International Journal of Management, IT & Engineering
Vol. 8 Issue 3, March 2018,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

THE IMPACT OF SEARCH ENGINES IN THE WORLD TODAY

Olagoke Olawale Israel*

Olatunji Olusoji Samson*

Abstract

Web search engines have major impact on people's everyday life. It is of great importance to test the retrieval effectiveness of search engines. Successful searching of information on the internet depends on techniques such as going straight to the information source, guess work and developing strategies for when to use subject directories and search engines. The utility of the web may match the skills of a professional reference librarian because it possesses an ever-changing and extremely heterogeneous document collection of immense proportions. Experiments on major search engines show that the approach to web research mines many high-confidence rules that help understand search engines and detect suspicious search results.

Keywords: World Wide Web, Information Retrieval, Indexing, Search Engines

* *Mass Communication Department The Polytechnic, Ibadan, Ibadan, Nigeria*

1. INTRODUCTION:

The World Wide Web (Web) was invented in 1989 and the World Wide Web Consortium (W3C) was established in 1994 to lead the World Wide Web to its full potential. By the turn of the century the Web had entered most aspects of our lives from communication to e-Government, e-Commerce and e-Learning, making it much more than just an information repository.

Sir Tim Berners-Lee invented the World Wide Web in 1989 while working as a software engineer at CERN, the large particle physics laboratory near Geneva, Switzerland. With many scientists participating in experiments at CERN and returning to their laboratories around the world, these scientists were eager to exchange data and results but had difficulties doing so. Tim understood this need, and understood the unrealized potential of millions of computers connected together through the Internet. Tim Berners-Lee, inventor of the Web and Director of the World Wide Web Consortium (W3C), is regularly cited for saying "The power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect" and more recently "One Web for anyone, everywhere on anything" – this is all part of the Web's "full potential".

In 1999 the W3C Web Accessibility Initiative (WAI) published the first set of international guidelines for Web accessibility, the Web Content Accessibility Guidelines 1.0 (WCAG), documenting the essential requirements for Web content to be accessible to people with disabilities. Accessibility requirements for authoring tools (ATAG) and user agents UAAG), including browsers followed. Traditional Web search engines mostly adopt a keyword based approach. When the keyword submitted by the user is ambiguous, search result usually consists of documents related to various meanings of the keyword, while the user is probably interested in only one of them.

With the advent of the World Wide Web (Web), a new category of searching now presents itself. The Web has had a major impact on society (Lynch, 1997) and comes the closest in terms of capabilities to realizing the goal of the Memex. In terms of quality, Zumalt and Pasicznyuk (1998) show that the utility of the Web may now match that of the skills a professional reference

librarian. The Web possesses an ever- changing and extremely heterogeneous document collection of immense proportions. . Although developed in an apparently unstructured environment, Web document discovery is extremely structured in terms of its hyperlinks. The user population of the Web is enormous and extremely diverse, albeit with certain groups over represented (Hoffman, Kalsbeek, & Novak, 1996; NTIA, 1999).

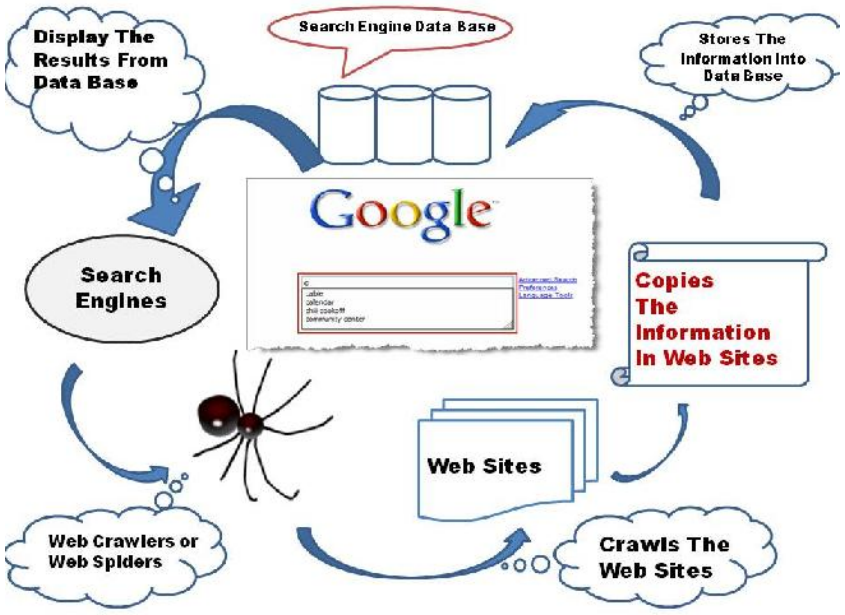

Figure 1: Process of Information searching through search engines

The Web's Information Retrieval (IR) systems are unique in terms of the interface, advertising constrains, bandwidth restrictions, and unique document indexing issues (e.g., spamming and URL hijacking). In sum, the Web appears to be a whole new searching environment (Sparck-Jones & Willett, 1997).

According to Clay & Esparza (2009), search engines are programs that search web pages or documents for a particular keywords or where the keywords were found. A Search engine is really a

collection of programs; however, the term is often used to specifically describe systems like Google, Bing and Yahoo! Search that enable users to search for documents on the World Wide Web. The components and tasks of web search engines, Crawling or Spidering is an automated process to gather the data with web spiders. They can be pictured as little spiders and are also known as crawlers, robots, software agents, web agents, wanderers, walkers, or know bots.

Spiders process the web page and give us information. The web pages are found by them by URL which is given by a web page holder to notify their web page, or through hypertext links embedded in most web pages (Sherman & Price, 2001).

In the latter case, spiders start by crawling a few web pages and follow the links on those pages. After fetching the pages they point to, they follow the links that are on the last pages. The same process will be continued until they have indexed a certain part of the web that includes pages they store across many machines, what leads to the next task. Indexing is the second part of search engines. It is the process of ―taking the raw data and categorizing it, removing duplicate information, and generally organizing it all into an accessible structure. The stored full-text indexes of the crawled web pages are organized in a database, typically in an inverted index data structure. It is ultimate for keyword based queries, so that documents that include the typed keywords can be quickly retrieved. Webmasters have taken many advantages of the web, especially for business commitments. A lot of power will be put into search engine optimization (SEO) or maximizing search engine visibility, online marketing

strategies (Clay & Esparza, 2009).

Ngwuchukwu (2009) noted in a study on ICT use for knowledge societies that some staff higher institutions who are also postgraduate students lack basic ICT skills which will help in achieving good results in the university and when using a search engine to do research.

2. LITERATURE REVIEW:

He and Göker (2000) conducted a study using 51,474 queries from an Excite log and 9,534 queries from a local version of the Alta Vista search engine. The researchers focused on identifying a time interval that could be utilized to specify a session. Based on the analysis, the

researchers concluded that a time interval of 10 to 15 minutes was the typically session length. There has been few studies that focus on Web sessions.

The study by Bifet et al. (2005) used many different factors in an estimation function derived for the ranking function of a search engine; with this function they compare their own predicted rankings with the actual rankings of Google. Although these studies obtain a number of interesting results, they are in themselves not decisive for many existing Search Optimization (SEO) factors.

Evans (2007) noted that search engines constantly work to improve their ranking calculations. As a result the calculated relevance of webpages has been observed to have varied over time, due to changes in the relative weights assigned to individual SEO factors, as well as the incorporation of new factors and the elimination or modification of others. In the case of Google, for instance, the identification of more than 200 factors which have varied over the lifespan of the search engine has, thus far, prevented the definition of a precise method by which the high ranking of a website on the search engine could be guaranteed.

Zimmer (2010) noted that when discussing the search engine market, it is often forgotten that while search engines are usually commercial enterprises, they also serve as facilitators of information and therefore, they serve the interests of the public.

Riemer and Bruggeman (2009) see search engine research at the crossroads between the design-science paradigm and the behavioural-science paradigm. An integrated approach would consider both and this would lead to a better understanding of existing systems and to the design of better systems in the future.

Primary and Secondary Web-Searching

Once read that the average person living in a modern industrialized society is exposed to as many different pieces of information in a single day as a person living 100 years ago would have seen in a year. That includes advertisements, newspaper headlines, websites, text messages, traffic signs, T-shirt slogans, and on and on and on. It's hardly surprising that attention spans are getting

shorter and that the majority of people believe themselves to be busier than ever. With this information overload, it is next to impossible to remember everything we need to, to call up names, dates, figures, phone numbers, email addresses and all the corporate and client information we need to do business effectively. That's why we use tools to do the remembering and information retrieval for us.

Many companies use Salesforce.com to handle the bulk of their customer relationship management information and also use Microsoft Outlook to manage their emails. When they want to find a product, service or piece of information online, they make use of a Search Engine. In the month of March 2006 alone, there were 6.4 billion searches. Assuming each user looks at an average of two search results pages, each of which displays 10 search results, that gives an average of 128 billion search results shown to Internet users in a single month. Search Engines are ubiquitous, and so accepted in contemporary culture that the word "Google" now appears in the dictionary as verb (as in "to Google something").

Search Engines essentially act as filters for the wealth of information available on the Internet. They allow users to quickly and easily find information that is of genuine interest or value to them, without the need to wade through numerous irrelevant web pages. The goal of the Search Engines is to provide users with search results that lead to relevant information on high-quality websites. The operative word here is "relevant". To attain and retain market share in online searches, Search Engines need to make sure they deliver results that are relevant to what their users search for. They do this by maintaining databases of web pages, which they develop by using automated programs known as "spiders" or "robots" to collect information. The Search Engines use complex algorithms to assess websites and web pages and assign them a ranking for relevant search phrases. These algorithms are jealously guarded and frequently updated. Google looks at over 200 different metrics when assessing websites, including copy, in-bound links, and website usability and information architecture. What this means is that the Search Engines provide users with the information they are looking for, and not necessarily the information that marketers would like them to see. Type the name of a major brand into Google, and you will most probably be served a wide range of search results that include not only the official website of the brand you searched for, but also other websites, consumer review sites, Blogs, online articles on Web 2.0 sites and press releases on news syndication channels. Of course, not all

searches are for brand names. The majority of searches are for non-brand key phrases - for example, "Hong Kong luxury hotel" rather than "The Peninsula Hong Kong". With key phrases that are service or product-specific rather than brand-specific, results pages will also include many competitors, which makes acquiring a prominent position at the top of the page even more crucial.

There are two major ways to make sure a website appears in a prominent location on the major Search Engines for relevant key phrases: Paid Search (also known as Pay-Per-Click) and Organic Search Engine Optimization. Of the two, Organic Search Engine Optimization tends to yield the best long-term results and the optimum return on investment, for the simple reason that Internet users are four times as likely to click an Organic search result as they are a Pay-Per-Click ad on the same results page. In a September 2006 poll by Marketing Sherpa, 68.7% of marketers in the US identified Search Engine Optimization as yielding the best Return on Investment for product marketing. I will discuss Paid and Organic search in much more depth in a separate article. It is enough here to state that companies doing business or marketing online should look at striking a healthy balance of both techniques to make the most of the potential of marketing through the major Search Engines.

Secondary Web-searching Studies, is a review of Web searching studies that are more limited in scope in that they do not present enough data to give a full picture of Web searching. Most of these studies analyzed Web searching on a singular Web site that was not a search engine.

3. MOST POPULAR SEARCH ENGINES

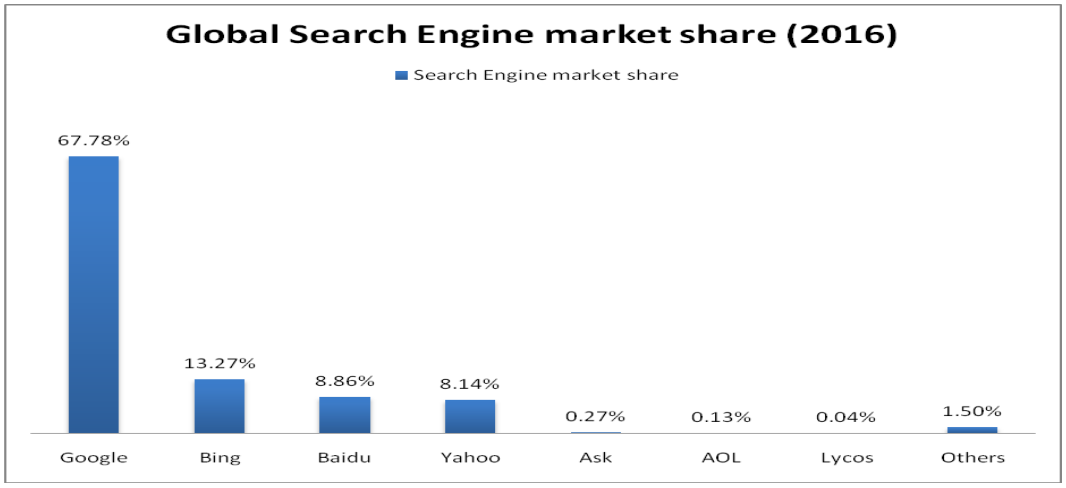

Figure 2: Bar chart of global search engine market share

Source: Net Market Share, (2016).

Knowing which search engines are getting the largest percentage of search traffic plays a big role in deciding the focus of optimization efforts. Google search statistics, for example, show us that Google is still the king of search traffic, accounting for 66.52% of all search traffic in July 2013. Bing and Yahoo! follow further behind with 11.40% and 8.40% respectively, while Ask is at 4.14% and AOL Search is at 1.84%.

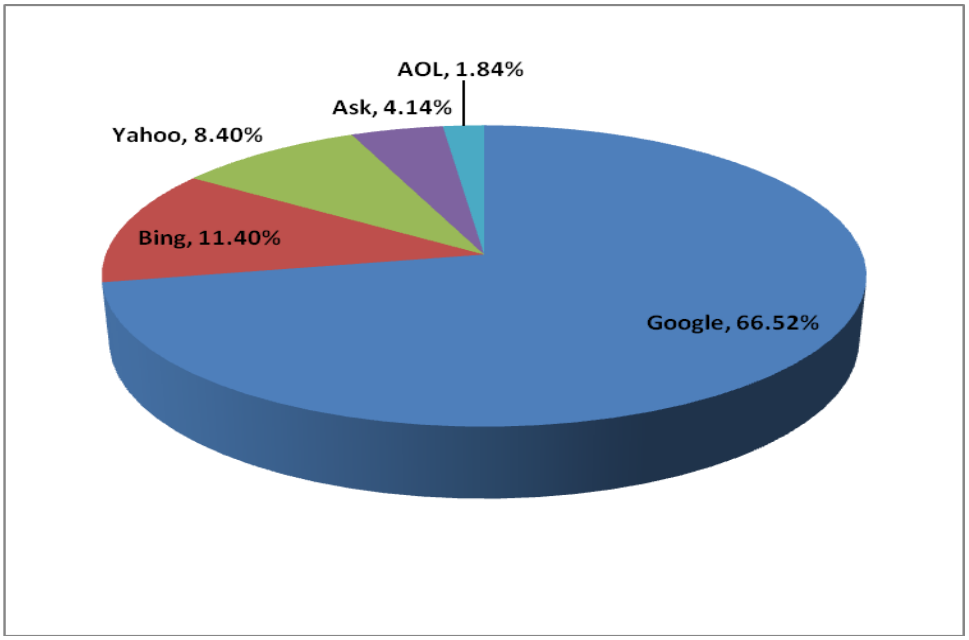

Figure 3: Pie chart of search engine usage (2013)

Most users want a single search engine that delivers three key features:

1. Relevant results (results you are actually interested in)
2. Uncluttered, easy to read interface
3. Helpful options to broaden or tighten a search

With these criteria, 5 Reader Favourite Search Engines come to mind. These 5 search sites should meet 99% of the searching needs of a regular everyday user.

1. Google

Google is the undisputed king of 'spartan searching'. While it doesn't offer all the shopping center features of Yahoo!, Google is fast, relevant, and the largest single catalogue of Web pages

available today. Make sure you try the Google 'images', 'maps' and 'news' features, they are outstanding services for locating photos, geographic directions, and news headlines.

2. Yahoo

Yahoo! is several things: it is a search engine, a news aggregator, a shopping center, an emailbox, a travel directory, a horoscope and games center, and more. This 'web portal' breadth of choice makes this a very helpful site for Internet beginners. Searching the Web should also be about discovery and exploration, and Yahoo! delivers that in wholesale quantities.

3. Bing

Bing is Microsoft's attempt at unseating Google. Bing used to be MSN search until it was updated in summer of 2009. Touted as a 'decision engine', Bing tries to support your researching by offering suggestions in the leftmost column, while also giving you various search options across the top of the screen. Things like 'wiki' suggestions, 'visual search', and 'related searches' might be very useful to you. Bing is not dethroning Google in the near future, no. But Bing is definitely worth trying.

4. Ask

The Ask/AJ/Ask Jeeves search engine is a long-time name in the World Wide Web. The super-clean interface rivals the other major search engines, and the search options are as good as Google or Bing or DuckDuckGo. The results groupings are what really make Ask.com stand out. The presentation is arguably cleaner and easier to read than Google or Yahoo! or Bing, and the results groups seem to be more relevant. Decide for yourself if you agree... give Ask.com a whirl, and compare it to the other search engines you like.

5. Webopedia

Webopedia is one of the most useful websites on the World Wide Web. Webopedia is an encyclopedic resource dedicated to searching techno terminology and computer definitions. Teach yourself what 'domain name system' is, or teach yourself what 'DDRAM' means on your computer. Webopedia is absolutely a perfect resource for non-technical people to make more sense of the computers around them.

4. SEARCH ENGINE OPTIMIZATION (SEO)

Search engine optimization (SEO) is the practice of optimizing a Web site so as to achieve preferred ranking on the search engine results pages (SERPs). Someone who practices SEO professionally is also known as an SEO (search engine optimizer).

Search Engine Optimation (SEO) is the process of improving the volume or quality of traffic to a website from search engines. The way websites come to be the most searched falls in the hands of the web developer. The web developer must keep in mind the SEO rules when creating any particular site. If the rules are neglected or purposely broken, to gain a competitive advantage, the site can be permanently shut down. The best way to become the most searched site is to follow the convention set forth by the highly skilled SEO web developer. Some of well known rules include the following.

i. Use keywords when describing the product or service of the particular website being developed. Usually, there will be multiple spots within the website to describe the product or service. The more descriptive the better, the web crawler picks up these keywords so the more descriptive keywords there are the better the chance is that the web crawler grabs the site.

ii. Use plenty of links to outside sources. Including links to other websites away from your site sounds counter intuitive, it is not. This is the best method of improving traffic to a particular site.

Technology is one of the fastest growing and ever-changing markets and since the web is highly technical its trends and developments should be paid close attention to. Although there are many avenues in web, the developments in SEO are especially important. In order for web developers to stay current on the rules of SEO they must research and use the information gained to be able to predict where the future of the web industry and SEO is headed. Currently HTML-based websites are the only types that can be optimized for search engines but the future may allow for newer computer coding languages such as Flash and JavaScript to be searchable as well.

The basic understanding of SEO starts with understanding how a search engine works. There are three basic types of web search engines; crawler-based search engines, human powered

directories, and hybrid search engines or mixed results. SEO only applies to the crawler-based search engine, which is what the larger search engines, Google, and Yahoo are using.

The spiders are programmed to looks for document changes to a site over time. It works by "crawling" around a website scanning the content. It to stores the content in the index and then follows the links to other pages within the site to repeat the same process. It returns about every month or two to check for changes that it can then store in the index. This is what is meant by a website being "spidered" or "crawled.

The index is where the content from the crawled site is stored. Much like a library that stores books, it collects and stores information to assist a fast and accurate retrieval later. For a website to be searchable it must be indexed first. The process of "crawling" may take months before the website is indexed. Once the website is indexed it is searchable using the search engine and the process of "crawling" continues and indexes any updates or changes made to the site. This is a process that Google does on its own without outside influence and speeding up the process cannot be requested (searchenginewatch.com).Search engine software sifts through millions of indexed sites to find matches to a search. It then ranks the searches in order based on relevance.

5. CONCLUSION

Based on this study, it is concluded that successful searching of information on the internet depends on techniques such as going straight to the information source, guess work and developing strategies for when to use subject directories and search engines. By indexing a target Web page more accurately, and allowing each user to perform more fine-grained search that satisfy his/her information need. Also, Search Engine Optimization is constantly changing as new aspects come into play, and others go. Fundamentally, businesses need to recognize the two crucial fundamentals for high-quality SEO, on-page and off-page. On-page SEO refers to what a publisher can control directly, whereas off-page SEO basically relies on user behaviour, social engagements, visitors and other publishers. Finally, the ability to search and locate information is very important skill for researchers and search engine users; the number of information on the internet has been arranged by experts to help researchers so there is need for its full utilization by these researchers.

6. ACKNOWLEDGEMENTS:

The authors appreciate all lecturers in the Department of Mass Communication,The Polytechnic,Ibadan for their unflinching support throughout this research paper.

7. REFERENCES:

Bifet, A., Castillo, C., Chirita, P-A. and Weber, I. (2005), "An analysis of factors used in search engine ranking", Proceedings of the Workshop on Adversarial IR on the Web, Chiba, Japan, 10-14 May, available at: http://airweb.cse.lehigh.edu/2005/bifet.pdf (accessed 1 December 2012).

Clay, B. and Esparza, S. (2009). "Building an SEO friendly site", in Clay, B. and Esparza, S. (Eds),

Search Engine Optimization All-in-One for Dummies, John Wiley & Sons, Hoboken, NJ, pp. 215-240.

Evans, M.P. (2007), "Analyzing Google rankings through search engine optimization data", Internet Research Journal, Vol. 17 No. 1, pp. 21-37.

He, D. & Göker, A. (2000). Detecting session boundaries from Web user logs. Paper presented at 22nd Annual Colloquium of IR Research. April 5 - 7, 2000, Cambridge UK.

Hoffman, D., Kalsbeek, W., & Novak, T. (1996). Internet and Web use in the U.S. Communications of the ACM, 39 (12), 106-108.

Lynch, C. (1997). Searching the Internet. Scientific American, 276(3), 52-56.

Ngwuchukwu, M. N. (2009). Information Communication Technology (ICT) Competencies among Library and Information Science Educators in two Universities in South East Zone of Nigeria. Global Review of Library and Information Science vol. 5.

NTIA. (1999). Defining the digital divide, the 3rd annual report by the National Telecommunications and Information Administration. Retrieved from the World Wide Web on 20 August 1999 from http://www.ntia.doc.gov/ntiahome/digitaldivide/.

Riemer, K., & Brüggemann, F. (2009). Personalisierung der Internetsuche - Lösungstechniken und Marktüberblick. In D. Lewandowski (Ed.), Handbuch InternetSuchmaschinen (pp. 148-171). Heidelberg: Akademische Verlagsgesellschaft Aka.

Sherman, C., & Price, G. (2001). The Invisible Web: Uncovering Information Sources Search Engines Can't See. Medford, NJ: Information Today.

Zimmer, M. (2010). Web search studies: Multidisciplinary perspectives on web search engines. In J. Hunsinger, L. Klastrup, & M. Allen (Eds.), International Handbook of Internet Research (pp. 507-521). Dordrecht: Springer.

Zumalt, J. & Pasicznyuk, R. (1998). The Internet and Reference Services: A Real-World Test of Internet Utility. Reference and User Services Quarterly, 38(2), 165 – 172.

AUTHOR'S BRIEF BIOGRAPHY:

Mr. Olagoke Olawale Israel: He is a lecturer at the department of Mass Communication, the Polytechnic, Ibadan. He holds a Master of Communication Arts degree from the University of Ibadan and Higher National Diploma (HND) in Mass Communication from the Polytechnic of Ibadan, Nigeria. He was a reporter of the Nigerian Tribune Newspaper and Federal Radio Corporation (FRCN), Ibadan, Nigeria. He is a member of the Nigerian Institute of Public Relations (NIPR) and Nigeria Union of Journalist (NUJ).

.

Mr. Olatunji Olusoji Samson: He is a lecturer at the department of Mass Communication, the Polytechnic, Ibadan. He holds a Master of Communication Arts degree from the University of Ibadan. He is a member of the Nigerian Institute of Public Relations (NIPR) and member, African Council for Communication Education (ACCE).