

К. А. Колосов*ГПНТБ России*

Библиометрический анализ обращений к электронному каталогу ГПНТБ России: что ищет пользователь

Рассмотрены результаты расширенного библиометрического анализа, проведённого на основе запросов, поступивших к электронным каталогам ГПНТБ России в 2018 г. и первом полугодии 2019 г. Такой анализ позволяет составить рейтинг популярности запросов от разных источников – поисковых машин интернета и физических пользователей. Отмечено, что в настоящее время версия ЭК, работающая на WEB ИРБИС, доступна для читателей, однако основная доля обращений к ресурсу приходится на поисковые машины интернета. Сделан вывод о том, что поисковые машины чаще всего ищут полные тексты публикаций тех или иных авторов в интернете. Изучение запросов физических пользователей показало, что большинство удалённых читателей обращаются к электронному каталогу ГПНТБ России за библиографическими описаниями авторефератов диссертаций. Наличие в фондах библиотеки печатных изданий авторов с наибольшим рейтингом позволит потенциальным читателям получить доступ к их произведениям в случае удаления правообладателями электронной версии. Проанализированные издания, тексты которых доступны для скачивания, следует в первую очередь рекомендовать для размещения в свободном доступе в Национальной электронной библиотеке.

Статья подготовлена в рамках Государственного задания ГПНТБ России на 2019 г.

Ключевые слова: библиометрия, альтметрики, библиотечная статистика.

Kirill Kolosov*Russian National Public Library for Science and Technology, Moscow, Russia*

Bibliometrical analysis of requests to RNPLS&T's e-catalog: What the users search for

The author reviews the results of the extended bibliometrical analysis of the requests to the RNPLS&T's e-catalogs in 2018 and the first half of 2019. This analysis enables to obtain popular requests rating for various sources – the Internet search engines and individual users. The author concludes that the search engines more often search for full texts on the Internet. However, individual users mostly search for bibliographic records of these abstracts' in RNPLS&T's e-catalog. The printed publications by the authors with the highest rating in the Library's collection would enable potential users to access these publications if the owner removes the digital copy. The publications which texts are available for uploading are recommended to be made accessible free through the National Electronic Library.

The paper is prepared within the framework of the State Order of RNPLS&T's for the year 2019.

Keywords: bibliometrics, altmetrics, library statistics.

As of 2019, our library has the following servers, the log files of which can be used for the bibliometric analysis: electronic catalog (IRBIS 128) – the main version of the electronic catalog; the previous version of the electronic catalog, requests for which continue to come from robots and bots, as well as from some readers; electronic library directory – in a log file requests for the search for electronic publications from readers and search engines are recorded; catalog of full-text publications of the electronic library (IRBIS64+) – readers' requests for reading full-text publications, including the statistics of accesses to individual pages of electronic documents, are recorded in the log file. We will consider the results on the basis of queries to the electronic catalogs of the in 2018 and in the first half of 2019. Almost all requests came from Internet search engines, i.e. robots and bots and to a greater extent reflect the popularity of these authors in user requests for Yandex and Google search portals. These requests are associated with the availability of full texts of “top” authors and the possibility of downloading. The list of the most asked authors for 6 months of 2019 includes: A. A. Ruchev, S. A. Gaindo-Legaria, A. N. Stirrts, A. R. Khuziakmetova, K. Yu, A. V. Prokopyev, A. A. Beylbman, Joao Lopes Cabral.

Physical users turn to this version of OPAC more often than robots, since the structure of the URL request for this directory is different. The analysis of bibliographic records the authors presented above showed that abstracts predominate in the search results. It can be argued that a significant number of remote users turn to him for bibliographic descriptions when preparing publications. The following conclusions can be drawn: the largest number of calls comes from Yandex bots and corresponds to the popularity of Yandex requests to educational publications available for download; the user searches for full texts of publications available for download; readers are unlikely to contact print publications with electronic versions, but the presence of printed books in the library should allow potential readers to get them if the electronic version is deleted by copyright holders; analyzed publications, the full texts of which are available for download from the Internet, should be recommended for the public domain of the National electronic library in the first place.

Расширенный анализ лог-файлов веб-сервера является первым этапом библиометрического анализа, цель которого – получение альтернатив наиболее спрашиваемых авторов, публикаций и подписных изданий различными категориями удалённых читателей библиотеки. В отличие от стандартных средств анализа статистики он позволяет: формировать базу данных, в которой фиксируются все запросы от удалённых и физических пользователей, роботов и «пауков», организаций (по справочнику IP-адресов); записывать пользовательские сеансы [1, 2].

По состоянию на 2019 г. в ГПНТБ России имеются следующие веб-серверы, лог-файлы которых могут быть использованы для проведения расширенного библиометрического анализа:

электронный каталог (ИРБИС128) – основная версия ЭК, доступного на сайте библиотеки;

электронный каталог (WEB ИРБИС) – предыдущая версия ЭК, запросы к которой продолжают поступать от роботов и ботов поисковых систем, а также от некоторых читателей;

каталог электронной библиотеки (ЭБ) – в лог-файле фиксируются запросы на поиск электронных изданий, поступающие от читателей и поисковых систем;

каталог полнотекстовых изданий ЭБ (ИРБИС64+) – в лог-файле фиксируются запросы читателей на чтение полнотекстовых изданий, включая статистику обращений к отдельным страницам электронных документов.

Востребованность электронных изданий из фондов библиотеки с учётом изменения технологии после перехода на ИРБИС64+ будет проанализирована в следующих публикациях.

В этой статье мы рассмотрим результаты библиометрического анализа, проведённого на основе запросов к электронным каталогам ГПНТБ России в 2018 г. и в первом полугодии 2019 г.

В настоящее время версия ЭК, работающая на WEB ИРБИС, доступна для читателей, однако основная доля обращений к ресурсу приходится на поисковые машины интернета. В табл. 1 представлены наиболее спрашиваемые авторы по всем видам запросов в 2018 г., а также в 2017 и 2019 гг.

Таблица 1

**Наиболее спрашиваемые авторы по всем видам запросов
к веб-версии ЭК ГПНТБ России**

Автор	Число запросов		
	2018 г.	2017 г.	I полугодие 2019 г.
М. А. Рогов	2 960	27	76
Л. Н. Чечевицына	1 483	800	229
В. В. Лебедев	1 389	263	336
В. А. Буканин	1 344	158	400
Н. Г. Мэнкью	1 132	20	172
Г. Н. Федорова	1 106	103	314
Э. А. Арустамов	1 045	88	204
И. И. Мазур	989	99	289
Н. П. Кондраков	908	122	183

Представим изменение числа запросов к этим авторам в течение 2017–2019 гг. в виде линейной диаграммы (рис. 1).

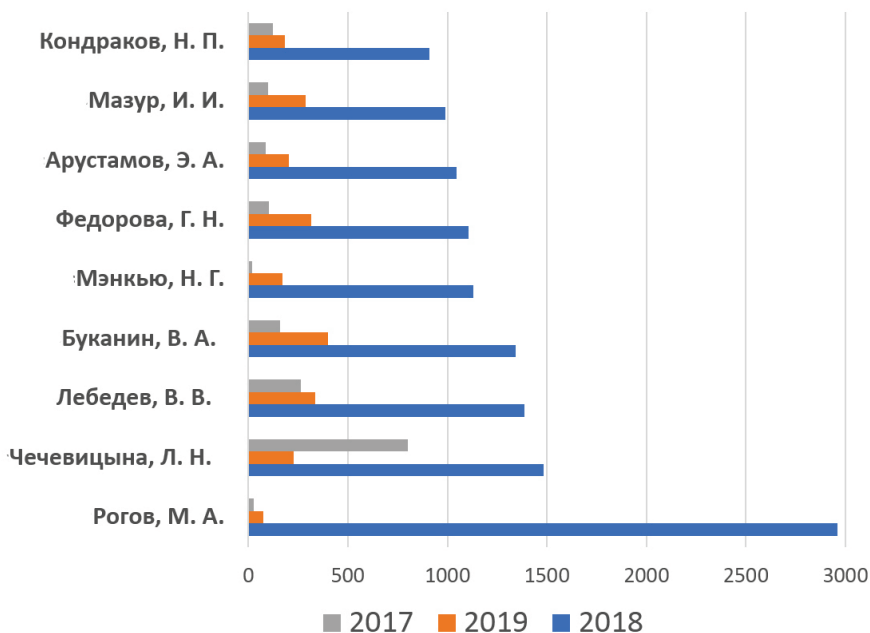


Рис. 1. Динамика изменений числа запросов к авторам в 2017–2019 гг.

По диаграмме, как и по аналогичным результатам анализа наиболее спрашиваемых авторов, понять причину столь резких изменений числа обращений к одним и тем же авторам не удалось. Следует учитывать, что практически все запросы поступили от поисковых машин интернета, т.е. роботов и ботов, и в большей степени отражают популярность названных авторов в запросах пользователей к поисковым порталам *Yandex* и *Google*.

Наиболее спрашиваемые авторы из топ-30 запросов меняются в течение рассматриваемого периода (рис. 2).

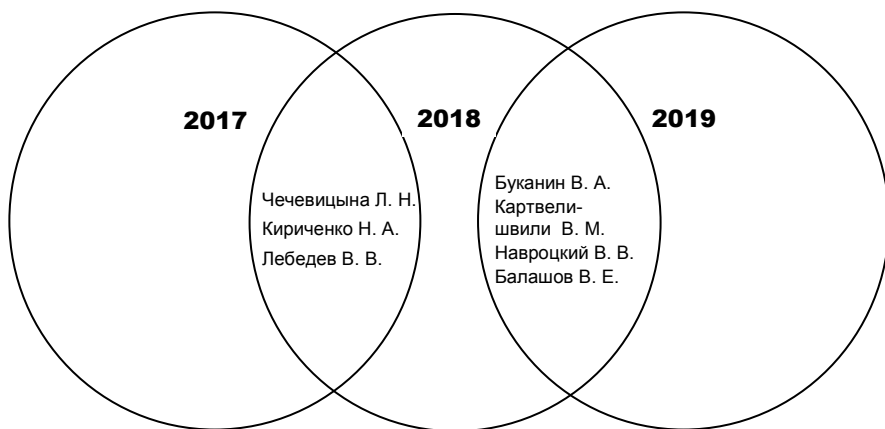


Рис. 2. Авторы из топ-30 запросов в 2017–2019 гг.

На следующем этапе библиометрического анализа была выдвинута гипотеза: запросы к ЭК от поисковых машин интернета связаны с наличием полных текстов «топовых» авторов и возможностью их скачивания пользователями. Рассмотрим, какие издания авторов представлены в ЭК ГПНТБ России, и выявим наличие их полных текстов в сети (табл. 2).

Таблица 2

Отдельные издания в ЭК ГПНТБ России из топ-30 запросов и наличие их полных текстов в сети

Запись в ЭК ГПНТБ России	Ссылка на полный текст издания, найденная в результате поиска в <i>Yandex</i>	Доступность для скачивания полного текста
Д8-01/61476 Рогов, М. А. Риск-менеджмент [Текст] / М. А. Рогов. – М. : Финансы и статистика, 2001. – 119 с. – 2000 экз. – ISBN 5-279-02379-5	https://www.twirpx.com/file/2281610/	Да, после регистрации
Д9-13/95157 Чечевицына, Л. Н. Анализ финансово-хозяйственной деятельности [Текст] : учеб. – 6-е изд., перераб. – Ростов н/Д : Феникс, 2013. – 368 с. – ISBN 978-5-222-20569-3	http://library.psu.kz/fulltext/buuk/b1216.pdf	Да

Запись в ЭК ГПНТБ России	Ссылка на полный текст издания, найденная в результате поиска в <i>Yandex</i>	Доступность для скачивания полного текста
Д8-01/78785 Буканин, В. А. Безопасность жизнедеятельности. Безопасность в чрезвычайных ситуациях. Пожарная безопасность : учеб. пособие / В. А. Буканин, В. Н. Павлов, А. О. Трусов. – СПб. : Изд-во СПбГТУ «ЛЭТИ», 2001. – 84 с. – 1000 экз. – ISBN 5-7629-0392-3	https://alleng.org/d/saf/saf94.htm	Да
Д8-94/14047 Мэнкью, Н. Г. Макроэкономика [Текст] / Н. Г. Мэнкью; общ. ред.: Р. Г. Емцов и др. – М. : Изд-во Моск. ун-та, 1994. – 735 с. : ил. – ISBN 5-211-03213-6	https://www.twirpx.com/file/8119/	Да, после регистрации
Д10-13/3569 Федорова, Г. Н. Информационные системы : учеб. / Г. Н. Федорова. – 4-е изд., стер. – М. : Академия, 2013. – 203 с. : ил. – ISBN 978-5-4468-0300-2	https://www.bolohovomt.ru/doc/informazionnie_sistemi.pdf	Да

Аналогичная информация была получена и в отношении остальных авторов из топ-30.

Таким образом, можно сделать вывод: запросы, поступающие от роботов и ботов поисковых машин интернета, отражают популярность полнотекстовых изданий авторов. Поскольку нет оснований полагать, что полные тексты были выложены в сеть с разрешения правообладателей, у библиотеки нет правовых оснований учитывать полученные результаты читательской активности в соответствующих библиографических записях ЭК. К сожалению, полные тексты этих изданий отсутствуют и в Национальной электронной библиотеке (НЭБ), хотя было бы логично разместить их там в первую очередь.

В следующей части библиометрического анализа рассмотрим, какие авторы попали в топ запросов по новой версии ЭК ГПНТБ России (ИРБИС 128). Список наиболее спрашиваемых авторов за 6 месяцев 2019 г. выглядит следующим образом: Ручьев А. А., Galindo-Legaria С. А., Штирц А. Н., Хузиахметова А. Р., Йу К., Прокопьев А. В., Бейлбман А. А., Жоао Лопеш Кабрал.

Результат, на первый взгляд, довольно неожиданный. Заметим, что к данной версии ЭК физические пользователи обращаются чаще, чем роботы поисковых машин интернета, поскольку структура URL-запроса у этого каталога принципиально иная по сравнению с WEB ИРБИС.

Анализ библиографических записей, найденных в ЭК ГПНТБ России по перечисленным выше авторам, показал, что в результатах поиска преоб-

ладают авторефераты. Можно утверждать: значительное число удалённых пользователей ЭК ГПНТБ России обращаются к нему за библиографическими описаниями при подготовке публикаций.

На следующем этапе анализа планируется сопоставить полученные выше результаты с книговыдачей печатных изданий из фондов библиотеки.

В заключение можно сделать следующие выводы:

наибольшее число обращений к ЭК ГПНТБ России поступает от ботов системы *Yandex* и соответствует популярности запросов *Yandex* к учебным изданиям, доступным для скачивания;

пользователь ищет полные тексты изданий, доступные для скачивания; обращение читателей к печатным изданиям, имеющим электронные версии, маловероятно, но наличие печатных книг в библиотеке должно позволить потенциальным читателям получить их в случае удаления электронной версии правообладателями;

проанализированные издания, полные тексты которых доступны для скачивания из интернета, следует рекомендовать для размещения в свободном доступе НЭБ в первую очередь.

СПИСОК ИСТОЧНИКОВ

1. **Земсков А. И.** Библиометрия в библиотеках / А. И. Земсков, К. А. Колосов // Науч. и техн. б-ки. – 2016. – № 11. – С. 5–23.

Zemskov A. I. *Bibliometriya v bibliotekah* / A. I. Zemskov, K. A. Kolosov // *Nauch. i tehn. b-ki.* – 2016. – № 11. – S. 5–23.

2. **Гончаров М. В.** Электронная библиотека ГПНТБ России: динамика пополнения, технологии, ресурсы / М. В. Гончаров, К. А. Колосов // Там же. – 2018. – № 12. – С. 34–41.

Goncharov M. V. *Elektronnaya biblioteka GPNTB Rossii: dinamika popolneniya, tehnologii, resursy* / M. V. Goncharov, K. A. Kolosov // *Tam zhe.* – 2018. – № 12. – S. 34–41.

3. **Гончаров М. В., Колосов К. А.** Разработка системы открытого архива ГПНТБ России // Там же. – С. 42–48.

Goncharov M. V., Kolosov K. A. *Razrabotka sistemy otkrytogo arhiva GPNTB Rossii* // *Tam zhe.* – S. 42–48.

Kirill Kolosov, *Cand. Sc. (Technology), Leading Researcher, Russian National Public Library for Science and Technology;*

kolosov@gpntb.ru

17, 3rd Khoroshevskaya st., 123298 Moscow, Russia