

# The Development Tendencies of Web Analytics Tools

N. S. Redkina

*State Public Scientific and Technological Library, Siberian Branch,  
Russian Academy of Sciences, Novosibirsk, 630200 Russia*

*e-mail: to@spsl.nsc.ru*

Received January 31, 2017

**Abstract**—The capabilities of website counters and methods of analyzing the tendencies of the development of various web-analytics tools are studied using Google Trends; promising approaches to more detailed research of the trends in the area of web analytics, a multi-aspect study of data from counters, services, and log analyzers aimed at optimization and promotion of websites of various organizations are presented. An overview is given for the sources of a comparative analysis of web-analytics tools, ratings, and data from the Ruward:Track laboratory on the number of added counters on the Russian Internet.

*Keywords:* web analytics, website counters, web analyzers, assessment, web metrics, Google Trends

**DOI:** 10.3103/S0005105517030050

## INTRODUCTION

The rapid expansion of web technologies and the information sphere resonates with virtually all aspects of human activities. The number of sites and users of the web is increasing. According to the All-Russian Center for Public Opinion Research, in 2016 70% of Russian citizens aged 18 or older used the Internet in Russia [1]. The number of sites on the web reached 1.8 billion by the beginning of 2017 [2]. Websites represent an integral part of activity for any organization. Their owners have been paying increasing attention not only to the content and improvement of user interface, search engines, navigation functions, but also to gathering data about their websites using web-analytics tools in order to determine their information value, friendliness, and popularity with users, and revealing the most demanded sections and reasons of rejection.

Systems of web analytics and various counters allow processing a large amount of data on users, such as the browser type, connection speed, and display size, as well as the type and age of visitors. The gathered information becomes useful in solving the problems of website optimization and making a dynamic response to constantly changing conditions of the external environment, information preferences, and demands of users. Web analytics tools allow one to adapt services to user demands, provide a good reputation and image of the organization in the virtual world, and promote its resources and services on the web, adding new channels of information distribution and improving the support service for users.

The main goal of this work was to study the approaches for assessing the potential capabilities and functional features of web-analytics systems.

## THE DEVELOPMENT OF WEB-ANALYTICS SYSTEMS

The systems of website statistics have been developing for several decades from the tools that allowed counting only the number of users and shown pages to complex analytics systems. It is believed that the onset of web analytics occurred in the 1990s [3, p. 26–27]. In that period Internet users started showing interest in web statistics. Server logs registered website requests and some additional information, including the file-name, time, referrer (a website or page from which the request was made), IP address, browser ID, and operating system. Commercial web analytics developed later as standard analyzers of server log files improved; the function of the presenting data as tables and graphs was added. By the 2000s, with the exponential growth of Internet popularity, web analytics became a solid area in which increasingly complex solutions are developed in order to assess the effectiveness and quality of web resources. The growth in popularity of web-analytics tools is linked to their availability, functionality, simplicity, and effectiveness. During recent years, the number of web-analytics tools has been growing, their capabilities have been expanding, and the popularity of certain counters has been increasing.

According to the data from the Ruward:Track research laboratory, which conducts analytical studies based on their native technology for quantitative measurements in various Russian Internet segments and publishes a quarterly independent rating/results of research on website counters and systems of web analytics on the Russian Internet, it can be concluded that such systems as Yandex.Metrika, LiveInternet, and Google Analytics are popular in the .RU

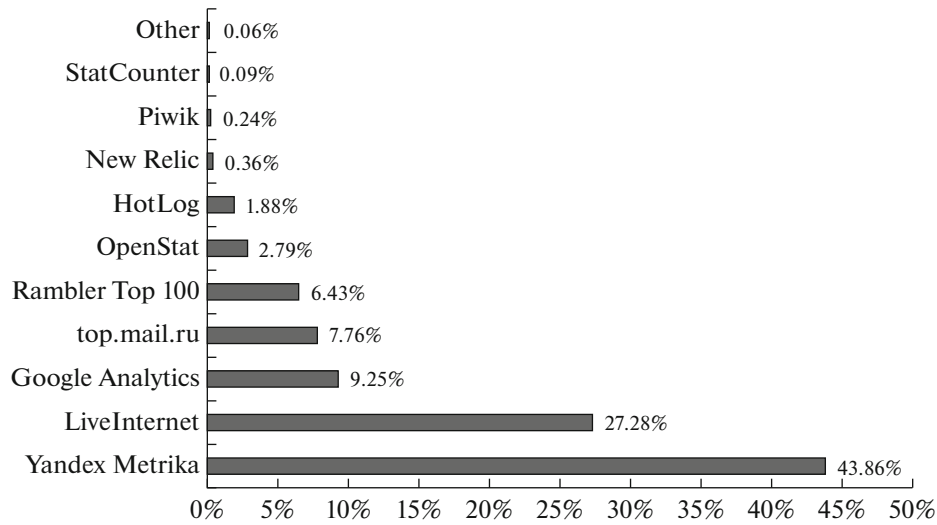


Fig. 1. Rating of web-analytics tools on the Russian Internet in 2016.

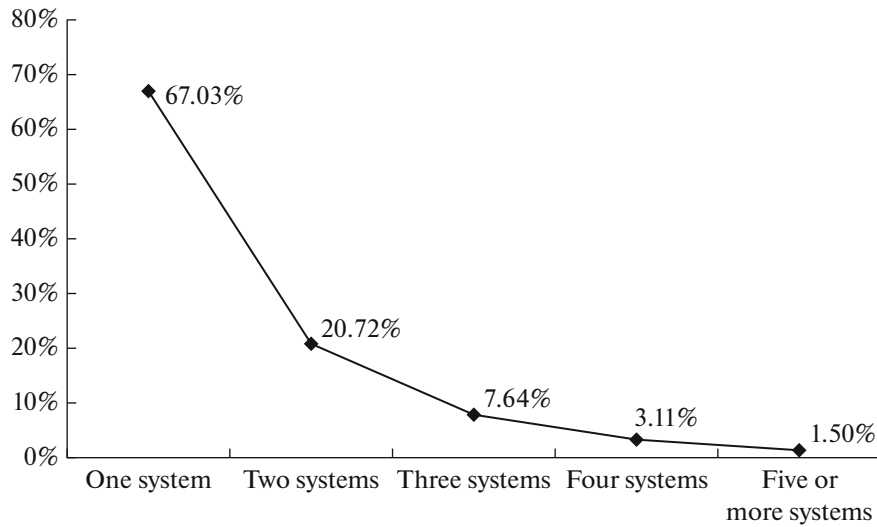


Fig. 2. The number of added web-analytics tools on the Russian Internet in 2016.

zone (Fig. 1). These were developed on the basis of a unique methodology, which allows fixing the location of a service by its “footprints.” Approximately 5 million websites are analyzed in the .RU zone. In June 2016, 5236 .RU domains were queried; 65.4% of them responded [4]. Web-analytics tools and counters were found on 42.7% of the responding domains.

In order to obtain more credible information, organizations use internal and external systems of website traffic statistics, or two or more options of free or commercial external tools. The data from the Ruward:Track laboratory confirm this fact (Fig. 2). More than 32% of the analyzed websites have two or more counters. However, the fact should be taken into account that using several counters on a website has disadvantages;

they slow down access of users to the website. As well, software tools that have different algorithms of collection of data and the results may differ substantially. As an example, in Google Analytics, if a user views only one page of the site during the entire time of a visit it is considered as a rejection; in the Yandex.Metrika system, rejection occurs when a user views only one page in less than 15 seconds.

#### THE FUNCTIONAL CAPABILITIES AND DISTINCTIVE FEATURES OF WEB-ANALYTICS TOOLS

Giving an explicit classification of web-analytics tools seems complicated considering the fact that counters are used in rating estimation and in analytical

systems. As well, for better understanding and further analysis of the modern structure of web-analytics tools, it is reasonable to note the basic methods of determining the website traffic used in web analytics: plugins (e.g., statistics plugin Word Press Stats), log analyzers (Analog, Webalizer, AWStats, etc.), counters (OpenStat (formerly Spylog), 24log.ru), ratings (LiveInternet, Rambler.Top100, top.mail.ru, Ruward, Directrix), and services (Google Analytics, Yandex.Metrika, Hotlog).

In the works of Russian and foreign researchers, broad capabilities of certain web statistics tools and metrics sites are described [3, 5–9]. Publications in which a comparative study of functional capabilities, advantages, and disadvantages, and indicators from various counters is given are of particular interest. The most popular systems, Yandex.Metrika and Google Analytics, have been studied in works by K.R. Enikeeva, I.A. Zadorozhnyi, E.S. Medvedeva, M.A. Olevinskii, and A.V. Radzevich [10–13]. The author of [14] presented a comparison of four systems of collection and analysis of website traffic (Yandex.Metrika, Google Analytics, Hotlog, and LiveInternet). A more detailed analysis was given in the work by E.Yu. Vas'kovskii and Yu.M. Brumshtein [15], which analyzed the functionality of free and paid counters used in Russia: SimilarWeb, Alexa, GoogleAnalytics, OpenStat (formerly SpyLog), LiveInternet, Yandex.Metrika, Hotlog, Clicky, Woorpa, GoSquared, Gauges, Chartbeat, StatCounter. The authors compared the above-mentioned counters according to the following subgroups: summary statistics; analysis of the content and navigation; web-design parameters; analysis of geography and behavior of users; analysis of commercial indicators; and other additional functions (the list of pages referring to the resource the ability to determine search robots, etc.).

Advantages and disadvantages were determined for each of the studied web-analytics tools. As an example, in Google Analytics it is impossible to track traffic if a user disabled the saving of cookie files; it is impossible to process data repeatedly if a user profile with configured filters was lost; there are also limitations on the number of reports and targets. The increase in the number of Yandex.Metrika installations is linked to a unique analytical tool, WebVisor, which allows one to view a video containing the actions of each user on the site pages. The disadvantages of Yandex.Metrika are insufficiently accurate traffic assessment and the lack of the possibility to configure one counter for a group of sites on subdomains and save columns in the reports. The LiveInternet system, which is popular in Russia, allows one to combine all sites in topic groups and make ratings by traffic; however, the counter does not register users with Java script support disabled.

It can be concluded that there is no single tool as yet that can solve all web-analytics problems in a

holistic manner and fully satisfy the needs of site owners. When choosing web-analytics tools, it is important to study the features of various counters and services in detail, as well as the metrics used by the systems and the reports which are formed by them, and to answer the following questions: Which web-analytics tools are installed most frequently? How many websites are there with particular counters? Which combinations of counters are used most frequently? Which counters are growing in popularity?

#### THE POPULARITY OF WEB-ANALYTICS TOOLS ACCORDING TO GOOGLE TRENDS

Google Trends is a public web application of Google Corporation, which shows how popular a certain word or phrase is in comparison with all queries during a period of time in different parts of the world and in different languages (<https://www.google.ru/trends/>). It is to be noted that Trends take consideration of only the queries entered by users in the search box at a sufficient frequency; the same queries from the same user in a short period of time and queries with apostrophes and similar characters are ignored.

When searching in Google Trends, one can view the traffic almost in real time, which reflects the popularity of a query and obtain information about indicators in a certain moment of time. The numbers on the graph reflect the popularity of a word or phrase in comparison with the total number of Google search queries in a certain time interval. If a graph is descending it may indicate a decline of interest in the topic, but not in all cases; it is possible that the popularity of other queries grew faster and fixed terms that are the subject of analysis dissolve in the avalanche of a stream of queries. As well, in computations Google Trends allows prediction of the popularity of a query in the nearest future, i.e., it predicts the interest in a particular topic. The data obtained in the analysis were depersonalized and normalized (divided by a scale variable) and the values are given from 0 to 100. Google algorithms define the point in a graph for the selected period where the query was the most popular and assume it equals 100. All other points in the graph are defined as a percentage of the maximum. If the data are insufficient, the value equals zero. Google Trends allows one to compare the statistics for search queries, regions, and periods in different languages and virtually in real time.

The content analysis of publications on The most popular traffic counters for websites topic, which was carried out on the basis of primary and secondary sources of information, revealed the most used web-analytics tools on the Russian Internet. The list was further analyzed in the Google Trends system and the queries were revealed for which a correct result is given considering the geographic filter; the list of terms for more detailed analysis was also defined.

**Table 1.** The popularity of basic terms by countries and cities

Search queries					
LiveInternet		Awstats		Google Analytics	
<i>Top countries (score)</i>	<i>Top cities (score)</i>	<i>Top countries (score)</i>	<i>Top cities (score)</i>	<i>Top countries (score)</i>	<i>Top cities (score)</i>
Latvia (100)	Moscow (100)	Switzerland (100)	Beijing (100)	Estonia (100)	San-Francisco (100)
Russia (88)	Donetsk (77)	Netherlands (99)	Amsterdam (71)	Netherlands (87)	London (92)
Ukraine (81)	Nizhny Novgorod (72)	China (96)	Shanghai (67)	Great Britain (83)	Barselona (91)
Belarus (68)	Zaporizhia (71)	Taiwan (81)	Frankfurt-on-Main (45)	Ireland (79)	Sidney (86)
Kazakhstan (52)	Saint Petersburg (70)	Germany (80)	Brussels (42)	Czech Republic (76)	Dublin (85)

At the same time, it is possible to compare up to five sets of words and phrases, each of which contains a maximum of 25 search queries. In order to form a search statement and carry out further comparative analysis, we experimentally determined five basic terms in English: LiveInternet, Google Analytics, Awstats, Spylog, and MetrikaYandex. Other variants of phrases related to web-analytics tools, for example, in Russian, such as Яндекс.Метрика or Гугл Аналитикс, etc. did not allow correct results in the Google Trends analytical system. In the first search session restriction by region was not set and the analysis was performed for the entire scope of queries (from 2004 until November 2016) in all categories.

The greatest interest in Google Analytics occurred in April 2013. This result was the highest in the entire period of study. Further, after increased interest, a tendency to a stable interest of users in Google Analytics is observed. During the 2004–2009 period, a moderate interest of users was observed in the Awstats system (the maximum was reached in November 2005 and was equal to 18). The number of LiveInternet, Spylog, and YandexMetrika search queries in the entire stream of queries does not correctly reveal these trends.

In the second search session, the restriction by region was set by regions of Russia, and the analysis was also performed for the entire scope of queries (from 2004 until November 2016) in all categories. The average value of the dynamics of the popularity during the studied period was LiveInternet (33), GoogleAnalytics (22), Awstats (9), Spylog (5), and YandexMetrika (0). Changing the YandexMetrika to ЯндексМетрика, Яндекс.Метрика, Metrika.yandex, etc. did not alter the result. Increased interest in Awstats was shown by users during the 2004–2006 period; the number of queries for LiveInternet then grew until 2010; since then, Google Analytics has been leading.

A more detailed assessment of the popularity of queries according to the geography of the users is given in Table 1. We could not determine the top regions and

cities via Spylog and YandexMetrika, because the data is too sparse to receive statistical forms.

The data show that the most interest in LiveInternet is observed among users from Latvia, Russia, Belarus, and Kazakhstan (in order of a decreasing index of queries and/or normalized scores). The Google Analytics service is popular among the citizens of San-Francisco, London, Barcelona, Sidney, and Dublin (the top cities). However, it is to be noted that these data are also relative for different regions. If the indicators related to a search query are equal in two regions it does not mean they are equally popular. As an example, if the numbers of Awstats queries in Germany and Taiwan are similar, this does not mean that users from these countries entered this word in the search box an equal number of times, but indicates its popularity across the total number of search queries. If there is a zero result for a region, this does not mean that the selected query was never used in it; it was just not sufficiently popular.

#### SIMILAR AND SUPERPOPULAR QUERIES

Google Trends also allows one to analyze similar queries and queries that are gaining popularity. These queries include words and phrases which have been searched for by users recently along with the specified word. Each query shows the percentage by which it is more popular than the previous query. If the interest in a word has grown by more than 5000%, the Superpopularity value displays instead of the percentage. An example of a studied query for Google Analytics is shown in Table 2. For the Google Analytics phrase, the words analytics, googleanalytics, and the adwords tool for creating advertisement messages are similar queries.<sup>1</sup> The maximum number of Superpopularity

<sup>1</sup> A 100-point grading scale is used, where 100 is the most popular topics; 50 is given for topics that were searched for two times less frequently than the most popular ones; 0 is given for the topics that were searched for 99% less frequently than the most popular ones

**Table 2.** A list of similar and trending queries from GoogleAnalytics terms

Top queries	Score	Trend of queries	Percentage
Analytics	100	googletagmanager	Superpopularity
Googleanalytics	65	Searchconsole	Superpopularity
Adwords	5	Googlesearchconsole	Superpopularity
Webmaster	0	Keywordplanner	+4300%
google a	0	analytics.js	+3300%

values with the query trends googletagmanager (a Google tag manager), searchconsole, and google-searchconsole (this a multifunctional tool that enables one to resolve technical issues and work on website optimization, track visibility, and present the website in Google search results). This information allows prediction of the interest in a term.

### CONCLUSIONS

The modern web-analytics systems are actively developing and popular tools with a rich functionality for analyzing websites. Using various sources of information regarding the trends of development and capabilities of various counters, services, and log analyzers, one can choose the most appropriate tool to study the website of a particular organization.

### REFERENCES

1. New about digital literacy, or Russians getting used to the web, *Press-Vyp. VTsIOM*, 2016, no. 3084. <http://wciom.ru/index.php?id=236&uid=115657>. Accessed September 13, 2016.
2. Web Server Survey. <http://news.netcraft.com/>. Accessed January 25, 2017.
3. Kaushik, A., *Web Analytics: An Hour a Day*, Sybex, 2007.
4. Systems of Web Analytics. June 2016. <http://track.ruward.ru/analytics>. Accessed November 21, 2016.
5. Yakovlev, A. and Dovzhikov, A., *Veb-analitika: Osnovy, sekrety, tryuki* (Web Analytics: Basics, Secrets, and Tricks), St. Petersburg: BKhV-Peterburg, 2010.
6. Shlyakhtina, S., Review of solutions for the analysis of site traffic. <http://compress.ru/Article.aspx?id=18178>. Accessed November 21, 2016.
7. Egorova, I.N. and Kadushkevich, O.N., Methods for efficient use of Google Analytics tools, *ScienceRise*, 2016, nos. 1–2, (18), pp. 40–43.
8. Andreeva, S.L. and Engibarova, M.M., Web analytics as an intellectual tool, *BI-tehnologii v optimizatsii biznes-protsessov: Materialy Mezhdunarodnoi nauchno-prakticheskoi ochno-zaochnoi konferentsii* (BI-Technologies in the Optimization of Business Processes: Proc. Int. Sci.-Pract. Conf.), 2014, pp. 60–62.
9. Skorodumov, P.V. and Kholodev, A.Yu., Analysis of the popularity of the website of a scientific organization with the help of various systems for collecting statistical data, *Vopr. Territ. Razvit.*, 2016, no. 1 (31), pp. 11–13.
10. Enikeev, K.R., Implementation of the system of analysis and monitoring of the web site, *8 Vserossiiskaya zimnyaya shkola-seminar aspirantov i molodykh uchenykh Aktual'nye problemy nauki i tekhniki, Ufa, 19–20 fevr., 2013* (8 All-Russian Winter School-Seminar of Post-Graduates and Young Scientists Actual Problems of Science and Technology), Ufa, 2013, vol. 1, pp. 127–130.
11. Zadorozhnyi, I.A. and Medvedeva, E.S., Comparative analysis of Yandex metrics and Google Analytics as tools for evaluating the effectiveness of digital marketing communications, *Nauch. Issled.: Teor. Prakt.*, 2015, vol. 2, no. 4 (5), pp. 161–164.
12. Olevinskii, M.A., Web analytics. Comparison of web analytics systems, *InSitu*, 2015, no. 4, pp. 46–48.
13. Radzevich, A.V., Web analytics for business: How to draw the right conclusions about the effectiveness of the site, *Internet-Mark.*, 2012, no. 4, pp. 218–225.
14. Skorodumov, P.V. and Kholodev, A.Yu., Analysis of approaches and tools for analyzing statistics of visits to a scientific organization's website, *Vopr. Territ. Razvit.*, 2015, no. 9 (29), pp. 11–12.
15. Vas'kovskii, E.Yu. and Brumshtein, Yu.M., System analysis of the functionality of the counters of site traffic, *Pri-kasp. Zh.: Upr. Vys. Tekhnol.*, 2015, no. 3, pp. 96–113.

Translated by A. Dunaeva