7-4-2019

Site Metrics Study of Koha OPAC through Open Web Analytics and Piwik Tools

Sukumar Mandal

Department of Library and Information Science, The University of Burdwan, sukumar.mandal5@gmail.com

Site Metrics Study of Koha OPAC through Open Web Analytics and Piwik Tools

Dr. Sukumar Mandal
Assistant Professor, Department of Library and Information Science
The University of Burdwan, Burdwan – 713 104
Email: sukumar.mandal5@gmail.com

Abstract

The extensive and comprehensive statistics of different users' characteristics and behaviour on a particular website has perform by using the Web analytics which offer increased visibility of website and greater user satisfaction. Measurement of website usage give a suitable and relevant source of user-centric information about the popularity of a library OPAC. The behavior of library visitors needs more attention so that a website can become a more effective, user satisfaction, library related activities and functions, and library users display statistics and it also provide site administrators a number of web metrics that relate users' activities on a individual library online public access catalogue. Information mashup and cloud computing are the major elements in the field of Web analytics. It is very important because increasing the library information services through open source software Koha in integrated library management and retrieval system. Web analytics gives many facilities for the websites metrics to the visitors and to improve the better organization of information resources. Web analytics have monitoring and evaluation of websites usage and statistics of different libraries. This is very important to the library professionals for easy and simple visual graph in different levels. Client server architecture is an important aspects in integrated library management and retrieval system. Actually client server architecture is a part of cloud computing. The prime objective of any library is to proper dissemination of information resources among the users, but the problem of suitable sites measures for viewing the library OPAC by different users. In this paper have select the two important open source software such as Piwik and Open Web Analytics. Apart from this it has also successfully integrate the Koha OPAC with Piwik and Open Web Analytics user interfaces.

Keywords : Web Analytics, Koha OPAC, Piwik, Open Web Analytics, and Bounce rate

Introduction

Internet is an important aspects in the field of library and information services. It has been rapidly growing the in this field. Most of the libraries have its own website and webpages which helps to proper management of library resources is very systematically and logically such as identification, collection, organization, and dissemination of useful information and knowledge. Web analytics is the popular method to analyzed the web pages of different sectors of library websites because it collects a large amount of data from the library OPAC or websites which would depends on certain factors such as screen size, Internet speed, types of browser, types of visitors, and etc. Generally web analytics or analytics have been gathering and measuring the data both the online and offline data directly or indirectly. The important features of web analytics are building block terms, visit characterization, visitor characterization, engagment, conversion which is aggregated and compiled the web related information for the users as well as library professionals. The integration of Koha OPAC with the helpf of Open Web Analytics and Piwik interface is very simple because its gives the javascript widget facilities. Tracking this code for displaying the different library web OPAC through these two web analytics tool which allows the bounce tracking and return visitors tracking. It is possible to easy measurement and optimizing the library web pages relating with online public access catalogue. It is not only measure the web traffic but it extends in the field of library and information services for increasing the usage of demand in library Web OPAC. Integrated library system is consists of two major elements such as housekeeping operations and information retrieval system. In this paper only analyzed the second part of ILS because its measure the scientifically and logically of end user satisfaction. Web analytics is also part of a cloud computing or information mashup. Koha is a popular web based software and fully support the information mashup features for analyzing and integration of library OPAC with Web Analytics Tools. The important tools of this paper are as follows :

Piwik	:	It is a well known popular web analytics open source software written in PHP high level programming language. Track Key consists of Performance Indicators such as visits, goal conversion rates, downloads, keywords, and other relevant elements (https://piwik.org).
Open Web Analytics	:	Open Web Analytics is also an important open source web analytics software written in PHP and developed by Peter Adams. It is support MySQL database connectivity for running the various web server (http://www.openwebanalytics.com/).

The goal of web analytics is to collect and analyze data regarding the usage patterns and web traffic and this can be comes from four sources such as (a) Direct HTTP request data (Directly comes from HTTP request messages); (b) Network level and server generated data (Not part of an HTTP request, but it is required for successful request transmissions.); (c) Application level data sent with HTTP requests (Generated and processed by application level programs, including session and referrals. These are usually captured by internal logs rather than public web analytics services.) (d) External data (It can be combined with on-site data to help augment the website behavior data described above and interpret web usage). It can be easily measure the uniques visitors to a library OPAC page in a specific time interval and display the sequence of library web page views in a single or multiple set. The visit can start when a patron first page views on the website and ends after few seconds of inactivity. A library OPAC page view happens when a patron views a web page of library web site and variable is defined after a specific time.

Review of Related Works

"Web analytics has become an essential part of every online marketer's toolkit. But you can't just rely on the flood of data alone—you need to interpret it, and in many cases, fine-tune reports to accurately reflect your own goals and objectives. The second edition of Brian Clifton's Advanced Web Metrics with Google Analytics is a comprehensive roadmap to helping you get the most from your metrics—an indispensable guide to helping you take your online marketing campaigns to the next level".

---- *Chris Sherman*, Executive Editor, Search Engine Land

Katuu in 2016 describe the web visitors from the Mandela Portal and explores the data regarding the portal information and resources. Taraghi & others in 2013 have discuss the two original contributions such as the analysis of user path tracing and a novel algorithm which performs efficient and effective integration of new articles of different journals into the existing comments and recommendations, because the fact that scientific journals have been published in a frequent and on the regular basis of time sequence. In 2017 Obrien, 2017 and others were discussed the present potential data that begin to detail the deficiencies of log file analytics reporting methods those are commonly built into institutional digital repository repository system and they have strongly proposed a innovative method for collecting and reporting of different items in the form of download statistics and metrics. They have also introduces a web enabled service prototype which captures activity that current analytics methods are likely to either miss. Plaza in 2009 is to develop a new user-friendly in-house tracking methodology for academics to analyse the effectiveness and smooth visits of websites depending on their traffic source such as direct visits, referring site entries and search engine visits. Showers & Stone in 2014 have discussed the shared analytics service for academic libraries to inform decision making a top priority in the next five years. The essential objective of this paper is to explore the effective and efficient shared analytics service for UK academic libraries and introduce the JISC Library Analytics and Metrics Project. Anna & Barbara in 2013 have explore the Web 2.0 concept through open source LibraryThing which allows users to catalogue books using data drawn from sources such as Amazon and the Library of Congress in different search facilities such as tagging and interest groups and this study tries to evaluate whether LibraryThing is a valuable tool for academic libraries. Clegg in 2014 has discussed the Web Analytics Strategies for information professionals. Klobas in 2014 is explores to measures of online open course success for institutional providers of massive open online courses (MOOCs) and other scaleable open online courses (SOOCs). Clegg & Evans in 2017 have discussed the big data and text analytics of unstructured data through visualization tools and it also manage the taxonomy of different terms. Academic library websites consists of a huge amount of information resources, often contributed using a large number of content creators with varying levels of technical expertise in the Penn State University Libraries site contains almost 10,000 pages contributed by over 200 content creators from all areas of the Libraries (Lush, 2015). This paper aims to describe the use of web statistics by libraries, archives and museums in the Netherlands (Voorbij, 2010). Chaurasia & Rosin, 2017 have explores and examine the applicability of Big Data in higher education institutions on the basis of competitive advantage and data complexity in majou four application areas were identified for the use of Big Data in higher education which consists of reporting and compliance; analysis and visualization; security and risk mitigation; and predictive analytics.

"Advanced Web Metrics is a unique book that combines high-level management advice and nitty-gritty detail in an easy to understand and, above all, useful way. It's great for web managers, analytics specialists, and marketers alike."

----- Dan Drury, Director, Bowen Craggs & Co., and Author of the
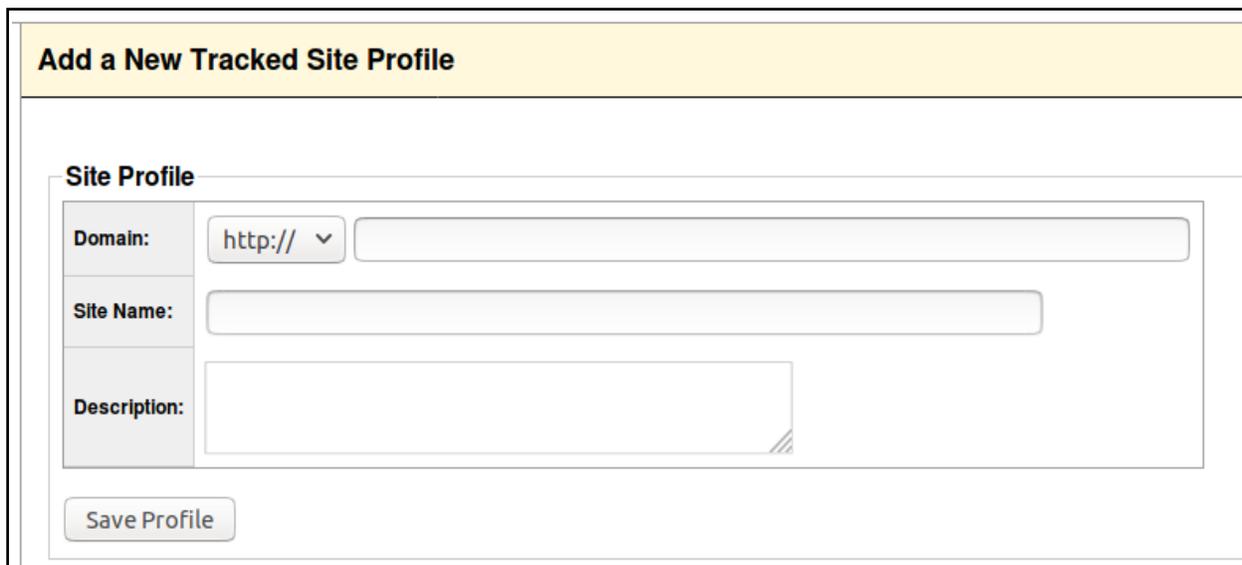Financial Times Index of Corporate Website Effectiveness

Objectives

The objectives of this paper are as follows :

(i)   To explore the web analytics for display the page statistics and visitors visit.

(ii)   To integrate the Koha OPAC visit statistics in Piwik and Open Web Analytics.

Tracked Site Interface

Virtual pageviews is the new concept in web application. Soingle Page Application is also an important concept in Web analytics for tracking the different sites. It is helps to load the relevant web resources available both the offline and online environment with different components such as page, subsequesnt content, uniform resource locator to manage the conventional page navigation except the full page request. It is dynamically measure the number of page views on a particular website. The tracked sites interface is represent in the Figure – 1 for open web analytics. By clicking on the option of add new site and it will automatically create the drop down menu of respective website of a particular web resources.

Figure – 1 : Tracked site interface in Open Web Analytics

List of Website

Website can be classified in major two categories such as static and dynamic. Dynamic site is a part of Web 2.0 community level for sharing the information resources between the site owner and site visitors. On the other hand static sites generally show the information. It is possible to direct access the information and multiple web resources. Dashboard is a core element in any website which helps to display the list of website after adding the relevant sites through the admin interface in Open Web Analytics tool. This dashboard is also display the site metrics of multiple and single pages in different statistics like visits, unique visitors, page views, average visit duration, bounce rate, pages per visit for efficient and effective statistics in an individual website. The Figure – 2 is represent the list of website in dashboard in Open Web Analytics.
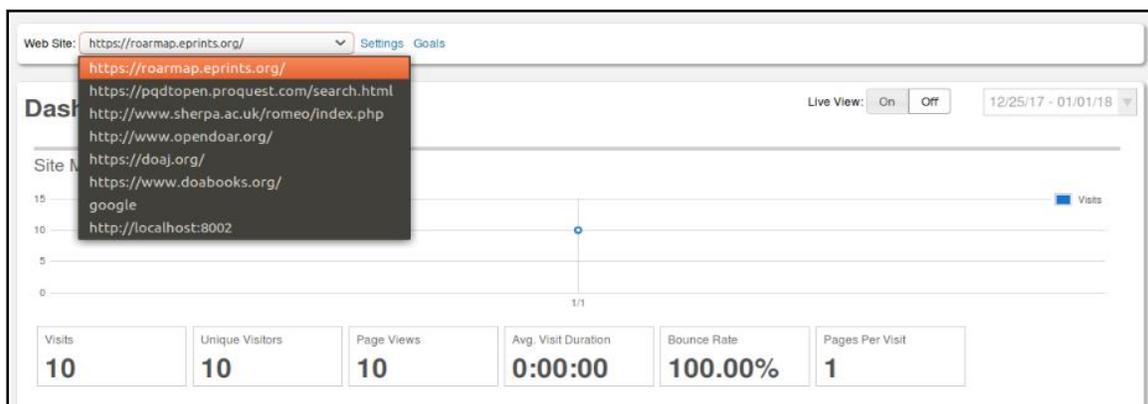
Figure – 2 : List of Website interface

There are many people have working in any institution or in any libraries. They can create their own user identification under the main categories in the level of general parameters. Create any website is very fashionable functions in different sectors such as personal website, commercial website, educational website, Government website, library websites, and etc. There are many websites have belongs to the particular area on a topic, entertainment, social networking, updated news and education. Websites can be accessed in two ways such as global and local through open source tools and techniques. The Figure – 3 is represents the add new user of Open Web Analytics for opening, closing and customizing the user interface by a particular librarians or library professionals.
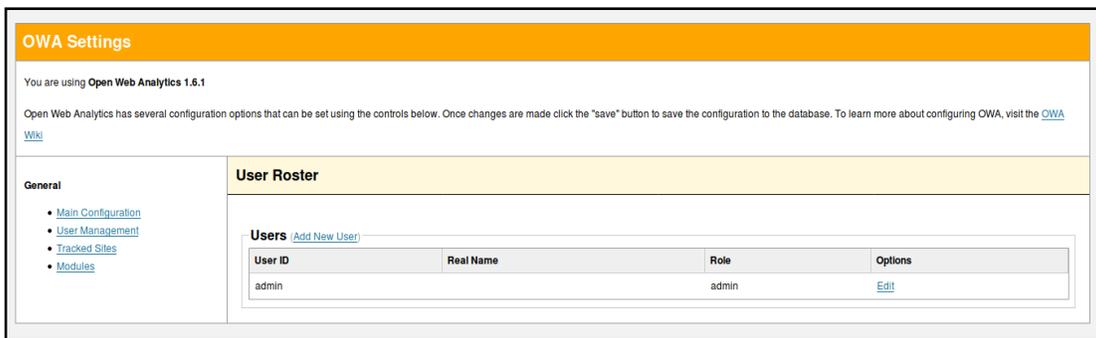

Figure – 3 : Add user interface

Koha Web Analytics

Koha is popular web based open source software in integrated library management and retrieval system. Integration of open web analytics with Koha is very simple and interesting to the users. In this section just copy the java widget script and paste into the opac credits in Koha under the global system preferences. The Koha Web analytics is represents in the Figure – 4 after click on option of save button under the global system preference settings. This is very user friendly for display the web pages and statistics through graph visualization in different components such as page views, visits, unique page views.

Figure – 4 : Koha web analytics interface
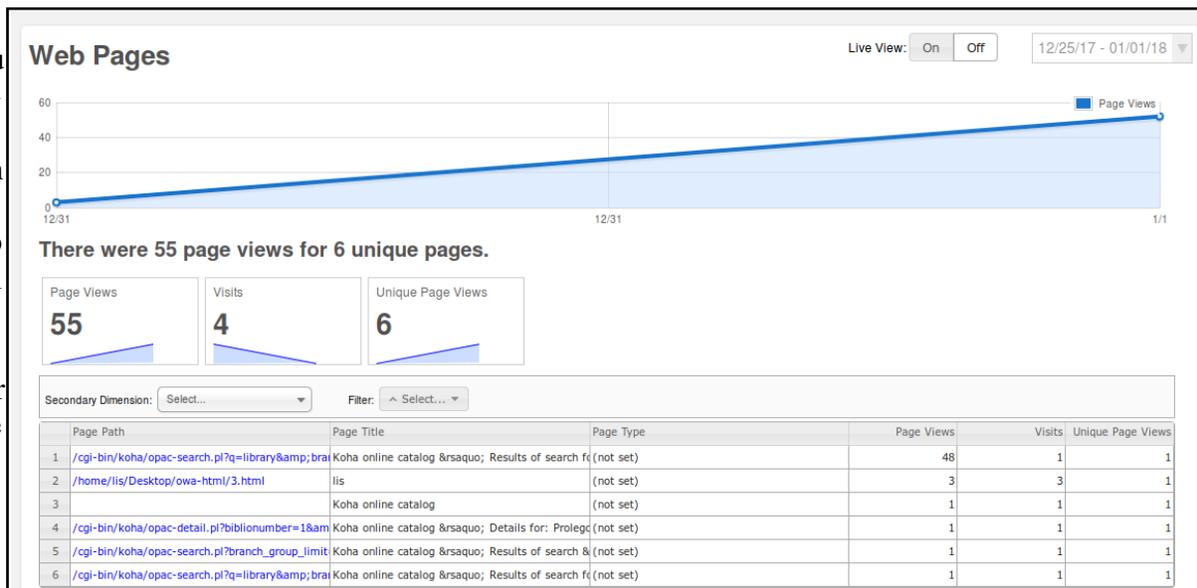
Piwik Web Analytics from Koha

Link analysis is one of the popular web 2.0 features in semantic web. Web analytics is also calculated from the Koha OPAC interface by using Piwik open source software. This is the powerful analytics interface of Piwik. Web analytics can be measure in two ways from Piwik such as statistics and graph visualization. The Figure – 5 is represents the web analytics interface of Koha web OPAC for visits and actions of a particular facets with proper link.

Figure – 5 : Piwik web analytics interface from Koha OPAC
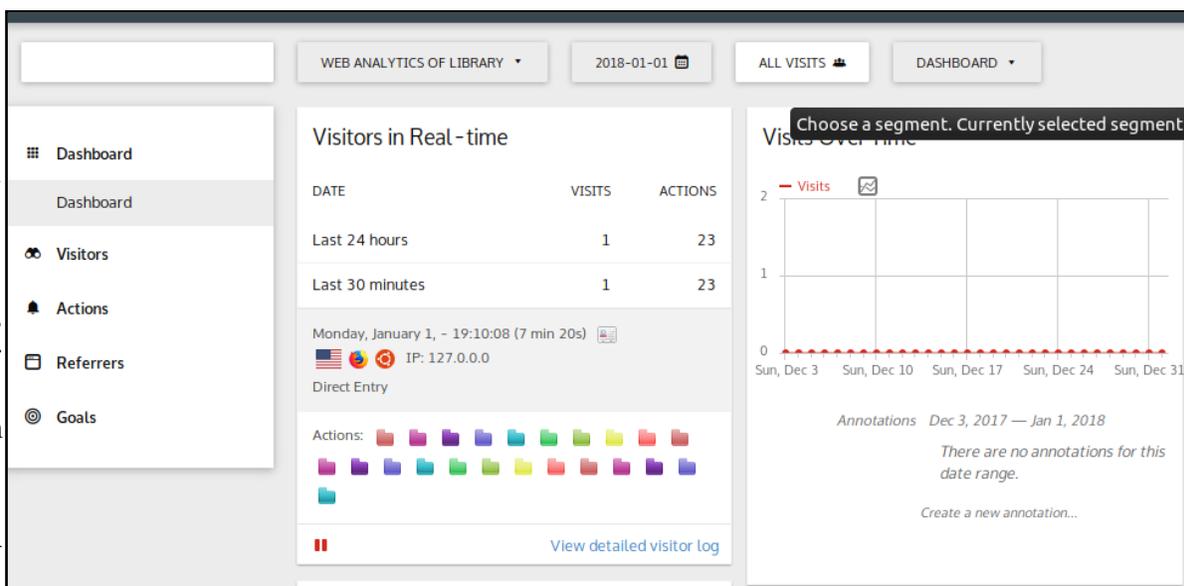
Visitors Log of Koha in Piwik

Web analytics is also give the visitor log of any web enabled interface and its the most advanced and matured level web analytics tool in web environment. It is calculate the real time analytics which display the invidual visitor records such as the total amount of time the visitor spent on Koha Library OPAC, the total number of page views and visits, the total conversions and the total conversions for each goal, the internal search keywords used by the visitor, and the average amount of time it took library servers to generate a page for the visitor. This is very much helpful to the library professionals because the web analytics is totally free service in-depth statistics of different visitors. It can be analyzed and measure the data from different sources such as search engines, traffic sources, referral sources on a particualr website. The Figure-6 is represents the log visitors of Koha OPAC through Piwik web analytics tool.

Figure – 6 : Log visitors of Koha

Conclusion

Dashboard consists of five elements such as content, action tracking, visitors, traffic, and goals. Further it can be devided as separately like content merge with different sub elements such as top pages, page types, feeds, entry pages, and exit pages. In the field of action tracking it can be sub devided in different fields such as action groups. Visitors includes different sub parameters in open web analytics such as geo-location, domains, visitor loyalty, visitor recency, visitor age, browser types, operating systems. Traffic fields has consists of many sub-parameters such as search terms, inbound link text, search engines, referring websites, campaigns, add performance, add types, creative performance, and attribution history. Goals has include only one fields such as funnel visualization. Site metrics reports is very essential in Web analytics. It is possible to display the site metrics of Koha Library OPAC through the Open Web Analytics and Piwik. Display the trend report and number of times library OPAC pages were viewed by different users for a spacific time period and visitors visit the library OPAC pages in more than one. This report is easily allows users to track page views for every library OPAC page and bibliographic search fields and aggregate of web page views for library OPAC entire site. Generally this visit is a sequence of web page views in library OPAC. It can be easily calculate the enhances use of library OPAC by different visitors. This metrics provide an effective and efficient process for users to quickly display the visibility and impact of bibliographic records in library OPAC and sketch the recent visibility of visitors against the author, title, subject and so many fields.

References

Chaurasia, S. S., & Rosin, A. F. (2017). From Big Data to Big Impact: analytics for teaching and learning in higher education. *Industrial and Commercial Training*, *49* (7/8), 321–328. https://doi.org/10.1108/ICT-10-2016-0069

Clegg, H., & Evans, H. (n.d.). Big Data and Text Analytics. In *The Emerald Handbook of Modern Information Management* (pp. 791–809). https://doi.org/10.1108/978-1-78714-525-220171033

Katuu, S. (2016). The Mandela Portal – an assessment of global visitors using Web analytics. *Digital Library Perspectives*, *32*(4), 270–286. https://doi.org/10.1108/DLP-01-2016-0002

Klobas, J. E. (2014). Measuring the success of scaleable open online courses. *Performance Measurement and Metrics*, *15*(3), 145–162. https://doi.org/10.1108/PMM-10-2014-0036

Lush, B. (n.d.). Managing Accessible Library Web Content. In *Accessibility for Persons with Disabilities and the Inclusive Future of Libraries* (pp. 169–189). https://doi.org/10.1108/S0065-283020150000040017

OBrien, P., Arlitsch, K., Mixter, J., Wheeler, J., & Sterman, L. B. (2017). RAMP – the Repository Analytics and Metrics Portal: A prototype web service that accurately counts item downloads from institutional repositories. *Library Hi Tech*, *35*(1), 144–158. https://doi.org/10.1108/LHT-11-2016-0122

Open Web Analysis (2018). About Open Web Analysis. Retrieved from http://www.openwebanalytics.com/. (Accessed on January 06, 2018)

Piwik (2018). About Piwik. Retrieved from https://piwik.org/. (Accessed on January 06, 2018)

Plaza, B. (2009). Monitoring web traffic source effectiveness with Google Analytics: An experiment with time series. *Aslib Proceedings*, *61*(5), 474–482. https://doi.org/10.1108/00012530910989625

Richards, A., & Sen, B. (2013). An investigation into the viability of LibraryThing for promotional and user engagement purposes in libraries. *Library Hi Tech*, *31*(3), 493–519. https://doi.org/ 10.1108 /LHT-03-2013-0034

Showers, B., & Stone, G. (2014). Safety in numbers: developing a shared analytics service for academic libraries. *Performance Measurement and Metrics*, *15*(1/2), 13–22. https://doi.org/ 10.1108/PMM-03-2014-0008

Taraghi, B., Grossegger, M., Ebner, M., & Holzinger, A. (2013). Web analytics of user path tracing and a novel algorithm for generating recommendations in Open Journal Systems. *Online Information Review*, *37*(5), 672–691. https://doi.org/10.1108/OIR-09-2012-0152

Voorbij, H. (2010). The use of web statistics in cultural heritage institutions. *Performance Measurement and Metrics*, *11*(3), 266–279. https://doi.org/10.1108/14678041011098541

Web Analytics Strategies For Information Professionals. (2014). *Library Management*, *35*(6/7), 527–529. https://doi.org/10.1108/LM-06-2014-0065