

М. В. Гончаров, К. А. Колосов

ГПНТБ России

Разработка системы открытого архива ГПНТБ России

Рассмотрена деятельность ГПНТБ России в области открытого доступа и открытых архивов информации. Приведены оценка разнородных информационных ресурсов ГПНТБ России, состав собственных цифровых ресурсов в 2018 г. К ним относятся: оцифрованные книги, размещённые в электронном архиве; статьи сотрудников ГПНТБ России, статьи из издаваемого ею журнала «Научные и технические библиотеки», из сборников материалов конференций, проводимых библиотекой, размещённые в ЭБ; информационные системы, поддерживаемые ГПНТБ России. Особое внимание уделено описанию поддерживаемых технологий доступа к цифровым ресурсам ГПНТБ России, включая различные поисковые системы как собственного, так и внешнего производства: доступ с использованием системы интегрированного поиска Summon, доступ к электронным каталогам библиотеки по протоколам Z39.50 и SRU/SRW и к цифровым ресурсам ГПНТБ России с портала НЭБ, доступ к метаданным ЭБ ГПНТБ России по протоколу открытых архивов OAI-PMH и ряд др. Отражено использование внешних сервисов интегрированного поиска. Дано подробное описание технологии доступа удалённых пользователей к отдельным подписным изданиям и лицензионным БД. Рассмотрено состояние разработки программных решений серверных модулей открытого архива ГПНТБ России с использованием Системы автоматизации библиотек ИРБИС64+.

Подчёркнуто, что цель всех перечисленных работ – дать пользователю библиотеки расширенный набор возможностей для информационного поиска и получения необходимых ему полнотекстовых изданий, в том числе и внешних.

Статья подготовлена в рамках Государственного задания ГПНТБ России на 2018 г. по теме 2.8441.2017/БЧ.

Ключевые слова: открытые архивы информации, автоматизированные библиотечно-информационные системы, поисковые системы, ГПНТБ России.

Mikhail Goncharov and Kirill Kolosov

Russian National Public Library for Science and Technology, Moscow, Russia

Development the RNPLS&T's Open Archive system

RNPLS&T's experience in providing open access and developing open archives is discussed. RNPLS&T's heterogeneous information resources are evaluated; the structure of the in-house digital resources in 2018 is described, including digitized books in RNPLS&T's Electronic Archive, employees' articles, articles in journals published by the Library, articles in the Library conferences proceedings, all hosted by the E-library, and the RNPLS&T's information systems. The technologies of access to RNPLS&T's digital resources are characterized, including in-house and commercial ones: access with Summon federated search system, access using Z39.50 and SRU/SRW protocols and access via the National Electronic Library access to the RNPLS&T's resources, access to RNPLS&T's E-library metadata using OAI-PMH protocol, etc. Integrated search external services are described. The technology of remote users access to individual subscriptions and licensed databases is specified. The current state of software solutions for server modules of the RNPLS&T's open archive based on IRBIS64+ ILS is characterized.

The article is prepared within the framework of the State Task Order for RNPLS&T for 2018 under the theme 2.8441.2017/BCh.

The goal is to provide users with the extended set of instruments for information retrieval and access to demanded full-text publications, including the external ones.

Keywords: open archives, integrated library and information systems, search systems, Russian National Public Library for Science and Technology, RNPLS&T.

The purpose of creating the open archive of the Russian national public library for science and technology is to provide access to all types of their own digital resources, as well as to the licensed research periodicals, to the National Electronic Library. The development of the open archive is the creation of a technological environment for storing and using open scientific data. The own digital resources of Library 2018 are: 1. Digitized books in the electronic archive. 2. Journal articles published by the library, proceedings of conferences held by the Library (8,322). 3. Other resources (number of records as of September 1, 2018): electronic catalog; Image catalogs; preprints; unpublished translations; Union catalog of ScieTech publications; UInion Catalog of EKBSON The "Scientific Archive" (392,528 full texts, and 2,678,490 processed metadata). Access to digital resources is provided via: 1. HTTP. Provided through the library website and websites of the supported information systems. 2. Access via the Summon system. 3. Access to electronic catalogs of the library via Z39.50 and

SRU/SRW protocols. 4. Access to metadata of electronic catalogs and the electronic library via the WEB-services technology. 5. Access via the OAI-PMH protocol. Licensing agreements allow to provide remote access to individual publications and databases, e. g. to the EBSCO databases: as well as to the Proquest database of dissertations. Access to full texts of licensed articles is carried out using a proxy server in authorization mode using a database of library readers and dynamic switching of proxy server / direct access modes. For authorization of users, a software module was developed and implemented, interfaced with the squid proxy server and the ILS IRBIS128. Development of software solutions for server modules of the open archive of the Russian National Public Library for Science and Technology using the ILS IRBIS64+. The system uses an electronic archive based on the XEROX DocuShare software solution and an electronic library module based on the ILS IRBIS64. In connection with the need to switch to Russian software, a new structure of open archive server modules is being developed using the IRBIS64+ software.

Цель создания открытого архива ГПНТБ России – предоставление пользователям доступа ко всем видам собственных цифровых ресурсов, а также к ресурсам научной периодики, в соответствии с заключёнными лицензионными соглашениями, к ресурсам Национальной электронной библиотеки (НЭБ), библиографическим базам данных других библиотек. Перспективная задача дальнейшего развития создаваемого открытого архива – создание технологической среды для хранения и использования открытых научных данных.

Состав собственных цифровых ресурсов ГПНТБ России в 2018 г.

1. Оцифрованные книги, размещённые в электронном архиве ГПНТБ России (табл.).

Показатель	Всего	Динамика пополнения по годам		
		2016	2017	2018
Количество оцифрованных книг	13 575	1 151	2 499	2 376
Количество страниц оцифрованных документов	3 032 973	301 118	692 723	379 696

2. Статьи сотрудников ГПНТБ России, статьи из издаваемого ею журнала «Научные и технические библиотеки», из сборников материалов конференций, проводимых ГПНТБ России, размещённые в электронной библиотеке, – 8 322.

3. Прочие собственные цифровые ресурсы ГПНТБ России (количество записей на 1 сент. 2018 г.):

электронный каталог – 1 388 324;

имидж-каталоги: отечественная книга – 711 791; зарубежная книга – 527 494; отечественная периодика – 272 371; зарубежная периодика – 392 451; препринты – 66 815; неопубликованные переводы – 428 569;

ИС СК НТИ – 1 223 712;

ИС ЭКБСОН – 46 399 335;

ИС «Научный архив» – 392 528 полных текстов документов, а также 2 678 490 обработанных метаданных.

Поддерживаемые технологии доступа к цифровым ресурсам ГПНТБ России

1. Доступ по протоколу HTTP. Предоставляется через веб-сайт библиотеки и сайты каждой из поддерживаемых информационных систем. Пользователям веб-сайта ГПНТБ России доступны для поиска и просмотра все её собственные цифровые ресурсы.

2. Доступ с использованием системы интегрированного поиска *Summon*. Предусмотрен поиск по ресурсам ЭК ГПНТБ России и по подписным БД научной периодики, к которым предоставляется доступ в читальных залах библиотеки.

3. Доступ к электронным каталогам библиотеки по протоколам Z39.50 и SRU/SRW. Позволяет предоставлять метаданные цифровых ресурсов ГПНТБ России сторонним информационным системам с целью проведения поиска, в том числе на корпоративных библиотечных порталах.

4. Доступ к метаданным электронных каталогов и ЭБ ГПНТБ России по технологии веб-сервисов. Позволяет использовать сторонние информационные порталы для проведения поиска по каталогам библиотеки.

5. Доступ к цифровым ресурсам ГПНТБ России с портала НЭБ.

Для передачи метаданных и цифровых ресурсов ГПНТБ России в НЭБ разработано программное решение на основе использования протокола OAI-PMH, которые планируется внедрить и в других библиотеках, использующих Систему автоматизации библиотек (САБ) ИРБИС64.

6. Доступ к метаданным ЭБ ГПНТБ России по протоколу открытых архивов OAI-PMH (<http://ellib.gpntb.ru/harvester3>). Реализован с использованием открытого ПО *Open Harvester*. Позволяет сторонним порталам осуществлять поиск и сбор метаданных ЭБ ГПНТБ России по технологии открытых архивов.

Технология доступа удалённых пользователей к отдельным подписным изданиям и лицензионным БД

Лицензионные соглашения с некоторыми поставщиками подписных изданий научной периодики позволяют предоставлять зарегистрированным читателям удалённый доступ к отдельным изданиям и базам данных. Так, ГПНТБ России имеет право предоставлять доступ удалённым читателям к БД *EBSCO: CASC, INSPEC* и *Science & Technology Collection*, а также к БД диссертаций *PROQUEST*. Доступ к полным текстам лицензионных статей осуществляется с использованием прокси-сервера в режиме авторизации с использованием БД читателей библиотеки и динамического переключения режимов «прокси-сервер»/«прямой доступ». Сценарий переключения задан в файле <http://ellib.gpntb.ru/user/proxy.dat>. Для авторизации пользователей разработан и внедрён программный модуль, сопряжённый с прокси-сервером *squid* и САБ ИРБИС128.

Разработка программных решений серверных модулей открытого архива ГПНТБ России с использованием САБ ИРБИС64+

Структурная схема взаимодействия программных модулей, обеспечивающих функционирование открытого архива ГПНТБ России, представлена на рис. 1.

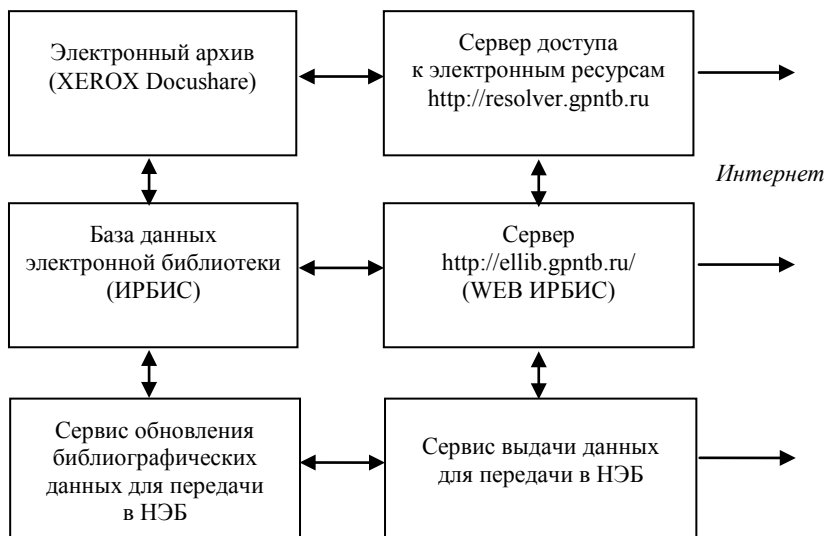


Рис. 1. Текущая структура серверных модулей открытого архива ГПНТБ России

В системе используются электронный архив на основе программного решения *XEROX DocuShare* и модуль ЭБ на основе САБ ИРБИС64. В связи с переходом на российское программное обеспечение ведётся разработка новой структуры серверных модулей открытого архива с использованием ПО САБ ИРБИС64+. На рис. 2 представлена структурная схема разрабатываемого решения.

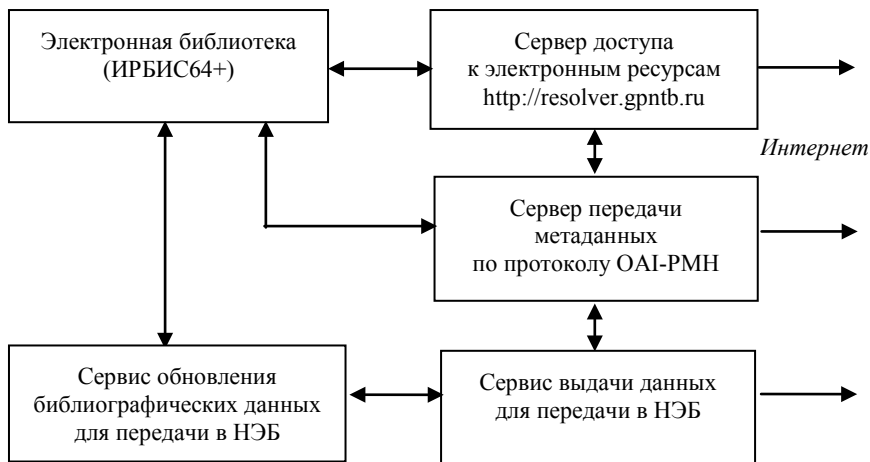


Рис. 2. Разрабатываемая структура серверных модулей открытого архива ГПНТБ России

ИРБИС64+ – это система, которая полностью включает возможности САБ ИРБИС64 и ИРБИС64 полнотекстовой БД, а также имеет дополнительный функционал, связанный со спецификой собственно электронной библиотеки.

В процессе разработки находится модуль обмена данными по протоколу OAI-PMN, внедрение которого позволит передавать метаданные из ИРБИС64+ сервис-провайдерам открытых архивов. Библиотеки вузов смогут более эффективно использовать программные средства для решения задач информационного обмена.

Поскольку ИРБИС64+ использует сервер баз данных семейства ИРБИС64, сохраняется преемственность с решениями, разработанными ранее для интеграции с внешними информационными системами. К таким решениям относится сервер Z64 (сервер Z39.50), который может работать по протоколу SRU/SRW, в том числе в режиме веб-сервиса.

***Mikhail Goncharov**, Cand. Sc. (Engineering), Associate Professor, Head, Perspective Research and Special Projects Group, Russian National Public Library for Science and Technology;*

goncharov@gpntb.ru

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

***Kirill Kolosov**, Cand. Sc. (Engineering), Leading Researcher, Perspective Research and Special Projects Group, Russian National Public Library for Science and Technology;*

kolosov@gpntb.ru

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