

فعالية مواقع ويب المعاهد العليا في مصر

محمد عبدالمولى

ملخص: تهدف هذه الدراسة إلى تقييم مدى فعالية الوجود على الويب لمواقع المعاهد العليا في مصر، وتجب الدراسة عن الأسئلة الآتية: كم معهداً عالياً في مصر لديه موقع ويب على الإنترنت؟ إلى أي مدى يتم اكتشاف مواقع المعاهد العليا في مصر؟ أيقوم مديرو مواقع المعاهد العليا باختيار صيغة مفضلة لعنوان الموقع canonical optimization أم لا؟ ما النطاق العلوي المفضل استخدامه مع مواقع المعاهد العليا في مصر؟ ما محرك البحث الأكثر تكشيفاً لصفحات مواقع المعاهد العليا في مصر؟ ما ترتيب مواقع المعاهد العليا في مصر وفقاً لقيم معامل تأثير الويب في صيغته العادية والمعدلة؟ وأخيراً هل هناك ارتباط بين معامل تأثير الويب في صيغته الأساسية والمعدلة؟ واعتمدت الدراسة على عينة من 65 موقعاً من مواقع المعاهد العليا في مصر، وأوضحت النتائج أن هناك 65 معهداً عالياً من أصل 153 في مصر لديها موقع على الويب، وفيما يتعلق بمدى التكتشف في محركات البحث حصل معهد العلوم الإدارية ببني سويف على المركز الأول (عند استخدام بيانات عناوين المواقع www.edu.eg)، ومعهد الجيزة للهندسة (عند استخدام بيانات عناوين المواقع بدون إضافة www.edu.eg)، النطاق العلوي الأكثر استخداماً هو edu.eg، محرك البحث الأكثر تكشيفاً لصفحات مواقع المعاهد العليا في مصر هو جوجل، أما ترتيب مواقع المعاهد العليا وفقاً لقيم معامل تأثير الويب فحصل معهد مصر القديمة للإدارة والسكرتارية على المركز الأول، وأخيراً أوضحت البيانات التي حلت أنه ليست هناك علاقة بين قيم معامل تأثير الويب في صيغته الأساسية والمعدلة.

الكلمات الدالة: الفعالية على الويب، الوجود على الويب، قياسات الويب، ويبوميترىكس، ويبوميترىكس، معامل تأثير الويب، المعاهد العليا، مواقع الويب، مصر.

The Web Presence of Higher Institutions in Egypt: A Web Metric Study

*Mohamed Abdelmawla**

Abstract: This study aims to evaluate critically the web presence efficiency of higher institutions' websites in Egypt. This main objective will be investigated through the next questions: how many higher institutions in Egypt have a website on the internet, to what extent higher institutions in Egypt are indexed in web search engines, do website administrators do a canonical optimization or not, what is the preferred Top Level Domains (TLDs) for higher institutions' websites in Egypt, what is the most indexing search engine for the higher institutions' websites in Egypt, what is the rank of higher institutions in Egypt according to Web Impact Factor (WIF) and Revised Web Impact Factor (RWIF) and is there a correlation between WIF and RWIF for higher institutions' websites in Egypt. The study sample consists of 65 higher institutions' websites. The result revealed that 65 of 153 higher institutions had a website (42%). According to visibility on web search engines Higher Institution of Administrative Sciences - Bani Sweif got the first place in the top ten places (Using URLs** with www) and Giza Engineering Institute (Using URLs without www). Higher institutions' website administrators did not do a Canonical Optimization. The most popular targeted TLD for the higher institutions developers was edu.eg. The most indexing search engine for the webpages whether under www domain or under root domain was google. After sorting gathered data according to WIFs values descending; Higher Institution of Management and Secretariat - Masr Kadema obtained the first place in the top ten places. Moreover, the data analyzed indicates that there was not a correlation between WIF & RWIF.

Key words: Web presence, Webometrics, Web Impact Factor, Higher Institutions, websites, Egypt.

* Libraries, Documentation & Information Technology Department, Arts Faculty, Cairo University, Egypt. mm_abdelmawla@yahoo.com

** URL (Uniform Resource Locator).

Introduction

The study of the quantitative aspects of web / website and web related phenomena is known as Webometrics. Almind and Ingwersen in 1997 introduced the term 'Webometrics'. It tries to measure the World Wide Web to get knowledge about the number and types of hyperlinks, the structure of the World Wide Web and usage patterns. According to that, webometrics is helping to recognize patterns of interlinking among web creators which is very alike to bibliometrics. The main tasks of hyperlink analysis are to extract, discover and link together from vast amounts of data sources to represent and evaluate the significance of the related evidences and to learn patterns to guide the extraction, discovery and linkage of entities. Hyperlink analysis helps to find out the nature of web impact of academic websites. Ingwersen in 1998 introduced Web Impact Factor (WIF) and how to calculate it for websites. The calculation of impact factors is based on average link frequencies. It has been considered one of the quantitative indicators. The WIF was developed by adopting Journal Impact Factor (JIF) by applying citation techniques to hyperlinks on the Web. There are various types of WIFs, such as Simple WIF, Self-Link WIF, External-Link WIF and Revised Link WIF. Ranking of websites is calculated on the basis of these WIFs (Prabakaran, 2014). Details about WIFs types are discussed in datasets and methodology section.

As The World Wide Web has now become one of the main sources of information about academic and research activities. therefore, it is an excellent platform to test new methods of evaluating webometric activities (Noruzi, 2005). This leads to consider the academic websites in many countries as large multifaceted communication terminals and are increasingly used for a wide variety of purposes like: announcing library activities, providing online library catalogues, and announcing the achievements of individuals, research groups, institutions and departments. They can also disseminate findings either by hosting online articles or by publishing summaries, data sets and tools. The pages themselves can be created centrally by administrators or webmasters or locally by individuals for themselves or their research team or projects. Potential benefits of an effective web presence include greater research impact, attracting students, media interest and commercial contacts. In this context it is logical to investigate measures of the effectiveness of websites

to study the communication activities that they represent and to build useful evaluation metrics (Vaughan & Thelwall, 2005).

Based on the above-mentioned information it was noticed that higher institutions in Egypt did not get enough attention whether from the government or from researchers. As most attention goes to universities sector, this study tries to give more attention to the higher institutions sector by evaluating its web presence. It should be taken into account that higher education in Egypt is divided into two main sectors: the governmental and private universities and the governmental and private institutions which give a lower academic degree than those given by universities.

Web presence could be evaluated through a lot of aspects such as the number of websites, the numbers of web pages indexed in the search engines and impact factor. As for the latter it “has been proposed as a method of gauging the influence of web sites on each other and analyzing links between institutions (Smith & Thelwall, 2002)”

Statement of the Problem

The effective presence of websites is not easy to be achieved as it requires the implementation of many procedures by web administrators. These procedures lead to make websites more effective as well as visible for search engines and the searchers of information. The researcher has noticed that some higher institutions' websites in Egypt are characterized by some features that indicate their ineffectiveness. Examples of these features are: not appearing in the results pages of search engines or appear in a late order.

Significance of the Study

Despite that many studies are conducted in the field of webometric at the global and regional levels, it was noted that the studies conducted were relatively few in the Arab world in general and at the level of Egypt in particular. As this study is concerned only with Egypt, there was only three previous studies; Ahmed Farag (Farag, 2006), Alireza Noruzi (Noruzi, 2006b) and Amgad Elgohary (Elgohary, 2008). They were concerned only with the websites of colleges and universities. Details about them will be discussed in the literature review section. Therefore, this study seeks to examine the effectiveness of the higher institutions'

websites in Egypt. Hence, the importance of this study is its quantitative and exploratory approach. Additionally, there are practical applications for this study: it helps to distinguish between good and bad practices by both webmasters and the boards of directors of major institutions in Egypt. It provides suggestions to correct the drawbacks to be taken into account.

Objectives & Questions

The primary objective of this study is to evaluate critically the web presence efficiency of higher institutions' websites in Egypt. This main objective will be investigated and achieved by answering the following questions:

- How many higher institutions in Egypt have a website?
- To what extent higher institutions' websites in Egypt are indexed in web search engines?
- Do higher institutions' websites administrators do a canonical optimization?
- What are the preferred Top Level Domains (TLDs) for higher institutions' websites in Egypt?
- What is the most indexing search engine for the higher institutions' websites in Egypt?
- What is the rank of higher institutions' websites in Egypt according to Web Impact Factor (WIF) and Revised Web Impact Factor (RWIF)?
- Is there a correlation between values of WIF and RWIF for higher institutions' websites in Egypt?

Literature Review

In recent years, there has been an increasing amount of literature on webometrics analysis in particular WIF. Most of these studies have focused on universities and have ignored higher institutions which in the academic system in the Arab world give a degree lower than that the universities give.

Generally, there were three key introductory studies. The first one was by Tomas & Peter (Almind & Ingwersen, 1997). Their study introduced the application of informetric methods to the World Wide Web (WWW).

The second study was by Peter (Ingwersen, 1998). He Introduced the

web impact factor (WIF) for the first time. This study reported the investigations into the feasibility and reliability of calculating impact factors for websites using what he called Web Impact Factors.

The third study was by Mike (Thelwall, 2001a). He provided an evaluation of Ingwersen's proposed external Web Impact Factor (WIF). In particular, he studied the case of the relationship between academic hyperlinks and research activities of British universities.

A lot of studies were followed investigating WIF under specific country code Top Level Domain (ccTLD) or under general Top Level Domain (gTLD). Mike (Thelwall, 2001b) designed a bespoke web crawler specifically for the calculation of reliable WIFs. This crawler was used to calculate WIFs for a number of UK universities. The principal findings of this study were, with certain restrictions, WIFs can be calculated reliably but did not correlate with accepted research rankings owing to the variety of material hosted on university servers.

In the study of Alastair and Mike (Smith & Thelwall, 2002) Counts of links into the websites of Australasian universities were calculated from the output of a specially designed crawler that covered universities in the UK, Australia and New Zealand. The results indicated that WIFs correlated quite well with research output for Australia but the relationship was not clear enough to be able to differentiate between the characters of the WIFs from the three different sources.

Alireza (Noruzi, 2005) investigated in his study the Web Impact Factors (WIFs) for Iranian universities and introduced a new system of measurement. Counts of links to the web sites of Iranian universities were calculated from the output of AltaVista search engine. There was a significant correlation between the proportion of English-language pages at an institution's website and the institution's backlink counts. This indicated that for linguistic reasons, Iranian (Persian-language) web sites may not attract the attention they deserve from the World Wide Web. This raises the possibility that information may be ignored due to linguistic and geographic barriers. Therefore, this should be taken into account in the development of the global Web.

Alireza (Noruzi, 2006b) in another study investigated the Web Impact Factor (WIF) for country code top-level domains (ccTLDs) of Middle-Eastern countries and sub-level domains (SLDs) related to

education and academic institutions in these countries. The results showed that the Middle-Eastern countries, apart from Turkey, Israel and Iran, have a low web presence. On the contrary, their websites had a low inlinks WIF. Specific features of websites may affect a country's Web Impact Factor. For linguistic reasons, Middle-Eastern websites (Persian, Kurdish, Turkish, Arabic, and Hebrew languages) may not receive and attract the attention that they deserve from the World Wide Web community.

Farag (2006) computed the WIF for eight universities in Saudi Arabia. The study used AltaVista to compute WIF and to provide a surface association between the WIF and productivities of the universities which were studied. The research suggested exploration on a large scale of the hyperlinks among Arab universities.

Elgohary (2008) examined the Web Impact Factor of Arab universities. The study included 99 university representing 20 Arab countries. The results indicated that Jordanian universities represented 40 percent of the top ten universities with the revised web impact factor. there was a strong correlation between external links and web presence.

Sebastian et al. (2008) described how search engines (SE) can be employed for automated and efficient data gathering for Webometric studies using well defined query specific URLs in SE. The results showed that rankings based on staff-related WIFs correlate much better with an established ranking from the Melbourne Institution than commonly used WIFs.

The next following studies were nearly similar in methodology but different in the scope. The observed note is that most of them was conducted by researchers from Asia. The majority of them concerned with academic entities in India included universities, institutions and libraries. The most significant results in these studies were found by Chakravarty & Wasan who found that there was a correlation between WIF and RWIF. These studies will be demonstrated in the next paragraphs.

Jalal, Biswas & Mukhopadhyay (2010) investigated the effectiveness and relevance of web impact factors (WIFs) for Indian universities' websites. They reviewed web impact factor and reported a case study on universities in West Bengal. SocSciBot 3.0 was used to generate link data

in order to develop micro-link topology under study. Result shows that all institutions under the National Institute of Technology (NIT) were closely related in the topology framework.

Jati (2011) explored the Web Impact Factors (WIFs) for Indonesian universities. Counts of links to the websites of Indonesian universities were calculated from the output of Yahoo search engine. The WIFs for Indonesian universities were calculated by dividing link page counts by the number of pages found in Yahoo for each university. These WIFs from the approach were then compared to study the impact, visibility, rich content and influence of Indonesian university web sites.

Mohammadamin & Fereshteh (2011) investigated the website performances of 18 Iranian research institutions based on the following factors: web impact factor (WIF), connectivity, visibility and traffic rank. AltaVista search engine was used to calculate the number of total links, WIF. data for traffic rank number of pages viewed and percentage of foreign and Iranian visitors were extracted from the Alexa data bank. Also, correspondence analysis was employed to analyze the data using John's Macintosh Project software.

Maharana, Panda & Sahoo (2012) examined and explored the web impact factor through a webometric study of 16 Indian Institutions of Technology (IIT). They identified the domain systems of the websites, analyzed the number of web pages and link pages, calculated the simple web impact factor (WIF) and calculated self link web impact factor and external web impact factor of all the Indian Institutions of Technology (IIT).

Satpathy & Maharana (2012) in their study revealed that some NIT institutions had higher number of webpages, but their link pages were very small in numbers, websites fall behind in their simple self-link and external link web impact factor. They covered only 21 NIT institutions out of 30 due to the absence of their official website.

Arif & Ismail (2013) evaluated the Public Universities in Malaysia based on the webometric perspective. Web Impact Factor (WIF) evaluation was carried out by using Majestic SEO* and Google Search

* SEO Search Engine Optimization

engine. It was found that UMS have the highest Average WIF with 13.013 followed by University Malaysia Kelantan (UMK) with 2.670 WIF and in the third place is University Pertahanan National Malaysia (UPNM) with 1.289 WIF.

Rassi et al. (2013) collected their data using webometrics formula via Yahoo search engine & Yahoo explorer during December 2010 to January 2011. The results indicated that the Institute of Food (FOI)** website got the highest impact factor among the websites research sample with WIF of 72 while the National Institute of Food and Agricultural (NIFA) with %57/41 received the highest number of inlinks having the highest visibility.

Prabakaran (2014) explored the webometric approach to library websites belongs to the observatory at national and international level. The paper aimed to investigate the observatory libraries' websites and their hyperlink analysis, web impact and ranking. The webometric research analysis has been conducted on 41 observatory library websites under the study. The analysis of the results revealed that overall External Links was higher followed by In-Links for the observatory libraries' websites.

Arif & Haroon (2015) explored the Web impact factors (WIFs) of Pakistani universities' websites. The paper discussed why RWIF is more meaningful than simple WIF. The study also attempted to rank the top-five websites of Pakistani universities' websites by considering four different website ranking systems. Also, it compared the WIFs of university websites to other developing countries as well.

Chakravarty & Wasan calculated WIF and RWIF of top ten library websites of HEIs (Higher Educational Institutes - India). They correlated both the formulas with Spearman's Rank Correlation. It was found that WIF and RWIF were correlated and associated which depicts that there was very less difference between the two ranking methods. Google search engine was chosen for the study.

Wimaladharma & Herath (2016) compared the significance of web impact of the state university websites in Sri Lanka based on the link analysis statistics obtained from well-known search engines: Google and

** they used FOI not IOF

Yahoo. In this analysis the RWIF, the ratio between the number of inlinks (external back links) and the number of web pages published in the website which are indexed by the search engines (not all pages of the website) was taken into account. The correlation coefficient between the rank of resultant Revised Web Impact Factors and Impact Factors taken from the Webometrics website derived by Cybermetrics Lab was calculated with the ninety percent of inference level. If an academic website increases its link density via Yahoo! Search engine, it is relatively significance for its Webometrics impact factor whereas Google indexing expresses less relevance for the Webometrics impact factor.

Recently there were a few studies concerned with WIF. One of them was by Kumar (2017) who analyzed the websites of Deemed universities in Andhra Pradesh and Telangana. He examined seven Deemed university websites and identifies number of webpages and link pages and calculates WIF using Google search engine. The websites were ranked based on these Webometric indicators. The study revealed that websites of Andhra Pradesh Universities got more number of web-pages but Telangana University websites got high web impact factor in most of the parameters.

Another study was by Verma (2017) who evaluated the websites of NITs of northeast region of India based on the webometric perspective. The study investigated eight NITs of northeast region through observation method and resolved that NIT Silchar of Assam has got the maximum number of webpages, highest number of Internal & External Links, highest Web Impact Factor and rank at top among the all websites of NITs.

Kokila et al (2017) calculated the Revised Web Impact Factor of university library websites in Sri Lanka, the correlations between RWIF with external backlink and page count.

Kutty (2018) analyzed the websites of Indian Council of Agricultural Research (ICAR) organizations according to the webometrics indicators. The study examined and explored the 92 ICAR organizational websites in India. It identified a number of webpages, link pages and calculated the Web Impact Factor (WIF). All websites were analyzed and data extracted using Google search engine. It suggested that Web

Impact Factors could be calculated as a way of comparing the attractiveness of web sites or domains on the Web.

Overall, these studies highlighted the need for WIF, its importance, the ways to calculate, its types, and its correlation with other variables. Also, they indicated websites ranks according the various types of WIF under a lot of various ccTLDs and/or gTLDs. The scope of these studies ranged between examining websites of universities, institutions and libraries. With regard to studies concerned with the Arabic world there were only three studies all of them were interested in universities websites only. This means that there is a gap in the research coverage and the current study seeks to fill this gap.

The Dataset and Methodology

The websites list under this study was obtained from website of Ministry of Higher Education (Education, 2017). Although, the website named the list with “private higher institutions” it was discovered that it also contained names of governmental higher institutions as well. The study dealt with this as an approved private and governmental higher institutions list by Ministry of Higher Education in Egypt. The initial list contained 153 higher institutions. There were no links to the websites. Google search engine was used to collect the website’s URLs by searching using its official name. Websites were examined carefully to be sure that it was the official website for each higher institution. There were 104 institutions had a website. Institutions with shared websites were excluded (39 websites) because they will give misleading results. The final list contained 65 websites of higher institutions in Egypt.

The final list was used to collect main study data. As google, yahoo and Bing stopped supporting “link:” command which was helping in calculating inlinks and outlinks. So, inlinks and outlinks datasets were collected from open site explorer (www.opensiteexplorer.org) free access which has in November 2017 a 190 billion URLs, 317 Million root domains, and 2 Trillion links. With regard to total web pages indexed in the search engine, the study depended on three search engines Google, Yahoo and Bing using command (site: www.example.com). Google search engine and its gathered data under URLs without www only were used to calculate WIF and RWIF because their data was the greatest one.

Data was collected twice on 22 and 23 November 2017. The first

time by using URLs with www (domains with www for example: www.website.com) and the second time by using URLs without www (domains without www for example: website.com). After gathering Data, it was saved in excel sheets and processed using statistical methods. Formulas of WIF, RWIF and correlation were applied.

Data was collected twice to investigate whether web administrators did a canonical optimization* or not. This is very helpful in avoiding the problem of splitting the inlinks on two sides by the search engines because search engines treat URLs with www and URLs without www as different websites. Search engines treat www, www1, www2,...etc. as a subdomain. Also, it deals with any domain starts with any of them as it may be owned by a different person. So, if the web administrators did a canonical optimization, the gathered data under URLs with www and under URLs without www would be equal.

There are three types of WIF: overall WIF, inlinks (revised) WIF, and selflink WIF. For the overall WIF of a web site, the numerator is the number of inlinks pages from outside the website and selflink pages within the website. For the inlinks (revised) WIF, the numerator is the number of inlinks pages counted from outside the site. For the selflink WIF, the numerator is the number of selflink pages counted from within the website. The denominators all remain the same and the number of webpages within the web site in question. The data needed to calculate the web impact factor could be coded as follow: A = total link pages (all inlinks and selflink pages), D = number of webpages published in the web site which are indexed by the search engine not all webpages available in the website. so that the formula of WIFs are as follows: overall WIF = A/D , revised WIF = $(A - (\text{self-link pages})/D$, and selflink WIF = $\text{self-link pages}/D$ (Noruzi, 2006a). This study will calculate overall WIFs and revised WIFs for the selected higher institutions' websites.

The correlation coefficient represents the relatedness of two variables and how well the value of one can be used to predict the value of the other. The correlation coefficient r ranges between -1 and +1. A

* choosing a preferred URL shape of website and make a redirect from the unfavorable URL shape using the web admin tools and website sittings.

positive r values indicate that as one variable increases, so does the other. An r of $+1$ indicates that knowing the value of one variable allows perfect prediction of the other. A negative r values indicate that as one variable increases the other variable decreases and an r of -1 indicates that knowing the value of one variable allows perfect prediction of the other. A correlation coefficient of 0 indicates no relationship between the variables (random scatter of the points) (Academy, 2017). In this study Pearson's product-moment correlation coefficient will be used. With The following formula. Where n = number of items, x = $rwif$, y = wif , and Σ = sum.

$$r = \sqrt{\frac{n(\sum xy) - (\sum x)(\sum y)}{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

Results

Higher institutions with a website

If the shared websites are taken into account it could be said that 104 of 153 higher institutions had a website (68%). Whilst if the shared websites are excluded it could be said that 65 of 153 higher institutions had a website (42%).

Measuring Visibility on web search engines

Using URLs with www

After sorting websites' data descending according to the average of indexed webpages in the three search engines Google, Yahoo and Bing. Higher Institution of Administrative Sciences - Bani Sweif got the first place in the top ten places followed by Social Work Institution in Cairo, International Academy for Engineering and Media Science in the second and third places respectively. The Higher Institution of Applied Arts - 6th October got the tenth place (Table 1). the Full list of higher institutions shows that 12 of 65 Institutions (19%) got the last place without any webpage was indexed (appendix 1).

Table (1)
Top ten places according to the average of indexed webpages
using URLs with www

Website Name	Domain	Google	Bing	Yahoo	Indexed webpages Avg	Rank
Higher Institution of Administrative Sciences - Bani Sweif	www.msa-bns.com	3370	32	13	3406	1
Social Work Institution in Cairo	www.swic-edu.com	3190	30	25	3228	2
International Academy for Engineering and Media Science, 6th October	www.iams.edu.eg	2190	75	140	2311	3
Future Academy, Ismailia Road	www.futureacademyegypt.com	970	740	248	1792	4
Giza Engineering Institute, Assiut Road	www.gei.edu.eg	1430	84	163	1568	5
Higher Institution of Islamic Studies, Meet Okba	www.iiss-egypt.org	681	439	1050	1470	6
Al Salam Higher Institute, Gesr Suez	www.alsalam.edu.eg	348	416	995	1095	7
Akhbar Elyom Academy, 6th October	www.akhbaracademy.edu.eg	430	70	140	546	8
Higher Institution for Tourism, Hotels and Computer - Alsyof, Alex	www.seyouf.org	400	54	87	483	9
The Higher Institution of Applied Arts - 6th October	www.appliedarts.org	331	56	105	422	10

Higher institutions were classified into eight groups with interval of five hundred. The group with the greatest frequency was (0-500 webpages) with 57 websites represented 88% of total number websites under study, while there were four groups with the least frequency with 1 frequency represented 2% for each (Table 2).

Table (2)
Institutions sets by total webpages indexed under www URL

Group	frequency	%
0-500	57	87.7
501-1000	1	1.5
1001-1500	2	3.1
1501-2000	1	1.5
2001-2500	1	1.5
2501-3000	0	0.0
3001-3500	2	3.1
3501-4000	1	1.5
Total	65	100%

Using URLs without www

After sorting websites' data descending according to the average of indexed webpages in the three search engines Google, Yahoo and Bing. The first place was obtained by Giza Engineering Institute followed by Social Work Institution in Cairo and Higher Institution of Administrative Sciences - Bani Sweif took the second and third places respectively. Higher Institution of Engineering and Technology in Borg El - Arab got the tenth place (Table 3). The Full list of Higher institutions shows that 14 of 65 (22%) institutions got the last place without any webpage was indexed (appendix 2).

Table (3)
Top ten places according to the average of indexed webpages using URLs without www

Website Name	Domain root	Google	Bing	Yahoo	Indexed webpages Avg	Root Rank	www Rank
Giza Engineering Institute, Assiut Road	gei.edu.eg	6350	477	1150	7210.333	1	5
Social Work Institution in Cairo	swic-edu.com	5750	79	160	5882.333	2	2

Cont. Table (3)
Top ten places according to the average of indexed webpages
using URLs without www

Website Name	Domain root	Google	Bing	Yahoo	Indexed webpages Avg	Root Rank	www Rank
Higher Institution of Administrative Sciences - Bani Sweif	msa-bns.com	3370	32	14	3406.667	3	1
Higher Institution of Islamic Studies, Meet Okba	iiss-egypt.org	955	1190	2940	3125	4	6
International Academy for Engineering and Media Science, 6th October	iams.edu.eg	2720	82	156	2854	5	3
Future Academy, Ismailia Road	futureacademy-egypt.com	1150	750	1830	2510	6	4
Alexandria Higher Institution of Engineering & Technology	aiet.edu.eg	244	736	1800	1580	7	12
El Minya High Institute	mhiet.edu.eg	1110	331	243	1522	8	17
Egypt Higher Institution for Trade and Computer in Mansoura	misr.academy	410	517	1240	1340.333	9	52
Higher Institution of Engineering and Technology Borg El - Arab	alborg-hiet.com	238	603	1440	1321	10	49

Under suggested twelve groups with interval of five hundred, institutions distributed over eight groups. The group with the greatest frequency was (0-500 webpages) with 47 websites represented 72% of total number websites under study, while there were 2 groups with the least frequency with 1 website represented 2% for each (Table 4).

Table (4)
Institutions sets by total webpages indexed under URL without www

Group	Frequency	%
0-500	47	72.3
501-1000	7	10.8
1001-1500	3	4.6
1501-2000	2	3.1
2001-2500	0	0
2501-3000	2	3.1
3001-3500	2	3.1
3501-4000	0	0
4001-4500	0	0
5001-5500	0	0
5501-6000	1	1.5
7001-7500	1	1.5
Total	65	100%

Canonical Optimization

Higher institutions' websites administrators did not do a Canonical Optimization because all the data gathered using URLs with www was not equal with data gathered with URLs without www.

Targeted TLDs

The most popular targeted TLD for the higher institutions developers was edu.eg as combination of gTLD (=edu) and ccTLD (=eg). There were 35 websites under it with 53.8 % of the total. in the second place of popularity came the.com gTLD with 19 websites represents 29.2 % of the total. while the gTLD.aca came in the last place with one website (Table 5).

Table (5)
Targeted TLDs

Tld	Frequency	%
.aca	1	1.5
.com.eg	2	3.1
.net	2	3.1
.org	6	9.2
.com	19	29.2
.edu.eg	35	53.8
total	65	100%

Which search engines index more webpages

All the search engines indexed more webpages under the root domain (URLs without www.) than under the www domain. The most indexing search engine for the webpages whether under www or root domain was google (Table 6).

Table (6)
total webpages indexed by the three search engines

S.E	Root	www	Total
google	31115	16893	48008
yahoo	14766	4569	19335
Bing	7101	3151	10252
total	52982	24613	

Calculating and ranking higher institutions according to WIFs & RWIFs

Under this section WIFs were calculated by gathered data from google search engine using the root domain. The reason for preferring URLs without www data than the URLs with www was that search engines under study indexed more webpages under it than under domain with www. Also, the reason for preferring google was that it indexed more webpages than the others under study.

Calculating WIFs

One website was excluded from calculating WIF as it had not any

indexed webpages in google. According to the overall data gathered under this section, twenty-one institutions got WIFs values equal or greater than one. total institutions could be classified according to WIFs value as shown in table 7. Full list of WIFs values for the sixty-four institutions could be overviewed through appendix 3.

Table (7)
WIF values

WIF value	Frequency	%
Equal Zero	24	37
Greater than zero and less than one	19	30
Equal or greater than one	21	33
Total	64	100

Higher Institutions Top Ten Ranks According to WIFs values

After sorting gathered data according to WIFs values descending, Higher Institution of Management and Secretariat - Masr Kadema obtained the first place in the top ten places followed by October High Institution for Engineering & Technology and Higher Institution of Islamic Studies - Meet Okba in the second and third places respectively. Agamy High Institution of Administrative Sciences got the tenth place. Table 8 shows the top ten places.

Table (8)
Higher Institutions Top Ten Ranks According To WIF values

Website name	Internal	External	Total Links	Indexed webpges	WIF	Rank
Higher Institution of Management and Secretariat, Masr Kadema	73493	0	73493	219	73493.00	1
October High Institution for Engineering & Technology	12676	17	12693	704	12676.02	2
Higher Institution of Islamic Studies, Meet Okba	6995	297	7292	955	6995.31	3
The High Institution of Computer and Information Systems, Mansora	6290	4	6294	150	6290.03	4

Cont. Table (8)
Higher Institutions Top Ten Ranks According To WIF values

Website name	Internal	External	Total Links	Indexed webpages	WIF	Rank
The Higher Institution of Applied Arts - 6th October	4607	8	4615	439	4607.02	5
Higher Institution of Administrative Sciences in Manzala	1688	0	1688	369	1688.00	6
Social Service Institution for Girls in Garden City	1083	111	1194	68	1084.63	7
Pyramids High Institution for Engineering & Technology, 6th October	1046	1	1047	217	1046.00	8
The Abbasia Institution for Computers and Commercial Sciences	658	0	658	90	658.00	9
Agamy High Institution of Administrative Sciences	579	0	579	72	579.00	10

Calculating RWIFs

One website was excluded from calculating RWIFs as it had not any indexed webpages in google. According to the overall data gathered under this section six institution got RWIF value equal or greater than one. Total institutions could be classified according to RWIFs value as shown in table 9. Overall RWIF values for the sixty-four institutions could be overviewed through appendix 4.

Table (9)
RWIF values

RWIF value	Frequency	%
Equal Zero	31	48
Greater than zero and less than one	27	42
Equal or greater than one	6	10
Total	64	100

After sorting gathered data according to RWIFs values descending, Al Alson Higher Institution obtained the first place in the top ten places followed by Higher Institution of Advanced Studies and High Institution for Management Science - New Cairo in the second and third places respectively. Higher Institution of Islamic Studies - Meet Okba got the tenth place. Table 10 shows the top ten places.

Table (10)
Higher Institutions Top Ten Ranks According to RWIF values

Website name	Internal	External	Total Links	Indexed webpages	RWIF	Rank
Al Alson Higher Institute	0	8917	8917	206	43.29	1
Higher Institution of Advanced Studies	0	7709	7709	423	18.22	2
High Institution for Management Science in New Cairo	0	2	2	1	2.00	3
Social Service Institution for Girls in Garden City	1083	111	1194	68	1.63	4
Higher Institution of Engineering and Technology in Arish	0	213	213	137	1.55	5
Zagazig Higher Institution of Engineering and Technology	0	3	3	3	1.00	6
New Cairo Higher Institution for Administrative Sciences and Computer	0	67	67	104	0.64	7
Higher Institution of Tourism and Hotels in Luxor	0	4	4	12	0.33	8
Modern Academy of Engineering and Technology	0	55	55	174	0.32	9
Higher Institution of Islamic Studies, Meet Okba	6995	297	7292	955	0.31	10

Correlation between Wif and Rwif

A Pearson product-moment correlation coefficient was computed to assess the relationship between the RWIF and WIF. Result indicated that there was not a relationship between the two variables, $r = -0.033$, n

= 64, $p = 0.813$. The result was *not* statically significant at $p < 0.05$. The next scatterplot summarizes the results (Figure 1).

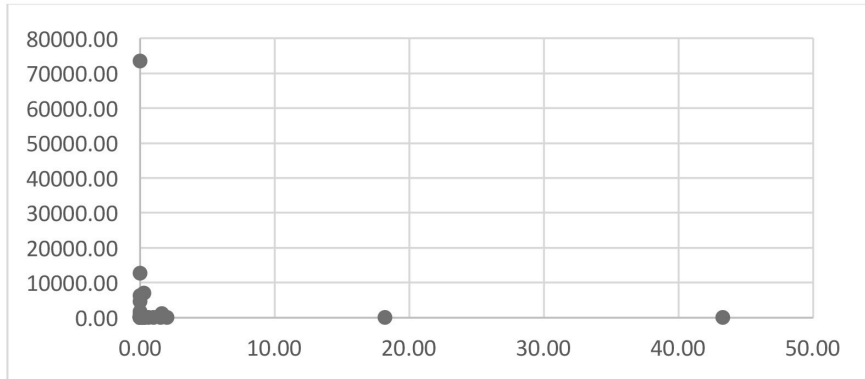


Figure 1 scatterplot for WIF & RWIF correlation

Discussion

The first finding of this study was the answer to the first question raised which sought to find out how many higher institutions in Egypt have a website. It was preferred to depend on the data that exclude the shared websites. Because it will give a misleading result. The current study found that there was 65 of 153 institutions (43%) had a website. This finding was unexpected because most of these institutions have technical and practical fields. The most common one is information technology. So, it can establish its own website by their self-efforts. this suggests that Strict procedures need to be done by the ministry of higher education in encouraging institutions' Administrative Boards to establish website for their institutions. A possible explanation for this might be Lack of awareness, among most of institutions' Administrative Boards, about the importance of having an official website.

The second question in this study sought to find out to what extent higher institutions in Egypt are indexed in web search engines. This study calculated the indexed webpages by search engines for the websites under study twice: by using www URL and root URL (without www). This research unfolded more than one interesting finding. The first most interesting finding according to the data gathered using www URL

indicated that 53 of 65 institutions (82%) had webpages indexed in one or more of the search engines used under study.

On the other side data gathered using root URL revealed that 64 of 65 institutions (98%) had webpages indexed in one or more of the search engines used under study. For the public, URL with www is the same URL without www, which is wrong. So, for the public, number of indexed webpages under URL with www should be the same under URL without www. But, in fact they are not the same for search engines. Search engines deal with each one of them as a separate URL. a possible explanation for this result is that web administrators did not use search engines admin tools effectively. One example of them is Google's web master tools. Web masters could confirm through it that they own both shapes of URLs. So, search engine deals with the two shapes of URL as one.

The second important finding is that most of websites whether under www URL or under root URL were having average of indexed webpages in the three search engines ranged between 0-500 webpages. This indicates that most of these websites have a slow growth rate. The explanation for this result is that most of these websites were static ones. This study intended to examine the growth rate of the websites under study but its data was not available easily. So a separate study is needed to discuss and compare the date of establishing every website and its total indexed webpages.

As an answer for the third question the results indicated that all websites administrators did not do a Canonical Optimization. A possible Explanation for this result that all of them are not aware of the web admins tools and its benefits to prevent duplicated links, under different URL shapes, pointing to their websites.

The fourth question in this study sought to know what is the preferred TLDs for higher institutions' websites in Egypt. In this matter there is a fact or a rule for any website administrators who want to establish an official website, which is to register it under TLDs indicates both the place and the scope of the website. the result belonging to this topic is consistent relatively with the mentioned rule. As, 53.8% of websites were established under edu.eg, but What was surprising is that the rest of websites were under various TLDs other than edu.eg. The

most used TLD for them was.com. putting an academic website under gTLD other than.edu sending a bad and negative message about it to the website's visitors. This result indicates that there was Lack of awareness between both websites administrators and institutions' Administrative Boards about the importance of establishing their websites under the appropriate TLDS.

The fifth question in this study targeted to know the most indexing search engine. The results of this study showed that google was the most indexing search engine Whether under URLs with www or under URLs without www. In fact, this result was not surprising. Because it is known that google is interested in the web middle east market at all, not only in the Egyptian market. This result suggests that other search engines, Bing and Yahoo should do more efforts to compete with google in the web middle east market. This result does not correspond to the result indicated by Alireza(Noruzi, 2006b). he reported that the Middle-Eastern web sites (Persian, Kurdish, Turkish, Arabic, and Hebrew languages) may not receive and attract the attention that they deserve from the World Wide Web community. This difference between the result of the current study and the result of Alireza's study could be attributed to the fact that his study took place at a time when the global search engines corporations were not yet interested in the Arab market.

The sixth question in this study sought to rank higher institutions according WIF and RWIF. The results showed evidence to prefer RWIF than WIF. As, there were illogical numbers of the inlinks for the most of top ten higher institution websites when compared to the total webpages indexed for each one. These abnormal totals of inlinks influenced the ranks in favor of some websites which got those results. On the other hand, when RWIF was calculated, numbers became logical again because the bias factor (internal links) was removed. Another significance result was that most of WIF and RWIF ranged between zero and less than one. This result indicates that most of websites administrators have a Lack of awareness about websites SEO (Search Engine Optimization).

The last question in this study aimed to find out if there is a correlation between WIF and RWIF for higher institutions' websites in Egypt. According to the result indicated there was no correlation between them. With regard to this result we cannot generalize the result of this relation. Because the result was *not* statically significant at $p <$

0.05. However, this finding of the current study does not support the previous research by Chakravarty & Wasan(Chakravarty & Wasan, 2015) they reported that there was a correlation between WIF and RWIF. A possible explanation for this result is that most of the websites under study lacks effective SEO practices which affected its data according to outlinks, inlinks and indexed webpages which were reflected in the correlation's result.

Conclusions

The main goal of the current study was to evaluate the web presence of higher institutions in Egypt. As mentioned in the literature review, on the Arab world level, there were only three studies concerned with universities. No prior studies were concerned with higher institutions in Egypt. This study has shown that most of higher institutions in Egypt had a relatively good web presence, but they have a low web impact factor because of the bad practices of their webmasters. It was also shown that there was not a correlation between WIF an RWIF.

This indicated that correlation between WIF and RWIF is relative and not absolute. This study has several practical applications. Firstly, it helps in distinguishing between good and bad practice from both websites' administrators and institutions' Administrative Boards of higher institutions in Egypt. Secondly, it gives suggestions to correct the negatives observed. Further research in this field would be great by conducting regional studies or local similar studies. Some of these further studies may take into account calculating websites growth rates. The study recommends that higher education ministry in Egypt should oblige higher institutions' boards to establish an individual website compliant with the international web standards for each higher institution before getting the official license (for new institutions) or before renewing the official license (for established institutions). Also, it is recommended that web administrators should raise their knowledge about using search engine webmaster tools and about SEO (search engine optimization).

Additionally, it is important that web administrators should not put their websites under TLDs other than.edu as this sends a bad and negative message to the website's visitors. Finally, a number of important limitations need to be considered. Firstly, the current study has only examined higher institutions officially which were certified by the

Ministry of Higher Education in Egypt. Other uncertified institutions were excluded because they are not recognized by the government or from most members of society. Secondly, the study did not examine the growth rate of higher institutions because belonging data was not easily available.

References

- Academy, N. (2017). Correlation coefficient. Retrieved from https://www.usna.edu/Users/oceano/pguth/md_help/html/satb4qgk.htm
- Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the world wide web: Methodological approaches to 'webometrics'. *Journal of Documentation*, 53(4), 404-426. doi:doi:10.1108/EUM0000000007205
- Arif, A., & Ismail, N. (2013). Web impact factor for Malaysian public universities. *International Journal of Future Computer and Communication*, 2(3), 151.
- Arif, K., & Haroon, I. (2015). Calculating web impact factor for university websites of Pakistan. *The Electronic Library*, 33(5), 883-895. doi:doi:10.1108/EL-01-2014-0022
- Boell, S. K., Wilson, C. S., & Cole, F. T. H. (2008). Usage of different Web Impact Factors for ranking Australian Universities. *COLLNET Journal of Scientometrics and Information Management*, 2(2), 57-70. doi:10.1080/09737766.2008.10700854
- Chakravarty, R., & Wasan, S. (2015). Webometric analysis of library websites of higher educational institutes (HEIs) of India: A study through Google Search Engine. *DESIDOC Journal of Library & Information Technology*, 35(5).
- Education, M. O. H. (2017). *Private institutions*. Retrieved from <http://portal.mohe.gov.eg/ar-eg/Pages/institute-names.aspx>
- Elgohary, A. (2008). Arab universities on the web: a webometric study. *The Electronic Library*, 26(3), 374-386. doi:doi:10.1108/02640470810879518
- Farag, A. (2006). Web presence of Saudi Arabian Universities on the web: an exploratory study of the web impact factor. *Arab Journal of Library & Information Science*, 26(3), 151-174.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236-243. doi:doi:10.1108/EUM0000000007167

- Jalal, S. K., Biswas, S. C., & Mukhopadhyay, P. (2010). *Web impact factor and link analysis of selected Indian universities*.
- Jati, H. (2011). Web impact factor: a webometric approach for Indonesian universities. Paper presented at the International Conference on Informatics for Development.
- Kumar, K. (2017). Web impact factor analysis for deemed universities in Andhra Pradesh. *DESIDOC Journal of Library & Information Technology*, 37(2), 98.
- Kumar, K. (2018). *Web Impact Factor and Link Analysis of Indian Council of Agricultural Research (ICAR) Organizations*. *International Journal of Knowledge Content Development & Technology*, 8(1), 5-23. doi:10.5865/ijkct.2018.8.1.005
- Maharana, R. K., Panda, K., & Sahoo, J. (2012). *Web impact factor (WIF) and link analysis of Indian Institute of Technologies (IITs): A webometric study*.
- Mohammadamin, E., & Fereshteh, D. (2011). *Visibility and impact of Iranian research institutions on the web*. *Library Hi Tech News*, 28(1), 4-9. doi:doi:10.1108/07419051111130367
- Noruzi, A. (2005). *Web impact factors for Iranian Universities*. *Webology*, 2(1).
- Noruzi, A. (2006a). *The web impact factor: a critical review*. *The Electronic Library*, 24(4), 490-500. doi:doi:10.1108/02640470610689188
- Noruzi, A. (2006b). *Web presence and impact factors for Middle-Eastern countries*. *Online*, 30(2), 22-28.
- Prabakaran, R. (2014). Hyperlink analysis and web impact factor of websites of observatory libraries. *International Journal of Library and Information Studies*, 4(4), 38-51.
- Ramanayake, K. H., Xianqiao, C., & Bing, S. (2017). Calculating web impact factor for university library websites in Sri Lanka. *Journal of the University Librarians Association of Sri Lanka*, 20(1).
- Rassi, M., Jalali Dizaji, A., Mousavizadeh, Z., Abbasgholi, N., & Haghighian Roudsari, A. (2013). Assessment of visibility and web impact factor (WIF) of Food Science & Technology Institutes websites based on webometrics scales. *Iranian Journal of Nutrition Sciences & Food Technology*, 7(5), 449-459.
- Satpathy, S. K., & Maharana, R. K. (2012). Web impact factor and webometric study of National Institute of Technology (NITs) of

- India. *Pearl: A Journal of Library and Information Science*, 6(1), 47-55.
- Smith, A., & Thelwall, M. (2002). Web impact factors for Australasian universities. *Scientometrics*, 54(3), 363-380.
 - Thelwall, M. (2001a). Extracting macroscopic information from Web links. *Journal of the American Society for Information Science and Technology*, 52(13), 1157-1168. doi:10.1002/asi.1182
 - Thelwall, M. (2001b). Results from a web impact factor crawler. *Journal of Documentation*, 57(2), 177-191. doi:doi:10.1108/EUM0000000007081
 - Vaughan, L., & Thelwall, M. (2005). A modeling approach to uncover hyperlink patterns: The case of Canadian universities. *Information Processing & Management*, 41(2), 347-359.
 - Verma, M. K., & Brahma, K. (2017). Ranking of National Institutes of Technology (NITs) of northeast region of India on the basis of web impact factor. *COLLNET Journal of Scientometrics and Information Management*, 1-14.
 - Wimaladharma, S., & Herath, H. M. (2016). *An analysis of significance of Revised Web Impact Factor for ranking the websites of state universities in Sri Lanka*. Paper presented at the Information and Automation for Sustainability (ICIAfS), 2016 IEEE International Conference on.

Submitted: February 2018

Accepted: February 2019

