By institutional domain
	By website
Number of insitutional domains	By domain or subdomain
Number of websites (virtual or physical)	By institutional domain
Depth (number of levels)	By website
Site quality (Search engine friendliness)	By website (secondarily by institutional domain)

Density (total hypertext links per page)	By website
Luminosity (different, not repeated, outlinks per page)	By website
Connectivity (different internal and external inlinks plus external outlinks)	By website
Link frequency (ratio between total number of outlinks by the number of different outlinks)	By website
Navigability (total, including repeated, internal outlinks per page)	By website
Endogamy (percentage between different internal outlinks and total number of different outlinks)	By website
Link quality (percentage of valid hypertext links)	By website
Outlink dispersion (by domain or subdomain)	By institutional domain
	By website
Popularity (relative position)	By website
Visibility (number of external inlinks)	By institutional domain
	By website
Impact (WebIF)	By institutional domain
	By website
Inlinks diversity (linguistic)	By institutional domain
	By website
Inlinks diversity (domain/subdomain)	By institutional domain
	By website

The URL addresses were identified for the universities of all countries in Europe, America and Australia/NZ and a representative sample of Asian and African ones. The sources of quantitative data has been major search engines that support field delimitation and the Alexa database. Some scripts were developed for automatic extraction of numeric data from these websites.

Results.

Over 5070 universities with web presence were identified in 73 different countries. Using Google and Altavista data, websize, visibility and impact indicators were calculated for each university. Popularity relative ranking were built using Alexa listings. The results were presented in different levels of geographical aggregation and listings of institutions with the highest values are showed.

Conclusions.

The sample analysed probably represents more than 90% of the world universities. The results show a correlation among some indicators and perceived scientific competence, productivity and impact of the universities involved. Open universities are specially favoured by the popularity indicator while visibility offer better performance than Web impact under certain limits. Rich files are more ramdonly distributed, but a general pattern show a correlation with productivity of formal papers in hard sciences oriented universities.