

# Информационные технологии

© Скородумов П.В., Холодев А.Ю.

## АНАЛИЗ ПОДХОДОВ И ИНСТРУМЕНТАЛЬНЫХ СРЕДСТВ АНАЛИЗА СТАТИСТИКИ ПОСЕЩЕНИЯ ВЕБ-САЙТА НАУЧНОЙ ОРГАНИЗАЦИИ



### СКОРОДУМОВ ПАВЕЛ ВАЛЕРЬЕВИЧ

кандидат технических наук, заведующий лабораторией интеллектуальных и программно-информационных систем отдела проблем научно-технологического развития и экономики знаний. Федеральное государственное бюджетное учреждение науки Институт социально-экономического развития территорий Российской академии наук  
E-mail: spv.vsc@gmail.com



### ХОЛОДЕВ АЛЕКСАНДР ЮРЬЕВИЧ

инженер-исследователь отдела проблем научно-технологического развития и экономики знаний  
Федеральное государственное бюджетное учреждение науки  
Институт социально-экономического развития территорий Российской академии наук  
E-mail: der.goodwin@ya.ru

*Веб-сайт научной организации является важнейшим элементом процесса популяризации деятельности научной организации в сети Интернет. Информационные технологии являются сегодня одним из самых динамично развивающихся направлений науки, технологий и техники. При этом использование самых современных и технологичных систем по организации веб-сайтов не гарантирует их популярности, необходима системная работа по их продвижению, поддержке и наполнению. Требуется постоянный контроль и оценка результатов выполняемой работы с целью принятия своевременных управленческих решений, для корректировки направлений дальнейшего развития веб-сайта. Одной из задач, направленных на достижение указанной цели, является оценка популярности веб-сайтов научной организации с помощью анализа статистики их посещения. В статье приведен обзор существующих методов оценки популярности веб-сайтов. Детально рассмотрены автоматизированные подходы. Приведено сравнение наиболее популярных систем сбора и анализа статистики посещения сайтов (Яндекс.Метрика, Google Analytics, Liveinternet и HotLog). В статье отражены широкие функциональные возможности рассмотренных систем. Обосновано решение Института социально-экономического развития территорий РАН (ИСЭРТ РАН) по разработке собственной системы сбора и анализа статистики посещения веб-сайтов. В работе приведено описание основных частей системы сбора и анализа статистики посещения веб-сайтов ИСЭРТ РАН. Представлены статистические данные по результатам июня и с начала 2015 года, полученные с помо-*

щью собственной системы сбора и анализа статистики посещения веб-сайтов, а также систем Яндекс.Метрика и Google Analytics. В статье выполнено сравнение числа просмотров веб-сайтов Интернет-портала ИСЭРТ РАН, полученное с помощью разных систем, отмечены возможные причины возникновения расхождения в результатах. В заключении отмечается, что анализ статистики посещения веб-сайтов научной организации с помощью различных систем сбора и анализа данных имеет особую значимость для современной научной организации, в том числе и с точки зрения положения учреждения в веб-метрических рейтингах, а также целесообразность продолжения работ по данному направлению.

*Веб-сайт научной организации, статистика посещения, системы сбора и анализа статистики посещения веб-сайтов.*

Популяризация научных достижений является одной из важнейших задач научной организации, что подтверждается проектом системы критериев оценки деятельности научных организаций Министерства образования и науки РФ, отмечающим необходимость учета количества упоминаний об организации в СМИ и посещаемости<sup>1</sup> официальных сайтов в сети Интернет [5].

Современное развитие информационных технологий и сети Интернет, в частности, существенным образом повлияли на деятельность научных организаций. Отдельные исследователи и организации в целом получили огромные возможности по продвижению результатов своих трудов и взаимодействию между собой посредством средств электронной коммуникации.

Важнейшим элементом этого процесса является веб-сайт<sup>2</sup> научной организации. Он служит не только площадкой для размещения справочной и новостной информации о работе учреждения, но и способствует формированию единого информационного пространства, объединяющего многочисленные ресурсы, от-

ражающие широкий спектр направлений деятельности учреждения [4].

Информационные технологии являются одним из самых динамично развивающихся направлений науки, технологий и техники и, кроме того, входят в перечень из 27 критических технологий Российской Федерации [6], что способствует совершенствованию и постоянному появлению новых решений, направленных на создание передовых веб-сайтов. Однако использование самых современных и технологичных систем по организации веб-сайтов не гарантирует их популярности, необходима системная работа по их продвижению, поддержке и наполнению. Требуется постоянный контроль и оценка результатов выполняемой работы с целью принятия своевременных управленческих решений, для корректировки направлений дальнейшего развития. Одной из задач, направленных на достижение указанной цели, является оценка популярности веб-сайтов научной организации с помощью анализа статистики их посещения.

Целью данной статьи является обзор и сравнение существующих подходов и инструментальных средств сбора и анализа статистических данных о посещаемости веб-сайтов, а также анализ их применимости к оценке популярности веб-сайтов научной организации.

Для оценки посещаемости веб-сайта широко применяется метод веб-

<sup>1</sup> Посещение – процесс взаимодействия посетителя с сайтом в течение некоторого времени.

<sup>2</sup> Веб-сайт – системное многоуровневое объединение разных ресурсов и сервисов. Отдельные крупные фрагменты сайта, а иногда и сами сайты называют порталом [11].

аналитики<sup>3</sup>. Критериями оценки могут быть: число уникальных пользователей<sup>4</sup>, количество просмотров<sup>5</sup>, число зарегистрированных пользователей, источники входящего трафика<sup>6</sup> и др.

В общем случае методы оценки популярности веб-сайтов можно разделить на ручные и автоматизированные [5].

В первом случае применяются такие подходы, как телефонные интервью, анкетирование и ручной анализ данных веб-сервера. Методы данного направления требуют значительных человеческих и временных затрат, в связи с чем по эффективности они уступают автоматизированным.

Вторые делятся на способы с участием или без участия пользователей. При участии пользователей проводятся вебинары и интернет-опросы. Недостатком такого подхода также являются значительные трудовые затраты, связанные с привлечением к работе пользователей, что кроме всего прочего не всегда вызывает положительный отклик.

Наибольший интерес вызывают автоматизированные методы оценки популярности веб-сайтов без участия пользователей, позволяющие экономить не только временные, но и человеческие ресурсы, т.к. полностью исключают участие человека в процессе сбора статистических данных, что занимает значительно большую часть времени по сравнению с самим анализом.

В этом случае способы сбора информации можно разделить на два типа:

<sup>3</sup> Веб-аналитика – метод изучения опыта пользователей в глобальном информационном пространстве для его улучшения [4].

<sup>4</sup> Уникальный пользователь – неповторяющийся пользователь, посетивший сайт за отчетный период (хотя бы один раз).

<sup>5</sup> Просмотр страницы – однократное отображение содержимого страницы сайта на экране компьютера.

<sup>6</sup> Сетевой трафик – объем информации, передаваемой по сети за единицу времени.

1. Активные, когда за сбор информации отвечают, например, поисковые роботы.

2. Пассивные, когда сбор осуществляется, например, на стороне сервера, собственными силами организации владельца сайта.

К активному способу сбора статистической информации относятся аналитические системы, реализованные самими поисковыми системами, например, Яндекс.Метрика и Google Analytics.

Яндекс.Метрика представляет собой бесплатный сервис, ориентированный на анализ посещаемости веб-сайтов и поведения пользователей на их страницах. Для подключения требуется размещение соответствующего программного кода на страницах ресурса [7].

Система позволяет оценить такие основные критерии, как количество посетителей, число просмотров, время пребывания пользователей на страницах сайта и др. С помощью сервиса можно сформировать географическое распределение пользователей, определять используемые ими в работе операционные системы, браузеры и другие технические параметры, позволяющие оценить возможности пользователей по просмотру того или иного сложного контента.

В ряде случаев (когда пользователь использует в работе поисковик или сервисы компании Яндекс) система позволяет определить и демографические характеристики посетителей, сформировать их «портреты» (возможно определение таких параметров, как пол, возраст, интересы, предпочтения и др.).

Особенностью системы Яндекс.Метрика является возможность анализа поведения пользователей на страницах веб-сайта. Сервис позволяет определять элементы сайта (ссылки, кнопки, формы и т. д.), пользующиеся наибольшей популярностью среди его посетителей.

Система не только формирует карты входящего трафика, но и дальнейшего поведения пользователей на нем (перемещение между страницами, перемещение мыши, нажатие клавиш). Также возможна оценка количества переходов по ссылкам, размещенным на странице сайта и ключевым фразам, по которым осуществлялись переходы со страниц поисковых систем, сортировка наиболее часто встречающихся ключевых слов.

Анализ поведения пользователей на страницах веб-сайта является важным элементом системы оценки статистики посещения, так как позволяет в короткие сроки определить предпочтения пользователей и своевременно повлиять на содержимое страниц сайта с целью повышения его популярности.

Система позволяет настроить контроль за работоспособностью (доступностью) ресурса. В случае возникновения неисправностей формируется специальное сообщение о сбое, направляемое администратору с рекомендациями о принятии мер по его устранению.

В целом система располагает богатым набором настроек и возможностей по оценке посещаемости страниц веб-сайтов с самых разных точек зрения.

Система Google Analytics также относится к активному способу сбора статистической информации. Для подключения, как и в случае с Яндекс.Метрикой, необходимо размещение соответствующего программного кода на страницах ресурса [8].

Система позволяет анализировать такие параметры, как число посещений, количество уникальных пользователей, число просмотров страниц, среднюю продолжительность посещения и др.

Как и рассмотренная выше Яндекс.Метрика, система Google Analytics позволяет анализировать источники входящего трафика, проводить оценку его значимо-

сти. С ее помощью можно формировать картину поведения пользователей на страницах ресурса.

К пассивным способам сбора статистической информации относятся счетчики и лог-анализаторы. Среди наиболее известных на сегодняшний день счетчиков можно отметить Liveinternet и HotLog.

Liveinternet является одним из наиболее удобных и популярных сервисов (в том числе среди российских пользователей), предназначенных для сбора статистической информации о посещаемости веб-сайтов в сети Интернет [10].

Для подключения счетчика пользователю необходимо пройти процедуру регистрации с указанием тематики сайта [11]. Последнее позволяет объединять сайты в группы по направлениям с последующим формированием рейтинга, отражающего реальное положение веб-ресурса среди своих основных конкурентов.

Сервис позволяет оценить статистику посещения за определенный период (день, неделю, месяц), выделить ключевые слова и определить количество переходов по ним из поисковых систем и социальных сетей. Система обладает графическим интерфейсом, необходимым для представления результатов в наглядном виде.

Счетчик способен собирать данные по таким критериям, как количество посетителей, размер аудитории, количество посещений одного пользователя, число просмотренных им страниц и т. д. [10]. Кроме того, с помощью Liveinternet можно сравнить результаты статистики посещения веб-сайта с другими ресурсами, рассчитать процентное соотношение и сравнить размер аудитории.

Среди минусов счетчика Liveinternet можно отметить открытость статистической информации всем зарегистрированным пользователям системы, приблизительный подсчет общего коли-

чества пользователей (это связано с особенностями программной реализации), возможность выбора временного промежутка только из имеющегося списка интервалов, ограниченное время хранения информации.

Сервис сбора статистической информации о посещаемости веб-сайтов HotLog также является бесплатным и для подключения требует размещения программного кода счетчика на страницах веб-ресурса. Как и рассмотренный выше счетчик Liveinternet, HotLog формирует рейтинг на основе собираемой статистической информации [9].

Следует отметить, что администратору сайта сервис HotLog предоставляет не только подробную статистику посещений, но и множество инструментов анализа эффективности проводимых мероприятий (анализ поведения пользователей на страницах ресурса, оценка результатов поисковой оптимизации и мероприятий по привлечению пользователей).

Сервис обладает расширенными возможностями веб-аналитики, такими, как оценка трафика, показатели отказов, глубина просмотра, расчет конверсии<sup>7</sup> и др.

Среди преимуществ сервиса необходимо отметить значительное число готовых отчетов по анализу большого количества показателей посещаемости веб-сайта: число просмотров, популярность отдельных страниц, география пользователей, анализ источников и поисковых запросов, анализ ссылочной массы и др.

HotLog позволяет фильтровать данные в отчетах, экспортировать результаты в популярный формат csv, в отличие от сервиса Liveinternet возможна настройка гостевого доступа. Кроме того, сервис позволяет настроить оповещения о резком изменении статистики посещения, фор-

мировании ежемесячных печатных отчетов и отправку результатов на e-mail.

Рассмотренные выше системы предназначены для анализа посещаемости всех веб-сайтов без разграничения по типам организации владельца. В первую очередь значительное число формируемых отчетов и анализируемых критериев говорят о направленности их на коммерческий сегмент сети Интернет. Причиной этому является стремительный рост Интернет-торговли и перевод многих операций в область электронной коммерции, что повлекло за собой необходимость привлечения пользователей на страницы Интернет-ресурсов организаций. А следовательно, особую значимость получила и оценка посещаемости сайтов (в том числе поведение пользователей на них) с целью повышения конверсии.

Стоит отметить, что использование описанных выше систем возможно и для некоммерческих организаций, в частности для научных учреждений, с целью анализа основных критериев посещаемости веб-ресурсов.

Выбор конкретного счетчика в каждом случае индивидуален и может определяться различными критериями, в том числе и теми параметрами, которые планируется учитывать при анализе статистики посещения веб-сайта (табл. 1).

Как видно из таблицы 1, в большинстве своем системы обладают широкими функциональными возможностями, особенно выделяются Яндекс.Метрика и Google Analytics. Ограничением является формирование набора статистических данных по уникальным характеристикам, таким как популярность статей научных журналов, выделение наиболее активных пользователей на страницах веб-сайтов, статистика анкетных данных и др. В тех случаях, когда планируется глубокий анализ статистики посещения пользователей веб-сайтов целесообраз-

<sup>7</sup> Коэффициент конверсии – отношение числа посетителей, выполнивших целевое действие на сайте к общему числу посетителей сайта.

Таблица 1. Сравнение систем сбора и анализа статистики посещения вебсайтов

Критерии сравнения	Google Analytics	Яндекс.Метрика	LiveInternet	HotLog
1. Составление демографических портретов пользователей	+	+	+	+
2. Стоимость использования	Бесплатно. Есть платная версия Google Analytics Premium для обработки значительного числа просмотров.	Бесплатно. Есть возможность подключения платных услуг	Бесплатно	Бесплатно. Есть платный пакет с дополнительными отчетами и серверными мощностями
3. Настройка собственных отчетов	Более 80 видов настраиваемых отчетов	Имеется конструктор отчетов	-	-
4. Оценка трафика	+	+	-	+
5. Настройка оповещений на электронную почту	+	+	+	+
6. Экспорт данных	PDF, CSV, Excel, TSV	PDF, CSV, XLSX	-	Excel, CSV, PDF
7. Анализ поведения пользователей на сайте	+	+	-	-
8. Графический интерфейс	+	+	+	+

нее использовать приложение, самостоятельно разработанное специалистами организации.

В этом случае организация-собственник такой системы получает неограниченные возможности по настройке и расширению всевозможных отчетов, независимо от поставщика программного продукта.

Именно невозможность оценки ряда показателей с помощью сторонних готовых систем послужила причиной создания сотрудниками лаборатории интеллектуальных и программно-информационных систем Института социально-экономического развития территорий РАН собственной системы анализа статистики посещения Интернет-портала ИСЭРТ РАН.

Необходимо отметить, что первая версия веб-сайта Института социально-экономического развития территорий РАН была запущена в 2000 году. Со временем количество сайтов Института росло, появились сайты научных журналов «Экономические и социальные перемены: факты, тенденции, прогноз», «Проблемы развития территории», «Вопросы территориального развития», сайты библиотеки, типографии, предприятия «Территория инноваций». В 2014 году было

получено свидетельство о регистрации программы для ЭВМ на систему управления Интернет-порталом [13].

За время существования Интернет-портала ИСЭРТ РАН техническая основа разработки менялась несколько раз, осуществлялся переход с одной платформы на другую. В конечном итоге по причине сложности и невозможности расширения системы с помощью готовых компонентов, было принято решение о разработке собственной CMS<sup>8</sup>.

Сегодняшнее развитие информационных технологий привело к необходимости разработки новых сервисов, без которых рост посещаемости веб-сайта просто невозможен. Современные системы проектирования предоставляют значительные преимущества для разработчиков по созданию дополнительного функционала такого рода по сравнению со старыми инструментами.

Опираясь на все вышесказанное, Институтом социально-экономического развития территорий РАН было принято решение о модернизации существующей

<sup>8</sup> CMS (Content management system) – система управления содержимым (контентом), информационная система или компьютерная программа, используемая для обеспечения и организации совместного процесса создания, редактирования и управления контентом.

технической платформы. В настоящее время сотрудниками лаборатории интеллектуальных и программно-информационных систем Института осуществляется перевод Интернет-портала ИСЭРТ РАН на современный фреймворк разработки веб-приложений Yii2 [16].

Использование в качестве основы для построения системы управления Интернет-порталом организации веб-ориентированных технологий, языка программирования PHP, фреймворка Yii2, паттернов проектирования позволит, опираясь на широкие возможности указанных решений, построить кроссплатформенную многофункциональную систему с широкими возможностями по модернизации и интеграции с другими приложениями, а также коммуникации между пользователями [14].

Реализация системы управления Интернет-порталом на основе фреймворка Yii2 будет способствовать не только разработке новых сервисов, повышающих удобство работы с веб-сайтом и влияющих на рост его популярности среди пользователей, но и развитию системы сбора и анализа статистики посещения веб-сайтов ИСЭРТ РАН.

Существующая система анализа статистики посещения Интернет-портала ИСЭРТ РАН состоит из двух частей. Первая часть отвечает за фиксацию посещения в базе данных.

Каждый посетитель сайта получает уникальный номер, называемый идентификатором сессии. Поддержка сессий заключается в способе сохранения данных между несколькими последовательными просмотрами веб-сайта [12]. С помощью механизма сессий сайт может подсчитать количество уникальных посетителей и идентифицировать пользователей, которые уже посещали сайт раньше.

Длительность сессии определяется промежутком времени между первым и

последним просмотром в сессии и влияет на показатель отказов<sup>9</sup> – процент от общего количества посещений, в рамках которых состоялось не более одного просмотра страницы.

В результате на сервере собирается статистика всех сайтов Интернет-портала Института. С помощью специального программного скрипта<sup>10</sup> присутствие пользователя на страницах Интернет-портала фиксируется в базе данных в соответствии с присвоенным ему идентификатором.

Вторая часть системы анализирует имеющиеся в базе данных записи и в соответствии с выбранными критериями и временным интервалом формирует результирующую таблицу статистических данных. В системе реализовано около 20 различных отчетов, позволяющих оценить посещения Интернет-портала с различных точек зрения. В том числе доступны отчеты о регистрации новых пользователей на каждом сайте, количестве просмотров страниц, с сортировкой посетителей по различным параметрам («зарегистрированный пользователь – гость», «сотрудник – не сотрудник», «англоязычные – русскоязычные страницы» и др.). Ведется статистика анкетных данных зарегистрированных посетителей, таких как география, пол, возраст, образование и др., отслеживается число просмотров рубрик и динамика популярных статей на сайтах научных журналов, формируется рейтинг наиболее активных пользователей с возможностью распределения их посещений между отдельными сайтами всего портала.

<sup>9</sup> Отказ – это визит, во время которого посетитель, зайдя на сайт, просмотрел только одну страницу, не совершив при этом перехода на другую страницу до времени окончания сессии.

<sup>10</sup> Скрипт (script – англ. «сценарий») – компьютерная программа, представляющая собой последовательность инструкций для работы некоторого приложения (или нескольких приложений).

Одним из наиболее распространенных показателей оценки популярности веб-сайта является количество просмотров всех его страниц. Данная характеристика находит отражение в проекте системы критериев оценки деятельности научных организаций Министерства образования и науки РФ [5], а также оценивается ведущими институтами РАН при анализе статистики посещения их веб-сайтов, например, Институтом социологии РАН [3].

Используя системы Яндекс.Метрика, Google Analytics, а также данные системы сбора статистики ИСЭРТ РАН, можно проанализировать популярность Интернет-портала Института с точки зрения просмотров страниц всех веб-сайтов организации в июне (табл. 2, рис. 1) и с начала 2015 года (табл. 3 – 5, рис. 2). Проанализировать работу счетчиков Liveinternet и HotLog невозможно, т.к. они не подключены к Интернет-порталу Института.

Если сравнивать данные, полученные с помощью собственной системы сбора и

анализа статистики ИСЭРТ РАН и сервиса Яндекс.Метрика, то можно говорить о соответствии результатов друг другу, итоговая разница составляет 4% (22010 и 21055 просмотров соответственно) от общего числа просмотров, полученного с помощью первой системы.

Если же сравнить результаты, полученные с помощью собственной системы сбора и анализа статистики ИСЭРТ РАН и сервиса Google Analytics, то разница составляет уже 25% (22010 и 16527 просмотров соответственно) от общего числа просмотров, полученного с помощью первой системы. Особенно заметно расхождение по сайтам научных журналов ПРТ и ВТР. Это может быть объяснено двумя причинами:

1. Система Google Analytics является разработкой компании Google и направлена на индексацию англоязычного сегмента сети Интернет.
2. В сервисах Яндекс.Метрика и Google Analytics используется такое понятие, как отказы (смотри выше).

Таблица 2. Количество просмотров страниц веб-сайтов ИСЭРТ РАН в июне 2015 года

Название сайта	Система сбора статистики ИСЭРТ РАН	Яндекс.Метрика	Google Analytics
1.Главный сайт	11637	11307	10967
2.Сайт журнала ЭиСП	3204	2973	2721
3.Сайт журнала ПРТ	2739	2600	245
4.Сайт библиотеки	665	616	596
5.Сайт НОЦ	2278	2144	1934
6.Сайт журнала ВТР	1487	1415	64
Итого:	22010	21055	16527

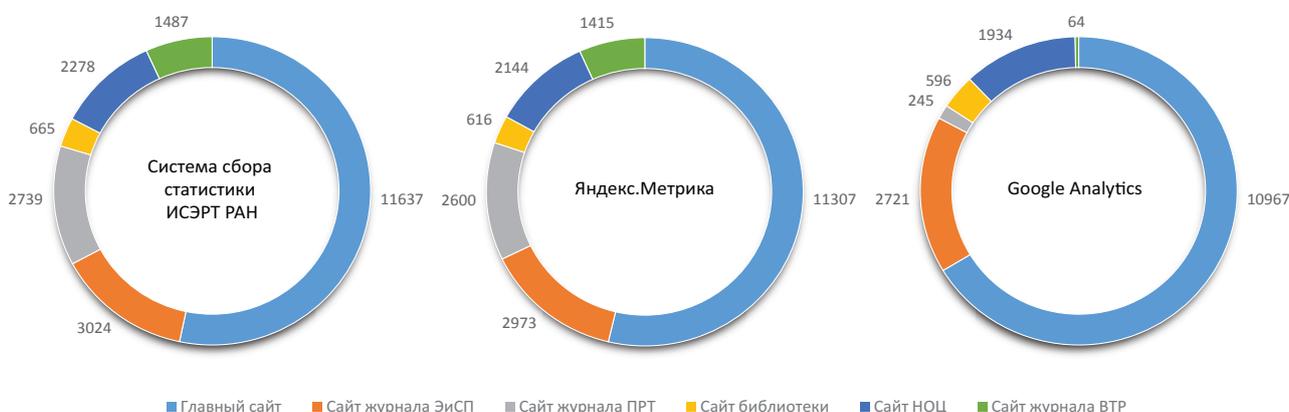


Рис. 1. Количество просмотров страниц веб-сайтов ИСЭРТ РАН в июне 2015 года

Таблица 3. Количество просмотров страниц веб-сайтов ИСЭРТ РАН в 2015 г., полученное с помощью системы сбора статистики Института

Название сайта	Январь	Февраль	Март	Апрель	Май	Июнь
1.Главный сайт	8358	9064	13634	12870	11874	11637
2.Сайт журнала ЭиСП	4470	4474	5303	4628	3572	3204
3.Сайт журнала ПРТ	2745	2727	2932	3192	2802	2739
4.Сайт библиотеки	705	911	1073	908	781	665
5.Сайт НОЦ	4561	3868	5271	3677	2915	2278
6.Сайт журнала ВТР	1595	1630	1715	1949	1463	1487
Итого:	22434	22674	29928	27224	23407	22010

Таблица 4. Количество просмотров страниц веб-сайтов ИСЭРТ РАН в 2015 г., полученное с помощью системы Яндекс.Метрика

Название сайта	Январь	Февраль	Март	Апрель	Май	Июнь
1.Главный сайт	7107	7975	13315	13519	11446	11307
2.Сайт журнала ЭиСП	3940	3860	4862	4364	3126	2973
3.Сайт журнала ПРТ	2359	2504	2767	3030	2500	2600
4.Сайт библиотеки	609	716	1030	853	684	616
5.Сайт НОЦ	4212	3546	4958	3619	2722	2144
6.Сайт журнала ВТР	1396	1387	1807	1856	1416	1415
Итого:	19623	19988	28739	27241	21894	21055

Таблица 5. Количество просмотров страниц веб-сайтов ИСЭРТ РАН в 2015 г., полученное с помощью системы Google Analytics

Название сайта	Январь	Февраль	Март	Апрель	Май	Июнь
1.Главный сайт	7519	8356	12904	12775	11369	10967
2.Сайт журнала ЭиСП	3180	3374	4056	3461	2543	2721
3.Сайт журнала ПРТ	887	313	527	444	168	245
4.Сайт библиотеки	573	759	935	766	625	596
5.Сайт НОЦ	3370	2761	3866	3035	2552	1934
6.Сайт журнала ВТР	34	27	423	57	47	64
Итого:	15563	15590	22711	20538	17304	16527

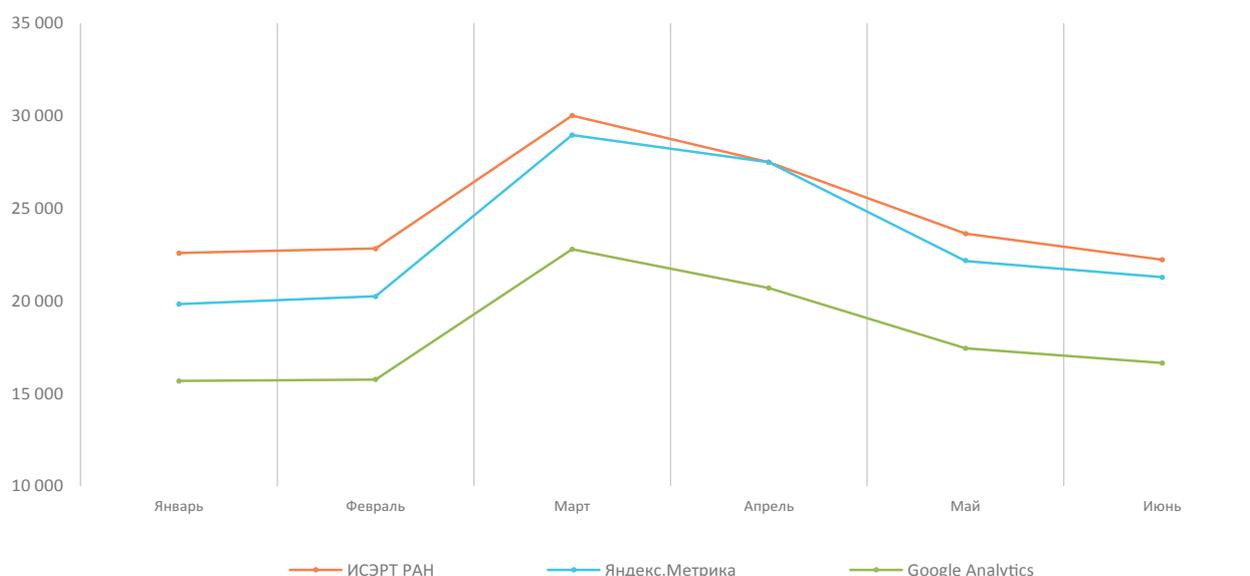


Рис. 2. Количество просмотров страниц веб-сайтов ИСЭРТ РАН с начала 2015 г.

Общепринятого стандарта минимального или максимального количества времени, которое посетитель должен провести на сайте прежде, чем выйти, чтобы состоялся отказ, не существует. Каждая из систем определяет этот критерий по-разному. Например, если пользователь просмотрел только одну страницу сайта, то счетчик Google Analytics посчитает такое посещение как отказ. А в системе Яндекс.Метрика отказом считается визит, в рамках которого состоялся просмотр лишь одной страницы, продолжавшийся менее 15 секунд.

Таким образом, следует отметить, что при сопоставлении данных различных систем сбора и анализа статистики посещения веб-сайтов, необходимо обязательно учитывать правила, по которым осуществляется подсчет тех или иных критериев. В различных случаях алгоритмы могут существенно различаться и приводить к значительной разнице в итоговых результатах.

Анализируя статистику посещения Интернет-портала ИСЭРТ РАН с начала 2015 года (см. табл. 3 – 5, рис. 2), можно говорить, что выводы, сделанные по итогам июня 2015 года, справедливы и для более длительного периода. Данные, полученные с помощью собственной системы сбора и анализа статистики ИСЭРТ РАН и сервиса Яндекс.Метрика, близки друг к другу (что говорит о соответствии используемых при подсчете алгоритмов), в то время как результаты, полученные с помощью сервиса Google Analytics ниже в среднем примерно на 25%.

Таким образом, можно отметить, что влияние указанных выше причин носит постоянный характер и не связано с колебаниями, вызванными активностью пользователей в течение года. Причина возникновения разницы показателей лежит в правилах подсчета параметров,

учитываемых при сборе статистики (в частности понятия отказа).

В заключение необходимо отметить, что на сегодняшний день практически не существует научных организаций, не представленных и не занимающихся популяризацией своих веб-ресурсов в сети Интернет. Результаты этой деятельности учитываются при формировании вебометрических рейтингов, получивших особое значение с точки зрения оценки деятельности научных организаций в современном мире [15], в частности, рейтингов Кибернетической лаборатории Центра научной информации и документации Национального исследовательского совета Испании (Cybermetrics Lab) и рейтинга ИПМИ КарНЦ РАН<sup>11</sup> [1; 2].

Все это обязывает научные организации следить за текущей статистикой посещения веб-сайтов и прилагать значительные усилия для увеличения показателей, учитываемых при анализе, применяя при этом различные системы сбора и анализа статистических данных.

Необходимо отметить еще раз, что выбор конкретной системы в каждом случае индивидуален. Каждая из систем обладает широким функционалом и способна представить различные отчеты по статистике посещения веб-сайта организации. Определяющим в этом вопросе является детализация статистических данных, насколько глубоким является анализ поведения пользователей на страницах ресурса.

Особое значение при анализе популярности веб-сайтов организации имеет сравнение результатов, получаемых с помощью различных систем. В частности, распределение пользователей по географии и демографии, а также подходы, применяемые в этом случае различными

<sup>11</sup> Институт прикладных математических исследований КарНЦ РАН.

системами (учет анкетных данных либо сохранение истории поведения пользователя в сети Интернет). Детальный анализ

статистики посещения веб-сайтов ИСЭРТ РАН планируется представить в дальнейших работах.

### ЛИТЕРАТУРА

1. Вебометрический рейтинг Кибернетической лаборатории Центра научной информации и документации Национального исследовательского совета Испании [Электронный ресурс] Cybermetrics Lab, Consejo Superior de Investigaciones Científicas. – Режим доступа : <http://www.webometrics.info>
2. Вебометрический рейтинг научных учреждений России [Электронный ресурс]. – Режим доступа : <http://webometrics-net.ru>
3. Дайджест Официального портала ИС РАН (декабрь 2014 – февраль 2015) [Электронный ресурс]. – Режим доступа : [http://www.isras.ru/files/File/publ/Digest\\_portal\\_zima\\_final.pdf](http://www.isras.ru/files/File/publ/Digest_portal_zima_final.pdf)
4. Зеленин, А. С. Организация информационного пространства научного учреждения (на примере ИСЭРТ РАН) [Текст] / А. С. Зеленин // Проблемы развития территории. – 2012. – № 1 (57). – С. 119–127.
5. Кабакова, Е. А. Веб-сайт научно-исследовательского учреждения: наполнение, посетители, развитие [Электронный ресурс] / Е. А. Кабакова, В. С. Усков // Вопросы территориального развития. – 2014. – № 3. – Режим доступа : <http://vtr.isert-ran.ru/article/1396/full>
6. Об утверждении приоритетных направлений развития науки, технологий и техники в РФ и перечня критических технологий РФ [Электронный ресурс] : Указ Президента РФ от 07.07.2011 № 899. – Режим доступа : <http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=116178>
7. Официальный сайт компании Яндекс [Электронный ресурс]. – Режим доступа : <https://metrika.yandex.ru>
8. Официальный сайт системы Google Analytics [Электронный ресурс]. – Режим доступа : [http://www.google.com/intl/ru\\_ru/analytics/index.html](http://www.google.com/intl/ru_ru/analytics/index.html)
9. Официальный сайт системы HotLog [Электронный ресурс]. – Режим доступа : <http://hotlog.ru>
10. Официальный сайт системы Liveinternet [Электронный ресурс]. – Режим доступа : <http://www.liveinternet.ru>
11. Райзберг, Б. А. Современный экономический словарь [Текст] / Б. А. Райзберг, Л. Ш. Лозовский, Е. Б. Стародубцева. – 5-е изд., перераб. и доп. – М. : ИНФРА-М, 2007. – 495 с.
12. Руководство по PHP [Электронный ресурс]. – Режим доступа : <http://php.net/manual/ru/intro.session.php>
13. Система управления интернет-порталом [Электронный ресурс] : свидетельство № 2014612043 от 17.02.2014. – Режим доступа : <http://www.vscs.ac.ru/storage/files/2014612043.pdf>
14. Скородумов, П. В. Система управления проектами организации: анализ подходов и существующих программных решений [Электронный ресурс] / П. В. Скородумов, Д. А. Баданин // Вопросы территориального развития. – 2015. – № 5. – Режим доступа : <http://vtr.isert-ran.ru/article/1505/full>
15. Третьякова, О. В. Возможности использования вебометрического анализа в оценке сайта научного института [Электронный ресурс] / О. В. Третьякова, Е. А. Кабакова // Вопросы территориального развития. – 2014. – № 2. – Режим доступа : <http://vtr.isert-ran.ru/article/1389/full>
16. Yii framework [Electronic resource]. – Available at : <http://www.yiiframework.com>

### ИНФОРМАЦИЯ ОБ АВТОРАХ

*Скородумов Павел Валерьевич* – кандидат технических наук, заведующий лабораторией интеллектуальных и программно-информационных систем отдела проблем научно-технологического развития и экономики знаний. Федеральное государственное бюджетное учреждение науки Институт социально-экономического развития территорий Российской академии наук. Россия, 160014, г. Вологда, ул. Горького, д. 56а. E-mail: [srv.vscs@gmail.com](mailto:srv.vscs@gmail.com). Тел.: (8172) 59-78-10.

Холодев Александр Юрьевич – инженер-исследователь отдела проблем научно-технологического развития и экономики знаний. Федеральное государственное бюджетное учреждение науки Институт социально-экономического развития территорий Российской академии наук. Россия, 160014, г. Вологда, ул. Горького, д. 56а. E-mail: der.goodwin@ya.ru. Тел.: (8172) 59-78-10.

**Skorodumov P.V., Kholodev A.Yu.**

## **ANALYSIS OF APPROACHES AND TOOLS IN THE ANALYSIS OF THE WEBSITE TRAFFIC OF A SCIENTIFIC ORGANIZATION**

*The website of a scientific organization is a crucial element in the process of promoting the activities of this organization in the Internet. Today, information technology is one of the most dynamically developing areas of science, technology and engineering. The use of the most modern and technologically advanced systems of websites organization does not guarantee their popularity; systematic work is required to promote, support and supplement them. It is also necessary to monitor and evaluate the results of the work performed for the purpose of making timely management decisions to adjust the directions of further development of the website. One of the objectives aimed at achieving this goal is to assess the popularity of the websites of a scientific organization by analyzing website traffic. The article presents an overview of existing methods for assessing the popularity of websites. Automated approaches are considered in detail. The authors compare the most popular systems of collection and analysis of website traffic statistics (Yandex.Metrics, Google Analytics, Liveinternet and HotLog). The article reflects wide functionality of these systems. It substantiates the decision of the Institute of Socio-Economic Development of Territories of RAS (ISED T RAS) to develop its own system for collecting and analyzing the statistics of visits to the websites. The paper describes the main parts of the system of collection and analysis of statistics of visits to the websites of ISED T RAS. It presents statistics on the results as of June and the beginning of 2015, obtained with the help of the Institute's own system for for collecting and analyzing the statistics of visits to the websites and also with the use of Yandex.Metric and Google Analytics. The article presents a comparison of the number of visits to the Internet portal of ISED T RAS obtained through different systems, and notes possible causes for discrepancies in the results. In conclusion, the authors note that the analysis of the statistics of visits to the websites of scientific organizations using different systems of data collection and analysis is especially important for modern scientific organizations, including from the viewpoint of position of the institution in webometrics rankings, and also the expediency of continuing the work in this area.*

*Website of a research organization, statistics, system of collection and analysis of statistics of visits to websites.*

### **REFERENCES**

1. Vebometrisheskii reiting Kiberneticheskoi laboratorii Tsentra nauchnoi informatsii i dokumentatsii Nacional'nogo issledovatel'skogo soveta Ispanii [Webometrics rank of a Cybernetic laboratory of the Center for scientific information and documentation, National Research Council of Spain]. *Cybermetrics Lab, Consejo Superior de Investigaciones Científicas*. Available at: <http://www.webometrics.info>
2. Vebometrisheskii reiting nauchnykh uchrezhdenii Rossii [Webometrics ranking of research institutions in Russia]. Available at: <http://webometrics-net.ru>

3. *Daidzhest Ofitsial'nogo portala IS RAN (dekabr' 2014 – fevral' 2015)* [Digest of the IS RAS official portal (December 2014 – February 2015)]. Available at: [http://www.isras.ru/files/File/publ/Digest\\_portal\\_zima\\_final.pdf](http://www.isras.ru/files/File/publ/Digest_portal_zima_final.pdf)
4. Zelenin A. S. Organizatsiya informatsionnogo prostranstva nauchnogo uchrezhdeniya (na primere ISERT RAN) [The organization of information space in the research institution (in the case of ISERT RAS)]. *Problemy razvitiya territorii* [problems of territory's development], 2012, no. 1 (57), pp. 119–127.
5. Kabakova E. A., Uskov V. S. Veb-sait nauchno-issledovatel'skogo uchrezhdeniya: napolnenie, posetiteli, razvitie [Website of research institute: content, visitors, development]. *Voprosy territorial'nogo razvitiya* [Territorial development issues], 2014, no. 3. Available at: <http://vtr.isert-ran.ru/article/1396/full>
6. *Ob utverzhdenii prioritnykh napravlenii razvitiya nauki, tekhnologii i tekhniki v RF i perechnya kriticheskikh tekhnologii RF: Ukaz Prezidenta RF ot 07.07.2011 № 899* [On approval of priority areas of development of science and technology in the Russian Federation and the list of critical technologies of the Russian Federation: Decree of the President of the Russian Federation of July 07, 2011 No. 899]. Available at: <http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=116178>
7. *Ofitsial'nyi sait kompanii Yandeks* [Official website of Yandex]. Available at: <https://metrika.yandex.ru>
8. *Ofitsial'nyi sait sistemy Google Analytics* [Official website of Google Analytics]. Available at: [http://www.google.com/intl/ru\\_ru/analytics/index.html](http://www.google.com/intl/ru_ru/analytics/index.html)
9. *Ofitsial'nyi sait sistemy HotLog* [Official website of HotLog]. Available at: <http://hotlog.ru>
10. *Ofitsial'nyi sait sistemy Liveinternet* [Official website of Liveinternet]. Available at: <http://www.liveinternet.ru>
11. Raizberg B. A., Lozovskii L. Sh., Starodubtseva E. B. *Sovremennyy ekonomicheskii slovar'* [Modern economic dictionary]. 5th edition, revised and supplemented. Moscow: INFRA-M, 2007. 495 p.
12. *Rukovodstvo po PHP* [The PHP manual]. Available at: <http://php.net/manual/ru/intro.session.php>
13. *Sistema upravleniya internet-portalom: svidetel'stvo № 2014612043 ot 17.02.2014* [Web portal management system: certificate No. 2014612043 dated February 17, 2014]. Available at: <http://www.vscs.ac.ru/storage/files/2014612043.pdf>
14. Skorodumov P. V., Badanin D. A. Sistema upravleniya proektami organizatsii: analiz podkhodov i sushchestvuyushchikh programmnykh reshenii [Project management system of the organization: analysis of approaches and existing software solutions]. *Voprosy territorial'nogo razvitiya* [Territorial development issues], 2015, no. 5. Available at: <http://vtr.isert-ran.ru/article/1505/full>
15. Tret'yakova O. V., Kabakova E. A. Vozmozhnosti ispol'zovaniya vebometricheskogo analiza v otsenke saita nauchnogo instituta [Opportunities for using webometric analysis when evaluating the website of a research institution]. *Voprosy territorial'nogo razvitiya* [Territorial development issues], 2014, no. 2. Available at: <http://vtr.isert-ran.ru/article/1389/full>
16. *Yii framework*. Available at: <http://www.yiiframework.com>

#### INFORMATION ABOUT THE AUTHORS

*Skorodumov Pavel Valeryevich* – Ph.D. in Engineering, Head of the Laboratory for Intelligent and Software-Information Systems at the Department of Scientific and Technological Development and Knowledge Economics. Federal State Budgetary Institution of Science Institute of Socio-Economic Development of Territories of Russian Academy of Science. 56A, Gorky Street, Vologda, 160014, Russia. E-mail: [spv.vscs@gmail.com](mailto:spv.vscs@gmail.com). Phone: +7(8172) 59-78-10.

*Kholodev Aleksandr Yur'evich* – Research Engineer at the Department of Scientific and Technological Development and Knowledge Economics. Federal State Budgetary Institution of Science Institute of Socio-Economic Development of Territories of Russian Academy of Science. 56A, Gorky Street, Vologda, 160014, Russia. E-mail: [der.goodwin@ya.ru](mailto:der.goodwin@ya.ru). Phone: +7(8172) 59-78-10.