



# International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: [www.ijarcsse.com](http://www.ijarcsse.com)

## Web Mining: Cybermetrics Analysis of the Nine (9) Newly Established Federal Universities in Nigeria in 2011

**Emeka Ogbuju\***  
Computer Science Dept.,  
Federal University Lokoja,  
Kogi State, Nigeria

**Prof. Sunday E. Adewumi**  
Computer Science Dept.,  
Federal University Lokoja,  
Kogi State, Nigeria

**Dr. Virginia Ejiofor**  
Computer Science Dept.,  
Nnamdi Azikiwe  
University Awka,  
Anambra State, Nigeria

**Edward Onyebueke Agu**  
Computer Science Dept.,  
Federal University Wukari,  
Taraba State, Nigeria

**Abstract -** *The National Universities Commission in Nigeria had made efforts to orient the Universities on the need to maintain her research activities in a web-presentable format. Most of the Universities that had ranked low in the World Webometrics University Ranking have quality educational or research outputs but lack functional online presentation culture necessary for a favourable ranking. The application of Web Analytics will help these Universities to understand the cyber metrics applied in the ranking process. Web Analytics collects online data using Web Mining approaches, report the trends and structures of those data and analyses them for institutional effectiveness. This paper discusses the cyber metrics of the nine (9) newly established Federal Universities in Nigeria in 2011 with the aim to examine the ranking parameters used and recommend compliance to standard measures that will enhance higher ranking of World Universities. The quantitative approach was used to analyse the Universities' websites by observation and comparison with the ranking metrics applied on Search Engine Optimization tools. The results showed that most of the Universities which lack qualitative research output online can improve using the recommended framework.*

**Keywords -** *Web Mining, Cybermetrics, Federal Universities, Webometrics, Web Analytics*

### I. INTRODUCTION

The Federal Executive Council of Nigeria at its 39th meeting on Wednesday, 10 November, 2010, approved the establishment of 12 new federal universities on the bases of equity and access to education spread across the six geo-political zones. Nine of these universities were approved for immediate take-off under Phase I of the initiative in Otuoke, Bayelsa; Dutsin-ma, Katsina; Dutse, Jigawa; Wukari, Taraba; Lokoja, Kogi; Lafia, Nasarawa; Oye-Ekiti, Ekiti; Ndufu-Alike Ikwo, Ebonyi and Kashere, Gombe States. The Phase II of the initiative in 2013 located the remaining three (3) at Birnin-Kebbi in Kebbi State, Gusau in Zamfara State and Gashua in Yobe State to address the twin challenges of access and equitable educational development of states in the federation. Nigeria today has 40 approved federal universities, 38 state universities and over 50 private universities [1]

By 2011, with the appointment of Vice Chancellors and Registrars for the nine (9) universities under Phase I, the universities took off with specific visions which they carved for themselves (see Table II). Design and development of their various websites followed. By 2012, all the universities were fully represented online (See Table I) and by January 2013 they had started participating in the Webometrics Ranking of World Universities conducted twice yearly (January and July) by Cybermetrics Labs in Spain (See Table III).

Table I List of the 9 New Universities and Their Website Addresses

S/No	Name of University	Website Address
1	Federal University, Dutse, Jigawa State	<a href="http://www.fud.edu.ng">www.fud.edu.ng</a>
2	Federal University, Dutsin-Ma, Katsina	<a href="http://www.fudutsinma.edu.ng">www.fudutsinma.edu.ng</a>
3	Federal University Kashere, Gombe State	<a href="http://www.fukashere.edu.ng">www.fukashere.edu.ng</a>
4	Federal University Lafia, Nasarawa State	<a href="http://www.fulafia.edu.ng">www.fulafia.edu.ng</a>
5	Federal University Lokoja, Kogi State	<a href="http://www.fulokoja.edu.ng">www.fulokoja.edu.ng</a>
6	Federal University Ndufu-Alike, Ebonyi State	<a href="http://www.funai.edu.ng">www.funai.edu.ng</a>
7	Federal University Otuoke, Bayelsa State	<a href="http://www.fuotuoke.edu.ng">www.fuotuoke.edu.ng</a>
8	Federal University Oye-Ekiti, Ekiti State	<a href="http://www.fuoye.edu.ng">www.fuoye.edu.ng</a>
9	Federal University Wukari, Taraba State	<a href="http://www.fuwukari.edu.ng">www.fuwukari.edu.ng</a>

The indicators used in assessing these websites fall under the research domain of Web Mining and the emerging Cloud Mining (Software-as-a-Service). It uses the implementation of applied cybermetrics techniques based on the positioning on search engines of Web domains and analysis of the information usage through Web data mining of log files. Web Mining can be broadly defined as the discovery and analysis of useful information from the World Wide Web.

This paper analyzes the nine (9) universities' websites to help the institutions benchmark themselves against each other in a healthy webometrics ranking to discover hidden patterns and analysing them.

Table II Vision Statements Of The Institutions  
(source: extracted from each institutions website)

<i>Fudutsinma</i>	<i>To build a World-Class, ICT-Driven, top ranking university committed to excellence in research and the production of generation of leaders with a passion for service</i>
<i>Fulokoja</i>	<i>To be the best among the nine newly established Federal universities in Nigeria in 2011 and one of the top ten ranking universities in Africa.</i>
<i>Fulafia</i>	<i>To be a renowned institution of learning, research and innovation for positive socio-economic transformation of the nation.</i>
<i>Funai</i>	<i>To become a vibrant centre of learning and research that will be reputable nationally and internationally and a hub for the economic transformation and development of the region through creative and innovative endeavours.</i>
<i>Fuoye</i>	<i>To become an institution of first choice recognized for providing critical opportunities for student success; acknowledged as a primary and engaged regional and global resource for entrepreneurial educational and best practices in engineering and farming technology; and valued as a university where there are faculty, students, and professional staff who are active in integrating cutting-edge technology in multidisciplinary and disciplinary research.</i>
<i>Fuotuoke</i>	<i>To gain and maintain a reputation as a world-class university that challenges all its students to achieve the highest levels of intellectual and personal growth, to promote sustainable development, as well as contribute purposeful and ethical service to the nation and mankind.</i>
<i>Fukashere</i>	<i>To become a world-renowned centre of learning where students are prepared with the knowledge, skills, and dispositions they need in order to serve their community, state, nation, and the world through excellence in teaching, research, and service.</i>
<i>Fuwukari</i>	<i>To be leader among world class public Universities by: advancing knowledge through high quality ICT centric educational experiences for students; encouraging entrepreneurship; conducting leading edge research and scholarship in all areas and promoting an intellectual environment that is anchored on the tenets of open dialogue and inquiry, and a deep and abiding appreciation of the entire spectrum of human experience.</i>
<i>Fudutse</i>	<i>To attract a diverse cast of lecturers and students, support research and teaching on local, national and global issues and create academic relationships with many Universities and higher education institutions in Nigeria and across the world.</i>

There is no way any of these institutions can reach their various visions without the application of Information Technology. Most of the institutions had enshrined this concept into their vision statements and are working hard towards making a good mark with ICT.

Table III Webometrics Ranking of the Institutions from 2013 to January 2015  
(source: www.webometric.info)

Domain/ Parameters	July 2013		July 2014		January 2015	
	World	Nigeria	World	Nigeria	World	Nigeria
<i>Fuoye</i>	12746	28	10760	27	7146	16
<i>Fudutsinma</i>	-	-	13762	43	5114	14
<i>Funai</i>	16303	57	15795	57	11443	34
<i>Fuotuoke</i>	17266	69	15830	58	16409	67
<i>Fulokoja</i>	18682	91	18318	90	17543	81
<i>Fuwukari</i>	19310	100	18745	98	18843	95
<i>Fud</i>	18534	83	18777	99	14795	52
<i>Fukashere</i>	19902	107	19088	104	20522	113
<i>Fulafia</i>	20035	109	19093	105	19089	98

## II. DEFINITION OF TECHNICAL TERMS

- A. *Web Mining*: The extraction of interesting and potentially useful patterns and implicit information from artefacts or activity related to the World Wide Web
- B. *URL*: Uniform Resource Locator, a complete address of a web page
- C. *Crawling*: The movement of a web search program across websites to extract web pages. The crawler (also called spider, ant, scutter or automatic indexer) is a web robot which systematically browses the World Wide Web, typically for the purpose of Web indexing.

- D. *Indexing*: Storage of web pages on the database of a search engine for future access
- E. *Visibility*: The total number of unique external links received (backlinks, inlinks or inbound links) by a site.
- F. *Search Engine Optimization (SEO)*: The process of affecting the visibility of a website or a web page in a search engine's unpaid results - often referred to as "natural," "organic," or "earned" result using specific keywords
- G. *Webometrics (also Cybermetrics)*: The measurement of the World Wide Web to get knowledge about the number and types of hyperlinks, structure of the World Wide Web and usage patterns.
- H. *Rich Files*: Files other than the conventional HTML/ASP web pages available online. Examples are Adobe Acrobat (.pdf), Adobe PostScript (.ps), Microsoft Word (.doc) and Microsoft PowerPoint (.ppt).
- I. *Traffic*: The number of hits or visits logged or recorded by a site

### III. STATEMENT OF THE PROBLEM

Many Universities rank low in the World Webometric University Ranking. Most of these institutions have quality educational or research outputs but lack good online presence and culture. Gaps have been identified for this challenge and this paper seeks to close those gaps. Maintaining quality contents online is a discipline which requires both technical and management approaches.

### IV. LITERATURE REVIEW

#### A. Web Mining

Web mining [6] is one of the major and important fields of data mining. Web Mining is the extraction of interesting and potentially useful patterns and implicit information from artefacts or activity related to the World Wide Web [2]. There are three knowledge discovery domains that pertain to web mining: Web Content Mining, Web Structure Mining, and Web Usage Mining [8]. Web content mining is the process of extracting knowledge from the content of documents or their descriptions. Web structure mining is the process of inferring knowledge from the World Wide Web organization and links between references and referents in the Web. And, the web usage mining, also known as Web Log Mining, is the process of extracting interesting patterns in web access logs.

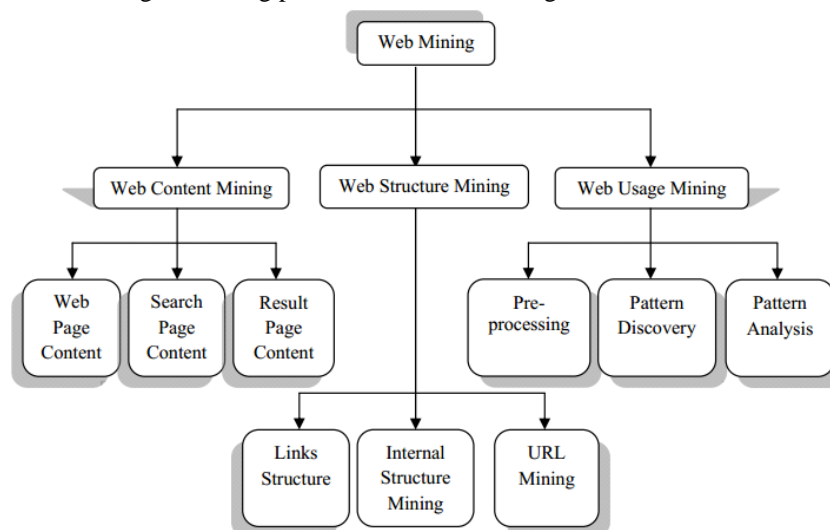


Fig I: Taxonomy of Web Mining (Source: Kavita et al, 2011)

The taxonomy of the studies in web mining presented in Figure I describes those three major areas of knowledge discovery from the web [3].

#### 1) Web Content Mining

Web Mining is information extraction from the textual contents of the web and the search engine result pages. Many pages are open to access the information on the web; these pages are the content of the web. Searching the information and open search pages is also the content of the web. Last accurate result is defined the result pages content mining.

#### 2) Web Structure Mining

We can define web structure mining in terms of a graph. The web pages are represented as nodes and hyperlinks as edges. Basically, it shows the relationship between the user & the web. The motive of web structure mining is generating structured summaries about information on web pages/webs showing the link of one web page to another.

#### 3) Web Usage Mining

This is the discovery of meaningful pattern from data generated by a client-server transaction on one or more web localities. The web is a collection of interrelated files on one or more web servers. It is automatically generated and the data is stored in server access logs, refers logs, agent logs, client sides cookies, user profile, metadata, page attribute, page content & site structure.[3] Web usage mining aims at utilizing data mining techniques to discover the usage patterns from web based application. It is a technique that predicts user behaviour as they interact with the web.

The web usage mining is categorized into three (3) phases:

- a) *Pre-processing*: According to client, server and proxy server, it is the first approach to retrieve the raw data from web resources and process them. The original raw data is automatically transformed. Usage, content and structure must be pre-processed.
- b) *Pattern Discovery*: Pattern discovery process is based on algorithms developed from different fields such as statistics, data mining, machine learning and pattern recognition. This section describes the kinds of mining activities that have been applied to the Web domain. Clustering, Classification and Association Rule Mining among others are applied. [4]
- c) *Pattern Analysis*: It is the last step in the overall Web Usage Mining process. Pattern analysis is needed to filter out uninterested rules or patterns from the set found in the pattern discovery phase. Pattern analysis consists of knowledge query such as SQL. Sometimes summarized information is useful such as loading usage data into a data cube and performing OLAP operations. A picture speaks thousand times better than words. Therefore visualization techniques such as graphs (bar chart, pie chart etc) of patterns or showing colours for different values can also give more focus on patterns or trends in the data.[4]

## B. Webometrics

Quantitative Studies of the web have been named as Webometrics by Almind and Ingwersen [5] who first coined the term in 1997. Webometrics covers the research of all network-based communications using metric and other quantitative measures. The science of Webometrics (also Cybermetrics) tries to measure the World Wide Web to get knowledge about the number and types of hyperlinks, structure of the World Wide Web and usage patterns. So, Webometrics is an application of Web Mining.

According to [9], the definition of Webometrics is "the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches." An alternative definition was given by [7] as "the study of web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study" The purpose of this alternative definition was to help publicize appropriate methods outside of the information science discipline rather than to replace the original definition within information science.

Webometrics as an application of Web Mining can be applied in any functional sector with an online presence; for example, aviation online booking activities and e-commerce websites can be mined. Likewise, the University education in Nigeria can be mined with the aim of determining best-ranking Universities with certain parameters. Though we are discussing Webometrics ranking, there are other areas of ranking in the University system of which Okebukola [12] noted six (6) different types of World Universities ranking as follows:

- 1) Times Higher Education-QS Ranking
- 2) Academic Ranking of World Universities
- 3) Webometrics Ranking
- 4) Professional Ranking of World Universities
- 5) Newsweek Ranking
- 6) Performance Ranking of Scientific Papers for World Universities

Webometrics ranking stresses on the importance and use of the web in teaching, learning and research. It portrays the web as the best showcase for Universities as it connects all stakeholders - professors, students and the general public.

The Nigerian Universities Commission [11] listed the following as factors responsible for Nigerian Universities' poor performance:

- 1) Scant attention paid to present findings of research conducted by scholars in Nigerian universities in a web searchable form which manifests in
  - Publishing in low impact local journals without Internet links.
  - Non-publishing in electronic journals

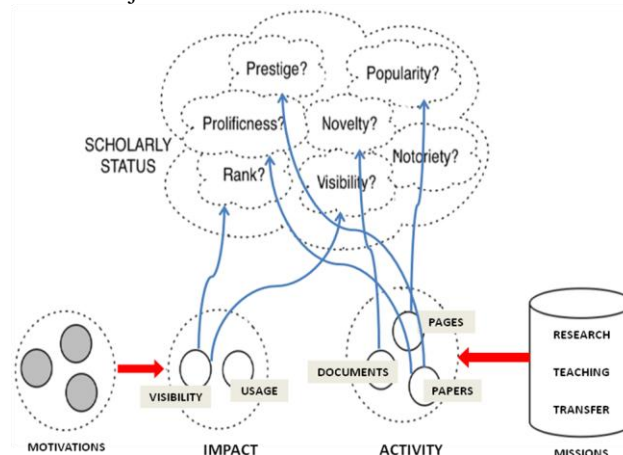


Fig II: Webometrics Ranking Model (Source: Okebukola NUC Ranking Workshop [12])

- 2) The absence of Nigerian universities on the Internet in a form that can be picked by the radar of Cybermetrics Research Group.
- 3) Lack of up-to-date and scanty content of the websites of Nigerian universities.

Using the model in Figure II, Universities are ranked based on contents size, quality resources, visibility and popularity. These four parameters are essential to the overall creation, design or management of any institutions' online presence. The parameters comply with the World Universities Ranking rating shown in Table IV below

Table IV Webometrics Rating  
(source: cybermetrics lab, 2011)

Criteria	Webometrics	Rating
Size	Web Size	20%
Research Output	Rich Files	15%
	Google Scholar	15%
Impact	(Link) Visibility	50%

To get a good ranking in any of the criteria presented, the university should showcase their knowledge production capabilities, academic works/resources of the faculty members, institution's policies and a well structured links network showing popularity and authority.

Ranking is achieved by considering the parameters in the formula presented in equation 1 [12] below thus applying the four identified parameters.

$$WR = (50\% * V) + (20\% * S + 15\% * R + 15\% * Sc) \quad (1)$$

The formula allows each one of the parameters to have a different weight but maintaining the ratio 1:1. The four (4) parameters were from the quantitative results provided by the search engines (Google, Yahoo, Bing and Exalead). They are: [11], [10]

- 1) Size (S): Number of pages recovered from four engines: Google, Yahoo, Live Search (Bing) and Exalead.
- 2) Visibility (V): The total number of unique external links received (inlinks) by a site can be only confidently obtained from Yahoo Search.
- 3) Rich Files (R): After evaluation of their relevance to academic and publication activities and considering the volume of the different file formats, the following were selected: Adobe Acrobat (.pdf), Adobe PostScript (.ps), Microsoft Word (.doc) and Microsoft PowerPoint (.ppt). These data were extracted using Google, Yahoo Search, Live Search and Exalead.
- 4) Scholar (Sc): Google Scholar provides the number of papers and citations for each academic domain. These results from the Scholar database represent papers, reports and other academic items.

The algorithm for the formula works by either normalizing the data collected from all of the search engines (equation 2) to determine the web size or by using the median method (equation 3). The number of rich files is also counted by using equation 4.

$$N_a = \frac{\log(n_a + 1)}{\log(\max(n_i) + 1)} \quad (2)$$

$$S_a = \frac{1}{2} * ((G_a + Y_a + L_a + E_a) - \max(G_a, Y_a, L_a, E_a) - \min(G_a, Y_a, L_a, E_a)) \quad (3)$$

$$R_a = PDF_a + DOC_a + PPT_a + Ps_a \quad (4)$$

## V. METHODOLOGY

The quantitative approach was used for the analysis of the Institutions' websites. The tools used are observation and comparison with the ranking metrics applied on search engine optimization (SEO) open tools online [14]. This study was carried out by repeated searches with the major indexing engines as well as visits to the University websites between April and May 2015 for data collection and analysis. Tables were presented to show the metrics and comparisons were made using the SEO tools. The overall workflow is presented in Figure III.

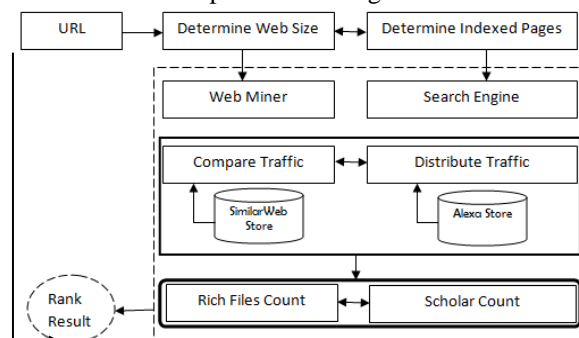


Fig III: Overall Framework

The metrics applied for the traffic rating was Alexa traffic ranking and Similar Web comparison tools. The research did not pay attention to design issues involving beauty and aesthetics of the sites.

The overall objective of the research in the long run is to develop a Web Miner tool that would incorporate the entire metrics so as to enable institutions to self crawl and index their sites as well as their competitors in order to reveal patterns and offer insights. Institutions can use this tool online to benchmark themselves.

## VI. RESULTS AND DISCUSSION

The result is presented using the three (3) major rating criteria - web size, impact (a measure of visibility, link and traffic) and institution's research output (Rich files and Google scholar). Table III shows the institutions' ranking as at January 2015 by Webometrics Ranking of World Universities with fudutsinma.edu.ng as the 14th University in Nigeria and 1st among the nine.

### A. Web Size

This is the number of pages collected from the four engines: Google, Yahoo, Live Search (MSN Bing) and Exalead. A search engine optimization search produced the following results:

Table V Web Size By Search Result And Indexing

URL	Google		Yahoo		Bing		Exalead	
	No of Results	No of Indexed Pages	No of Results	No of Indexed Pages	No of Results	No of Indexed Pages	No of Results	No of Indexed Pages
Fuoye	340, 000	108, 800	15, 100	2, 910	5, 920	6	1,299	-
Fudutsinma	22, 800	98, 300	15, 300	748	2, 520	4	79	-
Funai	8,240,000	22, 200	13, 100	887	2, 800	6	467	-
Fuotuoke	63,200	878	18, 600	894	1, 860	12	264	-
Fulokoja	91, 600	353	32, 800	374	2, 350	0	224	-
Fuwukari	65, 000	244	42, 800	220	2, 250	1	165	-
Fud	13,700,000	20, 700	238, 000	398	1, 500	1	2, 301	-
Fukashere	10, 900	148	10, 800	194	2, 180	0	9	-
Fulafia	55,600	98	37, 800	106	5, 550	2	73	-

Google, a popular search engine has a higher number of indexed pages in most of the websites. The institutions need to work harder to increase the indexing of Yahoo and MSN Bing. This can be achieved by direct submission of the sitemaps to these search engines.

Table VI Web Size by Mined Pages and Indexing

URL	No of Mined Pages	Google No of Indexed Pages	Yahoo No of Indexed Pages	Bing No of Indexed Pages
Fuoye	1, 455 (11dt)	108, 800	2, 910	6
Funai	1, 732 (10dt)	22, 200	887	6
Fuotuoke	2, 679 (102dt)	878	894	12
Fulokoja	1, 402 (4dt)	353	374	0
Fuwukari	41(1dt)	244	220	1
Fud	44(0dt)	20, 700	398	1
Fukashere	85 (5dt)	148	194	0
Fulafia	42(0dt)	98	106	2
Fudutsinma	45(0dt)	98, 300	748	4

Table VI shows the no of pages available in each institutions website at various depths and presents the corresponding no of indexed pages by each of the search engines. It shows that the search engines are not indexing all the pages available on the site as well as on the external pages. This is as a result of poor search engine optimization on the pages that are missed. For example, Fulokoja has a total of 1, 402 pages at 4 depths available on the site but the search engines are only able to index a maximum of 374 pages (Yahoo) thereby missing out 1, 028 pages which were poorly optimized. This will contribute in the institution losing marks on the 20% allotted for the web size rating. Note that some institutions which have more indexed pages than the mined pages have their external pages and their links also indexed. This is a result of good optimization techniques.

### B. Visibility

This is the number of unique external links received by a site. Inlinks can be from websites of partner institutions or search engines. The domain analysis of the institutions in Table VII presents a measure of the visibility of the institutions web presence using Alexa traffic tool. Note that Alexa measures traffic only from users with its toolbars installed.

Table VII Domain Analysis Of The Institutions

Domain/ Parameters	Fud	Fudutsin ma	Funai	Fuoye	Fuotuoke	Fukash ere	Fulokoja	Fulafia	Fuwuka ri
PageRank	5/10	4/10	5/10	5/10	5/10	4/10	5/10	4/10	0/10
Domain Authority	30%	34%	27%	42%	26%	21%	25%	22%	29%
Page Authority	27%	29%	33%	46%	23%	33%	22%	24%	30%
Alexa Site Links	61	64	60	400	61	16	40	35	38
External Back Links	106, 939	6, 052	3,410	44, 304	5, 393	131	442	208	4, 717
Referring Domains	156	816	175	611	169	57	110	80	84
EDU Back Links	104, 899	182	15	50	4, 172	5	9	9	9
PR Quality	Moderate	Strong	Moderate	Strong	Moderate	Weak	Moderate	Weak	Weak

Institution with higher site links and external back links commands high domain and page authority e.g. Fuoye and Fudutsinma have strong Page Rank Quality. This shows that the institutions have good contents available within the website to attract external linkages.

### C. Traffic

#### 1) Traffic Comparisons

Figure IV shows the institutions with strong page ranks between December 2014 and May 2015. This shows a good link structure management within and outside the universities' website.

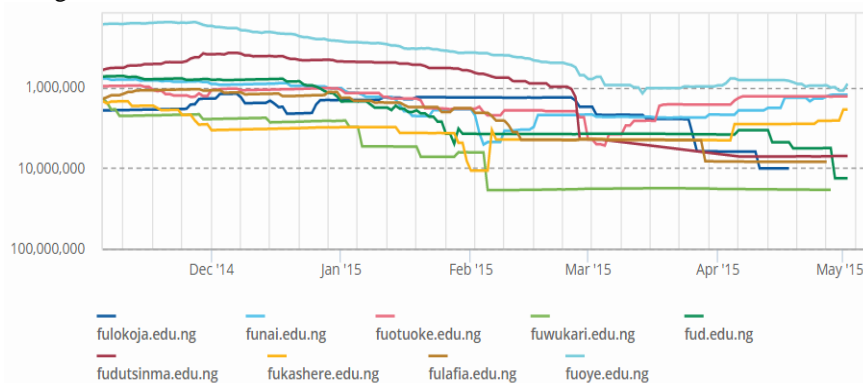


Fig IV: Traffic Comparisons of the Institutions

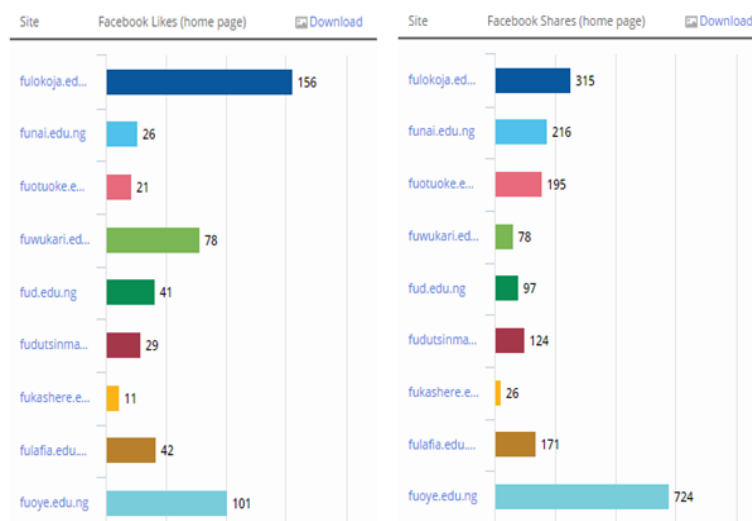


Fig V: Social Network Traffic

Figure V presents the social networks traffic of the institutions using Facebook. All the institutions are on Facebook, a major social networking for students. It reveals that Fulokoja had the highest Facebook audience indexed while Fuoye had the highest Facebook engagement. High engagement is achieved by updating the Facebook status message with interesting contents that could appeal to followers' sharing. When posts are shared, the message spreads across the users own audience thereby increasing the posts reach.

2) *Traffic Distributions*

Figures VI to XIV present the traffic distribution of each of the institutions' website. They show the percentages of visitors from direct visits, inlinks, search engines, social networks, institutions' email facilities and online paid advertisements. Those six (6) sources drive traffic to the Universities' website. The following observations may be noted by the distribution:

- Major traffic sources for the institutions came from direct visits which are generated from offline adverts with the URL while the others were from search results from search engines and inlinks from other websites.
- Though all the institutions have a social network presence but little or absolutely no traffic is received from those networks. This means that a majority of the institutions do not add their URL on their posts contents to direct followers to their websites.
- Only two (2) of the institutions (Fuoye and Fuotuoke) are using their University e-mail facilities yet with very little traffic.
- Only Funai had driven traffic from paid advertisements while all the other institutions had not used paid online advertisements.

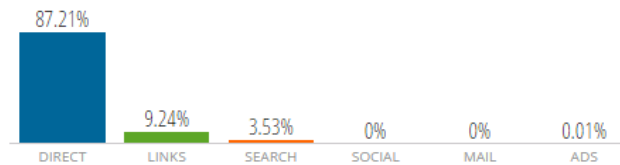


Fig VI: Traffic Distribution by Sources for Funai.edu.ng

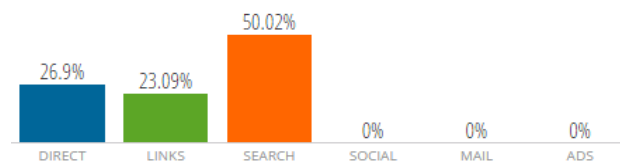


Fig VII: Traffic Distribution by Sources for Fud.edu.ng

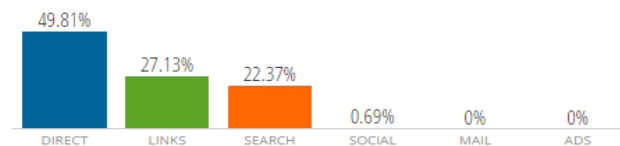


Fig. VIII: Traffic Distribution by Sources for Fudutsinma.edu.ng

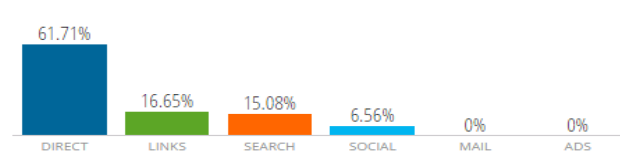


Fig IX: Traffic Distribution by Sources for Fukashere.edu.ng

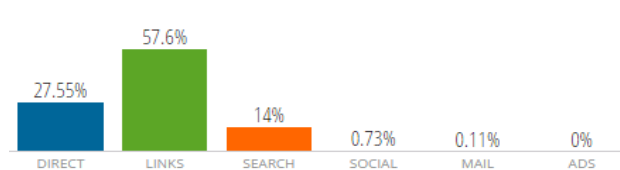


Fig X: Traffic Distribution by Sources for Fuoye.edu.ng

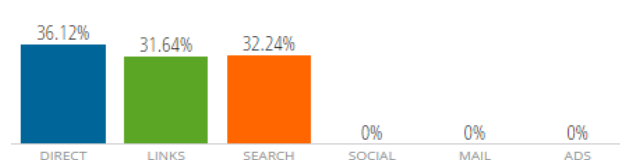


Fig XI: Traffic Distribution by Sources for Fuwukari.edu.ng

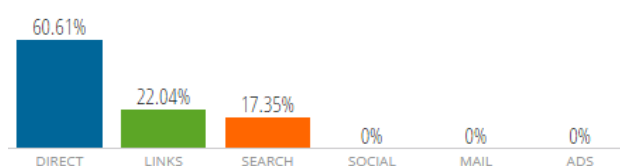


Fig XII: Traffic Distribution by Sources for Fulokoja.edu.ng



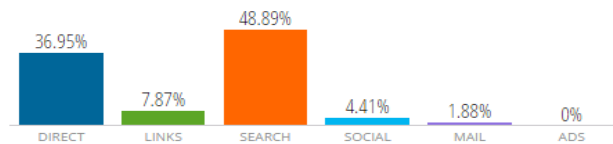


Fig XIII: Traffic Distribution by Sources for Fuotuoke.edu.ng

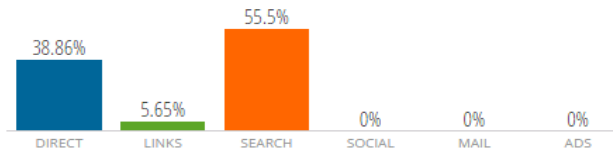


Fig XIV: Traffic Distribution by Sources for Fulafia.edu.ng

#### D. Institutions' Research Output

Research output is determined by the presence of rich files and availability in Google scholar. This takes 30% of the ranking parameters (15% for Rich Files and 15% for Google Scholar). A direct search to determine the rich files count using the classified indicator command *site:domain.tld filetype:pdf* on the institutions' website shows the following result in Table VIII.

Table VIII Rich Files Count Of The Institutions

Domain/ Parameters	fud	fudutsinma	Funai	fuoye	fuotuoke	fukashere	fulokoja	fulafia	fuwukari
DF	57	843	356	2180	84	15	134	113	22
DOC	5	179	107	32	0	0	0	1	1
PPT	2	6	3	72	5	0	0	0	0
PS	6	110	2	195	0	0	0	0	0
Total	70	1138	468	2479	89	15	134	114	23

The Fudutsinma has a well developed e-learning management system (Moodle) installed and all lectures at all levels are taken from there. This contributes to their traffic and high rich files ranking. Likewise, the Funai has DSpace management system installed and their count is rated high too. The Fuoye has a repository hosted within the university website and library management software KOHA which contributes in their highest PDF files count. It was observed that though Fuotuoke had a Moodle installed, but it's not in a good use. Fukashere, Fulokoja, Fulafia, and Fuwukari had no presence of .doc, .ps and .ppt files. The effect of this is that the Google scholar would not find any information to index for their ranking within the site.

Table IX Google Scholar Count Of The Institutions

fud	fudutsinma	Funai	Fuoye	fuotuoke	fukashere	fulokoja	fulafia	Fuwukari
6	66	13	74	16	5	7	13	9

The Table IX presents the Google Scholar counts for the institutions showing for example 66 academic papers within Fudutsinma out of the 1, 138 rich files shown as presented in Table VIII.

### VII. CONCLUSION/RECOMMENDATIONS

We have presented and analyzed data mined from the nine (9) newly established Federal Universities in Nigeria in 2011. We have discussed the results of the mining providing explanations to key criteria that would count for each institution's performance in the webometric ranking come July 2015 and beyond. Based on the findings, we make the following technical and general recommendations:

#### A. Technical

- 1) Appropriate search engine optimization techniques should be used in all the websites; especially good keyword practice should be promoted
- 2) The institutions should embark on paid advertisements to drive required and/or targeted traffic to the site
- 3) Social buttons from the social networks should be incorporated into the institutions websites to enable contents sharing and improved engagements.
- 4) Repositories and e-learning management system should be built within each institutions website from a sub-domain
- 5) The websites should be registered and submitted to relevant web indexing directories and listings.
- 6) The site maps should be properly generated, uploaded on a down-page and submitted to relevant search engines
- 7) All the institutions used content management systems (CMS) for design and development hence they can go a bit further to incorporate mobile versions (apps) because majority of the internet users in Nigeria now accesses the internet through mobile devices

- 8) The CMS should provide a language translation service to help non-English speaking nations accessing the website
- 9) Institutions' journals should be hosted within the university domain with a unique sub-domain

#### **B. General**

- 1) Academic staff should publish their papers with an email address bearing the University domain. Ranking gives credit to the domain whose email the author uses.
- 2) Regular updates of the institutions website with published and ongoing researches should be made
- 3) Academic publications (list and abstracts) should be included in the website with a rich file format.
- 4) Staff and students should be encouraged to use their official emails in the day to day University activities and reduce the use of public domains like Yahoo and Google mails. University should not accept external/public emails.
- 5) Memos, leave applications, notice and minutes of meetings, payslips, etc should go through the website to increase traffic

#### **ACKNOWLEDGEMENT**

The authors are members of the Nigeria Computer Society and academic researchers with their respective institutions. Emeka Ogbuju and Edward Onyebueke Agu are PhD students under the supervisions of Dr. Virginia Ejiofor and Prof. Sunday Eric Adewumi. We acknowledge the efforts of the accessed institutions in ensuring viable online presence.

#### **REFERENCES**

- [1] NUC Monday Bulletin, "FG Approves 3 New Federal Universities", Vol 8 No. 2, January 2013 Edition, ISSN 0795-3089
- [2] A. Abraham, "Business Intelligence from Web Usage Mining", Journal of Information & Knowledge Management, 2003.
- [3] K. Sharma, G. Shrivastava and V. Kumar, "Web Mining: Today and Tomorrow", 3rd International Conference on Electronics Computer Technology (ICECT), 2011
- [4] J. Srivastava, R. Cooley, M. Deshpande and P. Tan, "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", SIGKDD Explorations, 2000.
- [5] C. Almind and P. Ingwersen, "Informetric analyses on the World Wide Web: Methodological approaches to webometrics". Journal of Documentation 53 (4), 1997, pp 404–426.
- [6] T. Hussain, S. Asghar, N. Masood, "Web Usage Mining: A Survey on Preprocessing of Web Log File", International Journal of Computer Information and Emerging Technologies (IJCIET), 2010.
- [7] M. Thelwall, "Introduction to Webometrics: Quantitative Web Research for the Social Sciences", Morgan & Claypool, 2009, ISBN 978-1-59829-993-9
- [8] G. Stermsek and M. Strembeck, "A User Profile Derivation Approach based on Log-File Analysis", IKE 2007.
- [9] L. Björneborn and P. Ingwersen, "Toward a basic framework for webometrics", Journal of the American Society for Information Science and Technology 55 (14), 2004, pp 1216–1227
- [10] O. Osunade and C. Ogundele, "Evaluation of the University of Ibadan Website Using Webometric Ranking Parameters", Transnational Journal of Science and Technology, April 2012, Vol. 2, No.3
- [11] NUC, "'Webometric ranking' of World Universities: Matters Arising" NUC Monday Memo, March 13, 2006
- [12] P. Okebukola, "History, Types and Current Status of University Ranking", NUC Workshop on University Ranking, PPT File Accessed online 12/05/2015
- [13] Page Rank, <http://www.checkpagerank.net/> Accessed 05/05/2015
- [14] Search Engine Optimization Tools, <http://tools.seoachat.com/tools/domain-indexed-pages/> Accessed 28/04/2015