Application of Google Analytics Model for Evaluating the Visibility of Library Web Portals of the Uva Wellassa University, Sri Lanka

T. Pratheepan[1] and M. Jayakananthan[2]
[1]Uva Wellassa University, Sri Lanka
[2]Eastern University, Sri Lanka
E-mail: pratheepan@uwu.ac.lk, jeyakananthanm@esn.ac.lk

Abstract - **Google Analytics is the most prevalent page tagging technique to evaluate the visibility of web portals. The study was designed with the purpose of applying google analytics model on the library Web Portals (WP) of the Uva Wellassa University (UWU) which includes, Home Page (HP), Online Public Access Catalogue (OPAC) and Institutional Repository (IR). It was observed that the web resources of UWU have been viewed by ~366756 local & global users during the study period. The USA is the predominant country visited the UWU e. Repository with the count of 53,078 (15.82%) followed by Netherlands with 14,044 (4.78%) and France 15,775 (4.70%).**
Keywords: **Digital Library, Google Analytics, Library Web, Web Metrics**

I. INTRODUCTION

Advanced technologies and content delivery methods on the web have evolved in the world. Along with this context universities and other higher educational institutions are the essential focal point in providing formal and informal educational materials to the community. Designing and operating digital online tools require more technical support and decision-making skills for the developer in the education system. Since 2017 the Library of UWU has launched many online digitized services to the university community, such as Integrated Library Management System (ILMS), E- Repositories, Library Website for the library users. These services are built by considering published and unpublished scholarly work, text books, theses, past exam papers and other digital collection. Measuring the success or failure of the tool or a service is the key factor to determine its future expansion and popularity among the users.

In a study Sofos & Kostas, 2009 stated there are various administrative methods and procedures that can be applied to analyse and evaluate web based educational resources and its content, ranging from international quality standards to specific controls lists. However, these methods require human intervention and effort. Due to their subjective nature, it cannot guarantee accurate result and it is time-consuming. Barba *et al.,* 2013 pointed in a study that, academic libraries typically use a number of assessment tools including surveys, usability studies, and Web analytics.

While both surveys and usability studies have their place, Web analytics have proven extremely versatile. Web analytics is widely used method that collects, perform analysis, generate reports and visualize the web metrics data (Loftus, 2012). Also web analytics generates detailed reports about the visits to a website, and which is friendly application and tells the web owner how visitors found the site and how they interact with it. By Comparing the visitors behaviour, insight into how to improve the site content and design (Plaza, 2011) can be inferred. Web analytics is an approach that involves collecting, monitoring, analyzing and reporting web usage data to understand the visitors' experiences. Analytics can help to accomplish the goals and improve customer satisfactions and loyalty (Bekavac & Garbin, 2015).

"Google Analytics is the enterprise-class web analytics solution that gives you rich insights into your web traffic and marketing effectiveness. Powerful, flexible and easy-to-use features now let you see and analyze your traffic data in an entirely new way. With Google Analytics, you're more prepared to write better-targeted ads, strengthen your marketing initiatives and create higher converting websites." (Google, 2012).

In 2005 Google purchased Urchin Software Company and subsequently released Google Analysis to the public 2006 as a free web analytics tool. It is a valuable tool for those who need to determine their websites performance in a fast and reliable way (Manoj, 2006). Anyone can use Google Analytics tool with Google account. More importantly, no need of any specialized programming knowledge or experts to configure the GA account. Simply create GA account, set the web site profile to be tracked, and then embed the system generated HTML code in to web site. Within minutes GA can be functional and can reap the benefits.

II. OBJECTIVES OF THE STUDY

The prime objective of the study is to explore the potential of Google Analytics (GA) for investigating the usability behaviour of the library web portal. The specific objective of the study is to:

1. Investigate the visitor's behaviour.
2. Explore the visitor acquisition behaviour.
3. Investigate the content access behaviour.

III. METHODOLOGY

Google Analytics is the most prevalent page tagging technique found in the literature and the study directed to choose Web analytics study using page tagging GA tool. In order to use GA tool to track the usage of the library's online resources, Google Analytics account created by the administrator. After creating the GA account, the administrative interface was enabled. Then, the interface, separate track code and script for each and every online

resource have been embedded. The automatically generated GA code and script has been added to the online resource pages respectively. This study was analysed with the help of data generated from the Google Analytics console from January 2019 to September 2020 (21 Months). In this study, three important online resource interfaces of the UWU library were included in the data collection which includes Library Web, OPAC (Koha) and eRepository (DSpace) during the study period.
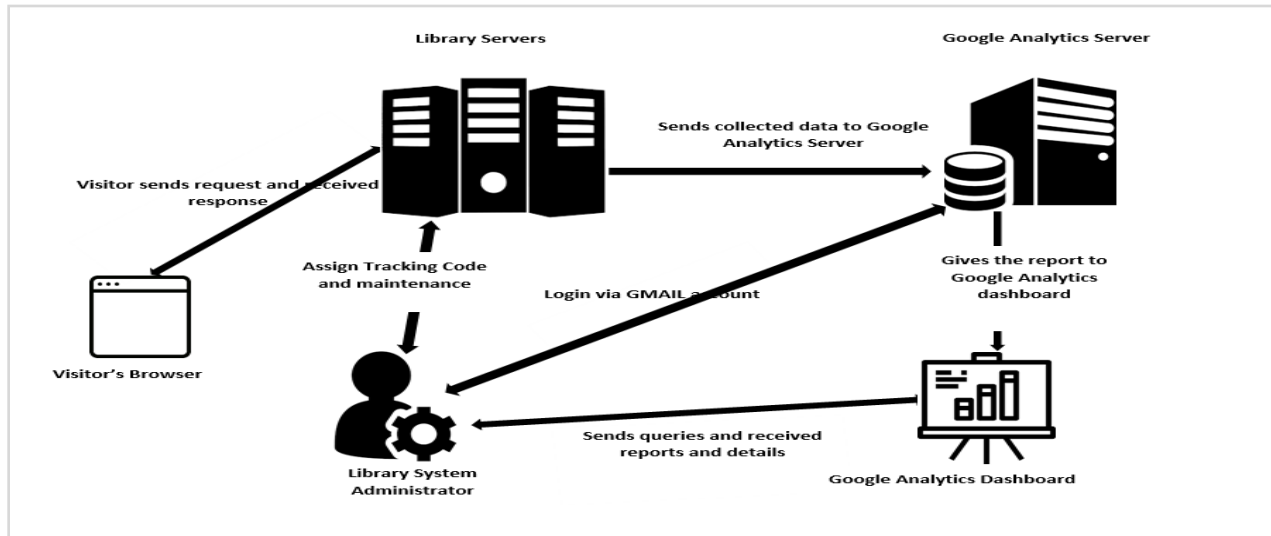

Fig. 1 Google Analytics Process Diagram

Generally, there are two type of web analytics methods used by experts in the field. The first one uses server based log files, and second is client based page tagging (Kaushik, 2007). Most of the earlier studies used server-based log file for web analytics and employed varies metrics to evaluate the web site. In most of the cases, they cover four areas: content, navigation, accessibility and design. However, inaccuracies of using log-files as data source were identified by both web analytics vendor and customer. This led to the page-tagging approach as new data source for collecting data from web site (Hasan, Morris, & Probets, 2009).

The page-tagging technique involves adding few line of script into the web pages and collects the data from them. Page-tagging method is far better when compared with web server log-file, because several reasons for this, one is that most page-tags are based on the cookies to indentify the visitor and not on the IP address, another is that non human user agents are excluded from the measurement and reporting (Engervall, 2017).

There are web analytics tools such as WebTrends, VisitStat, StatCounter, ClickTracks and Google Analytics (GA). Among these GA is the page-tagging approach web analytics tool and which had a major effect on the web analytics industry now days (Demirel Kutukcu, 2010).

Dyrli, 2006 in a study pointed GA is the most sophisticated web analytics tool for collecting web log data. The following table I shows the methods for embedding GA scripts in to the online resource systems.

TABLE I ONLINE RESOURCES AND EMBEDDING GOOGLE ANALYTICS CODE AND JAVASCRIPT

Sl. No.	Online Resource Interfaces	Tracking Method
1	OPACs of Integrated Library Systems (Koha)	OPAC -- Koha > Administration > System Preferences > OPAC >opaccredits
2	e Repository (eR) Software (Dsapce)	Add the web property ID to the dspace.cfg parameter xmlui.google .analytics.key.
3	Web Development Software (Joomla)	Joomla --- Extensions Directory under the Template Management—Custom Code

The GA provides set of factors under following general categories, i.e. audience, acquisition, behaviour, content & conversations. However, three of them, i.e. audience, acquisition, behaviour provide the data which are useful to address the research question of this study. The following Table II indicating required categories and metrics used for this study.

TABLE II GOOGLE ANALYTICS MAIN CATEGORIES AND SUB METRICS

Main Category	Metrics
Audience	Demographics
	Visits & Location
	Behaviour
	Technology
	Devices
Acquisition	Traffic Sources
	Search Console
	Social
Content	Site content
	Site speed
	Site search

IV. ANALYSIS OF AUDIENCE - WEB METRICS

The audience report present data on the number and location of people visiting the website. The number of visitors is recorded in several forms, including the number of visits, the number of unique visits, the number of page visits, and the average number of pages per visit. The audience report tends to give a good indication of the frequency and location of the visitors, to determine the reachability of the work being communicated through the website.

TABLE III DEMOGRAPHICS OF THE VISITORS

Metrics	Sub Metrics		Library Web (%)	OPAC (%)	eRepository (%)
Demographics	Male		35.55	46.38	41.75
	Female		64.45	53.62	58.25
	Age	18-24	10.66	17.08	15.94
		25-34	39.91	38.31	35.77
		35-44	25.09	25.53	27.70
		45-54	12.81	8.79	9.29
		55-64	6.76	7.63	7.81
		65+	4.77	2.66	3.49

According to the data that Google Analytics provides (Table III), the library online resources are mostly used by Female (Average 58.77%). And the significantly visits age groups are 25-34, 35-44 and 18-24. It clearly shows that the UWU academics and final year students have more attraction on the library online resources.

TABLE IV VISITS OVERVIEW

Particulars	Library Web	OPAC	IR	Total/Average
Users	2203	821	4028	7052
New users	2225	819	4046	7090
Sessions	3375	1302	10655	15332
Page Views	9284	5191	19680	34155
Pages/Session	2.75	3.99	1.86	2.86667
Avg. Session Duration	00:02:52	00:03:11	00:38:58	00:15:00
Bounce Rate	48.30%	48.62%	27.07%	41.33%
Top 5 Countries	Sri Lanka	Sri Lanka	USA	
	USA	India	Netherlands	
	India	China	Sri Lanka	
	South Korea	South Korea	Germany	
	Malaysia	USA	France	

According to the Visitors Overview presented in Table IV, during the study period of analysis, library online resources have 34,155-page views and the average 2.87 pages/visit. And totally 7052 users and 7090 new users have visited all the online resources. The average time spent by user on the online resources was 15 minutes, while the bounce rate was 41.33%.

It is clear that the e Repository (eR) has more impact among the students and academicians of the UWU. And the most of the users are from the USA Followed by Netherland & Sri Lanka. Surprisingly, Europe and US users have visited many times towards library online resources. The pattern of regular users and new users visits also similar as shown in the Fig. 2.
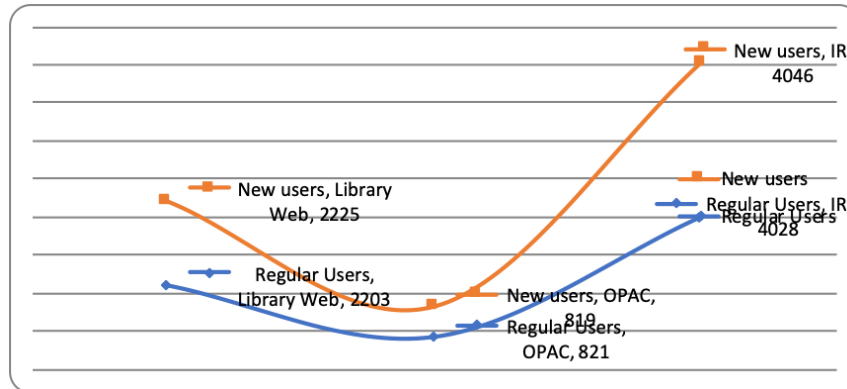

Fig. 2 New and returning users

The Table V indicates the visitor total frequency. The most of the users have visited once on the Library Web and OPAC. The reason behind this is the frequent visitors were familiar with online contents so they did not need to visit as many pages to find what they were looking for. They could also bookmark the specific pages and eR has more visits by the audience.

minutes). This implies majority of visitors spent longer time on the library online resources.

TABLE VI ENGAGEMENT VISIT DURATION

Duration	Library Web	OPAC	IR
Visit Duration	Page Views	Page Views	Page Views
0-10 seconds	2,008	721	3,245
11-30 seconds	502	165	875
31-60 seconds	686	274	1,134
61-180 seconds	1,654	943	3,071
181-600 seconds	2,123	1,376	3,570
601-1800 seconds	1,651	1,070	2,444
1801+ seconds	660	642	5,341

According to the Fig. 3, Chrome was the most popular web browser and library online resources received 52% of visits sent through the Chrome. Firefox was the second most popular browser and owned 21% visits. Safari, Edge and IE are not much popular.

TABLE V FREQUENCY OF VISITS

No. of Visits	Library Web	OPAC	eRepository
Count of Visits	Visits	Visits	Visits
1	2226	819	4051
2	358	76	469
3	178	39	124
4	92	24	49
5	71	17	27
6	51	14	20
7	34	14	15
8	27	14	14
9-14	111	56	67
15-25	107	72	98
26-50	82	110	127
51-100	38	47	60
101-200	0	0	100
201+	0	0	5344

The Table VI shows that data about how user engages with library online contents and how many pages they viewed. The majority of visits stay on 181-600 seconds (3-10
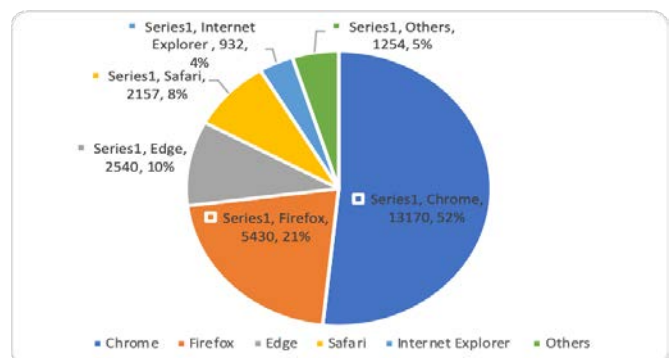

Fig. 3 Most popular web browsers used by visitors

The following Table VII shows that the most library online resource users mostly using Windows operating system shared 62.25%. The rest of the operating systems are mobile and Linux.

TABLE VII OPERATING SYSTEM USED BY VISITORS

Operating Systems	Percentage
Windows	62.25%
Android	30.18%
iOS	5.21%
Linux	0.77%
Macintosh	0.55%
Windows Phone	0.12%
Tizen	0.07%
Samsung	0.05%
BlackBerry	0.02%

The Lanka Education and Research Network (LEARN) is the leading internet service provider of the visitors. It is clearly shown that most of the visitors are from the educational institutions. And another 23% of visits came through the Dialog Axiata others Mobital and ADSL shared rest of the visits.
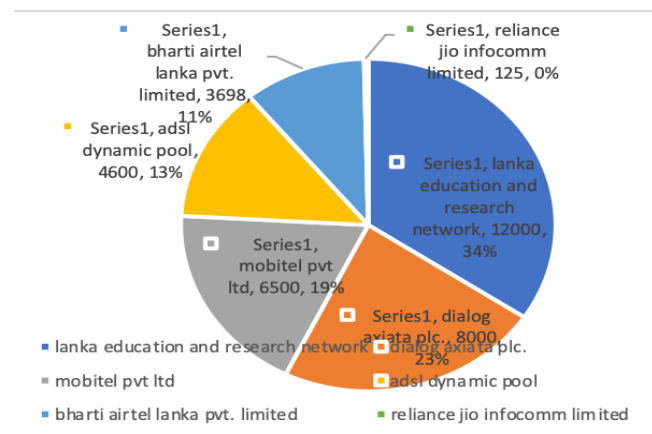

Fig. 4 Visitors Internet Service Provider

Table VIII, shows that the only 2.83% visits were made through the tablet and 65.13% of them were regular laptop and desktop computers

TABLE VIII DEVICES USED BY VISITORS

Devices	Percentage (%)
Desktop	65.13
Mobile	32.04
Tablet	2.83

V. ANALYSIS OF ACQUISITION - WEB METRICS

The traffic source report reveals data on the visits resulting from: search engines, referrals from other websites, directly entering the web address, and from search keywords. The traffic source overview indicates the pathways that were used by the visitors to access the website, as some of the details regarding search term access are not available.

Fig. 5 illustrates the different ways that visitors have accessed the library's online resources. During the 21 months research period, the Direct access was the most common method in connecting to the online resources, with above half of the visitors (50%) access the library online resources either by typing the URL into a browser, or by clicking on a bookmark, or following a hyperlink in an email or other electronic documents. Another 28.66% of visits were made from the Search engines, such as Google, Yahoo, Bing, etc. And 12.70% of visits were made through the referral sources. This type of visitors accessed through the links on external party servers. Above results clearly implies that majority of the visitors had enough knowledge about the library online resources.
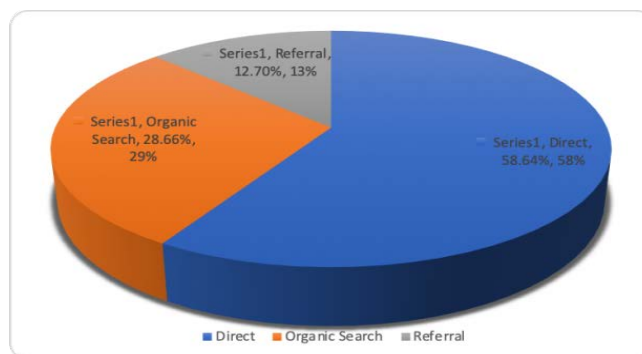

Fig. 5 Traffic Sources of the Library's portal

The top direct access pages were the library's home page, eR, and OPAC other direct access pages. Mobile versions of library home page, eR and OPAC were also having significant direst access by the visitors. The top referral source to the library's digital contents is UWU Library main web page. The social media facebook library page also had a decent number of referral visits to the library online contents. Among the search engine traffic sources, Google was the preferred search engine that visitors have used. And the top keywords were used by the visitors through the search engines include "Uva Wellasa University", "UWU library", "Uva Wellassa University online library", "opac library UWU" and "digital library UWU".

VI. ANALYSIS OF CONTENT WEB METRICS

GA's Content category tracking the visitors' access related to the content available in the library online portal. According to the Table IX, the library web home page received more visits by the visitors, while e-resources shared 32.80% accessed by visitors. Rest of the visits includes online catalogue, library profile, rules and periodical collection pages. The most of the visitors did not rely on library rules pages because Average Time on Pages is very low and Bounce rate is very high.

TABLE IX LIBRARY HOME WEB PAGE CONTENT ACCESS

Content/Pages	Page Views (%)	Average Time on Pages	Bounce rate
Home	41.89%	00:01:29	44.10%
e-Resources	32.80%	00:01:36	52.00%
OnlineCatalogue	19.37%	00:02:57	58.33%
Online Journals	2.43%	00:01:48	43.75%
General Rules	1.90%	00:00:13	81.38%
JournalSelection Tools	0.71%	00:00:52	49.02%
Open Access	0.66%	00:01:14	41.67%
Other pages	0.24%	00:00:31	92.00%

Table X illustrates the OPAC contents accessed by the visitors. The OPAC home page made the most visits and advanced search, author search is hold significant visits by the users. Further Table X shows the top bibliographic data accessed by the users according to the Dewey Decimal Classification System.

TABLE X TOP CONTENTS USED IN THE OPAC

Sl.No.	Content/Pages
1	OPAC - Home
2	Advanced Search
3	Author Search Results
4	Search Results - 630 Agriculture & Related subjects
5	Search Results - 549 Mineral Science
6	Search Results - 658 Management
7	Search Results - 338 Tourism Management & Related Subjects
8	Search Results - 664 Food Science

TABLE XI TOP CONTENT ACCESSED BY VISITORS ON E-REPOSITORY

Sl.No.	Content/Pages
1	Impact of Social Media Marketing Activities on Consumer Buying Behaviour for Casual Dining Restaurants in Sri Lanka
2	The Impact of Brand Exposure through Brand Ambassador on Consumer Purchasing Intention in Sri Lanka (With Special Reference to Millennial Audience)
3	Assessment of women centered agricultural extension programmes in Balangoda Agriculture zone in Sri Lanka
4	Colour Development of Geuda Stones after the Gas-Fired and Electrically Operated Furnace Heat Treatments
5	Analysis of Bioactive Properties of Fish Protein Hydrolysates from Scomber japonicus Fin Wastes
6	Recognition of Vowels for Sri Lankan Traditional Pirith Chanting Using Formant Variation
7	Big Data: A Scientometric Analysis Based on Indian Publications
8	Organic Fertilizer Producers in Sri Lanka: An Assessment of Constraints along the Value Chain
9	Factors Affecting Customer satisfaction in Micro Life Insurance an Assessment with Special Reference to Micro Life Insurance Policyholders in Uva Province
10	Job Satisfaction Among Public Library Staff in Trincomalee District

The eR of UWU consist of Thesis, Journal articles, Conference papers and Past papers from the faculties and units. According to the Table XI, digital theses collection was the most accessed online resources by the visitors. Specifically, students' undergraduate reports and postgraduate theses shared most visits. Further, article and research papers were mostly identified by the visitors from the eR.

The access of online digital resources should be considered with loading and downloading time. Table 12, illustrates the average loading time of page, redirection, domain lookup, server response and downloading time of the library web, OPAC and IR. According to the data the OPAC performed very quickly and eR, library web also have very decent time period for visitor usage.

TABLE XII AVERAGE CONTENT ACCESS TIME

Time in Sec	Library Web	OPAC	IR
Avg. Page Load Time (sec)	6.89	5.84	5.99
Avg. Redirection Time (sec)	0.11	0.04	0.02
Avg. Domain Lookup Time (sec)	<0.01	<0.01	0.12
Avg. Server Connection Time (sec)	<0.01	0.04	0.36
Avg. Server Response Time (sec)	0.66	0.34	0.74
Avg. Page Download Time (sec)	0.16	<0.01	0.22

VII. DISCUSSION

The aim of the research was to explore the potential of Google Analytics (GA) for investigating the behaviours of the online visitors to the library web based resources. Based on the objectives the previous results section covers all required basic and advanced metrics including Demographics, Visits & Location, Technology, Devices, Traffic Sources, Search Console, Social, Site content and Site speed.

The students and academics from UWU were the major users in this study and approximately 6786 members. GA recorded 77184-page views based on the data during the study period. This implies UWU members and outside community have significant attention about the library online resources. Further, the study revealed that the library online resources have created much attention to attract user around the globe. GA demographics do not allow depth analysis of users but data imply, the academics and final year students were have more awareness about library web based resources. In an article Morgan, 2010 stated that between 1-5 minutes can be estimated on a site with exception of social network like the Facebook, the library web based resources were doing a good job in keeping the users spend a considerable time on it. The electronic Repository has the strong engagement with visitors.

Therefore, it is confirmed that the library web-based resources consist targeted contents for the education community. It is vital to know that the new visitor and returning visitor interacts with library web-based resources. GA reported both new visitor and returning visitor have small different with their accessed pattern. This implies the library online resources kept the regular user as well as attract the new user to capture the online resources. The bounce rate can be defined as the amount of time visitors visits the site and then immediately leave it without completing any tasks. This implies visitor spent little time and no interactions with site. The high bounce rate indicated poor design, low usability, loading time high and weak traffic sources. GA reported the average bounce rate is 41%, so it is confirmed the library web-based resources contents were designed very attractive, have high usage and more prone to leave the library portal. Also, the library portal visitors connecting through latest web browser, decent network and used desktop, mobile as well.

Regards to the GA traffic sources report the visitors have enough attention about the library portal and they accessed directly or already saved the address favourites. This also suggests that email and other awareness campaign succeeded in promoting the library portal. Further, GA reported there are numerous amounts of traffic sources from the search engines, social network and other external referral links.

Some keywords also identified which are used by visitors in order to identify the library portal. This could be very much useful for web admin to develop future expansion of the portal. The contents of the library portal seems to achieve its main purpose because GA reported there are high amount of visits to the e-resources page, OPAC page, services page, etc. Also indicated OPAC and eR were accessed required documents downloaded by visitors for their educational purpose.

VIII. LIMITATIONS OF THE STUDY

GA evaluation used in this study had many limitations. Because the visitor can hide their actual data with configuring network setup of their own server/browser. And analysis about the least percentage of visitors' metrics has not been included in the study. So, the study left some information about the minor groups of the visitors. However, the study counts all quantitative data according to the visitors' interaction through the network.

IX. CONCLUSION

In general, GA has proved its potential in investigating the usability of the library web portal with respect to the audience behaviour, traffic sources and content. The research identified specific web metrics that can be used quick, easy and cheap quantitative data about the usability of the library web portal. In some cases, those metrics will

help to identified problems on the specific resource pages from the web portal. Further, this study recommends librarians and information professionals to use the GA open source tool to investigate the effectiveness and impact of the web portal and to upgrade them based on the usage.

REFERENCES

[1] Barba, I., Cassidy, R., De Leon, E., & Williams, B. J. (2013). Web Analytics Reveal User Behavior: TTU Libraries' Experience with Google Analytics. *Journal of Web Librarianship*, *7*(4), 389–400. Retrieved August 21, 2018 from https://doi.org/10.1080/19322909.2013.828991.

[2] Bekavac, I., & Garbin Praničević, D. (2015). Web analytics tools and web metrics tools: An overview and comparative analysis. *Croatian Operational Research Review*, *6*(2), 373–386. Retrieved September 11, 2018 from https://doi.org/10.17535/crorr.2015.0029.

[3] Chande, S. (2015). *Google Analytics -Case study*. Retrieved August 25, 2018 from https://www.optimizesmart.com/page-tracking-in-google-analytics/

[4] Coates, M. (2013). Using google analytics to explore ETDs use. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, (January 2012), 437–438. Retrieved October 10, 2018 from https://doi.org/10.1145/2467696.2467770.

[5] Crutzen, R., Roosjen, J. L., & Poelman, J. (2013). Using google analytics as a process evaluation method for internet-delivered interventions: An example on sexual health. *Health Promotion International*, *28*(1), 36–42. Retrieved August 15, 2018 from https://doi.org/10.1093/heapro/das008.

[6] Demirel Kutukcu, S. (2010). *Using Google analytics and think-aloud study for improving the information architecture of METU informatics institute website: A case study*. Retrieved October 03, 2018 from https://etd.lib.metu.edu.tr/upload/12612584/index.pdf.

[7] Engervall, E. (2017). *A method to empirically investigate design manipulations*. Retrieved October 03, 2018 http://uu.diva-portal.org/smash/record.jsf?pid=diva2:1139283.

[8] Fagan, J. C. (2014). The suitability of web analytics key performance indicators in the academic library environment. *Journal of Academic Librarianship*, *40*(1), 25–34. Retrieved september 24, 2018 from https://doi.org/10.1016/j.acalib.2013.06.005.

[9] Fang, W. (2007). Using Google Analytics for Improving Library Website Content and Design: A Case Study. *Library Philosophy and Practice*, *9*(2), 22. Retrieved October 13, 2018 from https://doi.org/10.1002/14651858.CD004937.

[10] Hasan, L., Morris, A., & Probets, S. (2009). Using google analytics to evaluate the usability of e-commerce sites. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5619 LNCS, pp. 697–706). Retrieved August 25, 2018 from https://doi.org/10.1007/978-3-642-02806-9_81.

[11] Kaushik, A. (2007). *Web analytics: an hour a day. Manager*. Retrieved from Retrieved October 04, 2018 from http://www.amazon.com/Web-Analytics-Hour-Avinash-Kaushik/dp/0470130652.

[12] Loftus, W. (2012). Demonstrating success: Web analytics and continuous improvement. *Journal of Web Librarianship*, *6*(1), 45–55. Retrieved October 19, 2018 from https://doi.org/10.1080/19322909.2012.651416.

[13] McGuckin, C., & Crowley, N. (2012). Using Google analytics to evaluate the impact of the cyber training project. *Cyberpsychology, Behavior, and Social Networking*. Retrieved October 12, 2018 from https://doi.org/http://dx.doi.org/10.1089/cyber.2011.0460.

[14] Morgan, J. (2010). Google Analytics for Libraries. *Management*. Retrieved October 06, 2018 from https://doi.org/10.4271/2013-01-0515.

[15] Plaza, B. (2011). Google Analytics for measuring website performance. *Tourism Management*, *32*(3), 477–481. Retrieved October 05, 2018 from https://doi.org/10.1016/j.tourman.2010.03.015.

[16] Sofos, A., & Kostas, A. (2009). Pedagogically-Oriented Evaluation Criteria for Educational Web Resources. *World Wide Web Internet And Web Information Systems*, (December), 1–9. Retrieved October 09, 2018 from https://www.slideshare.net/elearningpapers/pedagogically-oriented.

[17] Tidal, J. (2011). Using web metric software to drive mobile website development. *Computers in Libraries*, *31*(3), 19–23. Retrieved October 12, 2018 from https://juniortidal.com/publications/

[18] Turner, S. J. (2010). Website statistics 2.0: Using Google analytics to measure library website effectiveness. *Technical Services Quarterly*, *27*(3), 261–278. Retrieved October 15, 2018 from https://doi.org/10.1080/07317131003765910.

[19] Xing, W., Guo, R., Fitzgerald, G., & Xu, C. (2015). Google analytics based temporal-geospatial analysis for web management: A case study of a K-12 online resource website. *International Journal of Information Science and Management*, *13*(1), 87–106. Retrieved October 12, 2018 from https://ijism.ricest.ac.ir/index.php/ijism/article/download/505/241.