

И. И. Михайленко

ГПНТБ России

Технологии сбора статистической информации в библиотечных информационных системах и открытых архивах

Представлены результаты работ по развитию Открытого архива ГПНТБ России в части создания и совершенствования системы сбора и анализа статистической информации. Отмечено, что для выполнения основной функции ГПНТБ России – повышение качества библиотечно-информационного обслуживания всех категорий пользователей – необходимо формировать и анализировать использование фондов, баз и банков данных, включая электронные каталоги, специализированные тематические базы данных и базы данных научных периодических изданий. Для решения этих задач выстраивается система сбора и анализа статистической информации, которая на новом уровне обеспечивает возможность получения нового среза данных как для дальнейшей оптимизации работы информационных систем, так и для обслуживания пользователей в целом. Констатируется, что в результате внедрения новых технологий появилась возможность отслеживать потоки пользователей, оптимальные пути выдачи заказов, популярность тех или иных тематик, отдельных изданий. То есть появился мощный универсальный инструмент и технология анализа действующих в ГПНТБ России технологий и информационных систем, а также прогнозирования их развития с учётом реальных данных о поведении пользователей.

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

Ключевые слова: открытые архивы информации, статистическая информация, автоматизированная система ГПНТБ России, электронный каталог, электронная библиотека.

Ilya Mikhailenko*Russian National Public Library for Science and Technology, Moscow, Russia*

Technology of acquiring statistical information in library information systems and open archives

The author reviews the experience of RNPLS&T in developing the Open Archive and in particular, in building the system of statistical information acquiring and analysis. The author emphasizes that the RNPLS&T's main goal is to increase the standards of user library and information services, which is accomplished through building, maintaining, analyzing its collections, databases and databanks, including its e-catalogs, specialized topical databases and periodical databases in sciences. For this purpose, the system of acquiring and analyzing statistical information is being built. This system, at a new level, enables to acquire a different segment of data essential for further optimization of information systems and for user services, on the whole. The new technologies enable to keep track of user flows, efficient order processing, and demand for certain subjects and individual publications. The author concludes that this makes a powerful and universal instrument and technology for analyzing the RNPLS&T's technologies and information systems and forecasting their development with actual information on user behavior.

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

Keywords: open archive, statistical information, RNPLS&T automated system, e-catalog, e-library.

As part of the previous stages of development of the Open Archive, a transition was made to the new automated system of the State Public Library for Science and Technology of Russia. Its distinguishing feature is the provision of open access to all resources of the library based on a single access point. With this approach, the creation of a new system for collecting statistical information, which will allow analyzing the operation of the entire system in multivariate service schemes and technologies for delivering an information product to a user, is of particular importance. As a result of the work carried out as part of the state assignment, a new system was developed for collecting and analyzing statistical information accumulated in the library. It is a set of applied programs that allow the collection of statistical indicators. Each metric is determined by the following parameters: date; name of the statistical indicator; value. In addition to the daily collection of statistical indicators, special reports are created in which data are grouped in accordance with the reporting period. Monthly, quarterly, semi-annual reports are provided, as well as for 9 months and a year. Aggregation function is

indicated for each parameter when generating a report for constructing a report. Examples of such functions: “Sum”, “Average”, “Last”, “Maximum”, etc. As a result of the analysis of the main indicators, it was determined that for the automated collection of parameters the following can be used: the number of domestic books issued in the Book Fund subdivision – is calculated automatically based on information on book issuance from the library automation system; the number of issued domestic books in the entire library is calculated as the sum of the parameters “the number of issued domestic books in the Book Fund”, the number of issued domestic books in the “Reader Services Department” and several other similar units. Each parameter can have its own automated collection algorithm. Such algorithms run after a full day, i.e. “Today” is collecting data for “yesterday”. Each collection algorithm takes as a parameter the date on which data is collected (due to this, in case of failures that do not allow collecting information on any parameter, if the cause of the failure is eliminated, the algorithm will automatically restart and the necessary data will be collected).

Для обеспечения информационно-библиотечной поддержки и сопровождения отрасли образования и науки, выполнения основной функции ГПНТБ России – повышение качества библиотечно-информационного обслуживания всех категорий пользователей – необходимо формировать и анализировать использование фондов, баз и банков данных, включая электронные каталоги, специализированные тематические БД и БД научных периодических изданий, электронные носители информации, Электронную библиотеку ГПНТБ России [1].

Для решения перечисленных задач необходимо постоянно разрабатывать и совершенствовать технологии и современное программное обеспечение системы Открытого архива ГПНТБ России. Для выявления тенденций в использовании комплекса электронных информационных ресурсов нужно также постоянно проводить мониторинг и анализировать потребности и обслуживание пользователей, что является нетривиальной задачей [2].

В рамках предыдущих этапов развития Открытого архива ГПНТБ России был осуществлён переход на новую автоматизированную систему ГПНТБ России. Её отличительная черта – обеспечение открытого доступа ко всем ресурсам библиотеки на основе единой точки доступа. Если раньше пользователь был вынужден работать с разными интерфейсами составных частей информационной системы библиотеки (электронного каталога, электронного архива, электронной библиотеки), то после перехода на новую версию АС ГПНТБ России стал возможен доступ с помощью «единого окна», что, безусловно, более эргономично и удобно [3].

Особое внимание при проектировании и обосновании принципов построения Открытого архива и в целом новой версии АС ГПНТБ России было уделено реализации принципа единых подходов и технологий обслуживания пользователей вне зависимости от того, локальные они или удалённые, в какое отделение ГПНТБ России он пришли, какой документ запрашивают – печатный, электронный или мультимедийный.

При таком подходе особую важность приобретает создание новой системы сбора статистической информации, которая позволит проводить анализ работы всей системы в многовариантных схемах обслуживания и технологий доставки информационного продукта пользователю.

В результате работ, проведённых в рамках государственного задания, была построена новая система сбора и анализа статистической информации, накапливаемой в АС ГПНТБ России. Она представляет собой комплекс прикладных программ, позволяющих вести сбор статистических показателей ГПНТБ России, каждый из которых определяется следующими параметрами: дата, наименование показателя, значение.

Пример наименования параметра: «Общее количество посещений в стационарных условиях зала литературы по экологии». Каждый параметр может быть введён вручную, также его значение может быть собрано на конкретную дату автоматизированным способом.

Все параметры группируются по листам ввода: например, «Зал литературы по экологии». Необходимо это для того, чтобы сотрудники, занимающиеся учётом статистических показателей, видели только те из них, которые относятся к их зоне ответственности, и не могли менять другие.

Помимо ежедневного сбора статистических показателей создаются специальные отчёты, в которых группируются данные в соответствии с периодом отчёта. Предусмотрены ежемесячные, квартальные, полугодовые отчёты, а также за 9 месяцев и за год. Для получения итоговых цифр при построении отчёта для каждого параметра указана функция агрегации. Примеры таких функций: «Сумма», «Среднее», «Последнее», «Максимальное» и т.д.

Ежемесячные отчёты строятся на основе соответствующих ежедневных отчётов; квартальные – интегрируют данные соответствующих месячных отчётов. Аналогично полугодовые и отчёты за 9 месяцев формируются из соответствующих квартальных, отчёт за год – из полугодовых. В результате оптимизируется и минимизируется время, необходимое для построения отчётов.

Основные задачи при вводе такой системы в эксплуатацию – это формализация каждого из статистических показателей, определение для каждого из них возможностей для их автоматизированного сбора и написание соответствующих алгоритмов сбора.

Даже для параметров, собираемых вручную, формализация их сбора позволяет более чётко описать, что понимается под тем или иным параметром, и закрепить это в виде подсказки в поле ввода данных.

Семинары:	<input type="text"/>
Мероприятия:	<input type="text"/>
Экскурсии:	<input type="text"/>
▲ Выдача электронных документов и копий	
Обслуживание электронными носителями информации CD:	<input type="text"/>
Выдача копий фрагментов документов из Интернет-ресурсов и полнотекстовых баз данных (в бумажном виде):	<input type="text"/>
	Обслуживание электронными носителями информации CD Количество документов
	Количество документов
	Количество страниц

Пример формы ввода параметра вручную

В результате проведённого анализа основных показателей было определено, что для автоматизированного сбора параметров могут быть использованы:

количество выданных отечественных книг в подразделении «Книжный фонд» – рассчитывается автоматически на основе информации о книговыдаче из системы автоматизации библиотеки;

количество выданных отечественных книг во всей библиотеке – рассчитывается как сумма параметров «количество выданных отечественных книг в подразделении “Книжный фонд”», «количество выданных отечественных книг в подразделении “Отдел обслуживания читателей”» и ещё нескольких подобных по подразделениям ГПНТБ России.

Для каждого параметра может быть установлен собственный алгоритм автоматизированного сбора. Такие алгоритмы запускаются по прошествии полных суток, т.е. «сегодня» собираются данные за «вчера». Каждый алгоритм сбора принимает в качестве параметра дату, на которую происходит сбор данных (благодаря этому при устранении причины сбоя, не позволившего собрать информацию по какому-либо параметру, алгоритм перезапустится автоматически и необходимые данные будут собраны).

В результате внедрения АС ГПНТБ России появились новые возможности: получить новый срез данных, необходимых как для дальнейшей оптимизации работы информационных систем, так и для обслуживания пользователей в целом; потоки пользователей, оптимальные пути выдачи заказов, популярность тех или иных тематик, отдельных изданий.

Таким образом, созданы мощный универсальный инструмент и технология анализа действующих в ГПНТБ России технологий и информационных систем, а также прогнозирования их развития с учётом реальных данных о поведении пользователей.

СПИСОК ИСТОЧНИКОВ

1. **Открытый** доступ, открытые архивы и открытая наука : коллектив. моногр. / М. В. Гончаров, И. И. Засурский, А. И. Земсков [и др.] ; под общ. ред. Я. Л. Шрайберга ; Гос. публ. науч.-техн. б-ка России. – Москва : ГПНТБ России, 2017.

Открыту доступ, открытые архивы i otkrytaya nauka : kolektiv. monogr. / M. V. Goncharov, I. I. Zasurskiy, A. I. Zemskov [i dr.] ; pod obshch. red. Ya. L. Shrayberga ; Gos. publ. nauch.-tehn. b-ka Rossii. – Moskva : GPNTB Rossii, 2017.

2. **Гончаров М. В., Михайленко И. И.** Интеграция информационных ресурсов ГПНТБ России в рамках Системы открытого архива // Науч. и техн. б-ки. – 2018. – № 4. – С. 5–13.

Goncharov M. V., Mihaylenko I. I. Integratsiya informatsionnyh resursov GPNTB Rossii v ramkah Sistemy otkrytogo arhiva // Nauch. i tehn. b-ki. – 2018. – № 4. – S. 5–13.

3. **Гончаров М. В., Колосов К. А.** Разработка системы открытого архива ГПНТБ России // Там же. – № 12. – С. 42–48.

Goncharov M. V., Kolosov K. A. Razrabotka sistemy otkrytogo arhiva GPNTB Rossii // Tam zhe. – № 12. – S. 42–48.

Июа Mikhailenko, Senior Researcher, Russian National Public Library for Science and Technology;

irimikhaylenko@gmail.com

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