



ИПМ им.М.В.Келдыша РАН

Абрау-2017 • Труды конференции



С.А. Кириллов, И.Н. Соболевская,
А.Н. Сотников

**Актуальное состояние электронной
библиотеки «Научное наследие
России»**

Рекомендуемая форма библиографической ссылки

Кириллов С.А., Соболевская И.Н., Сотников А.Н. Актуальное состояние электронной библиотеки «Научное наследие России» // Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2017. — С. 250-259. — URL: <http://keldysh.ru/abrau/2017/59.pdf>
doi:[10.20948/abrau-2017-59](https://doi.org/10.20948/abrau-2017-59)

Размещена также [презентация к докладу](#)

Актуальное состояние электронной библиотеки «Научное наследие России»

С.А. Кириллов, И.Н. Соболевская, А.Н. Сотников

Кириллов Сергей Александрович – зав. сектором Межведомственного Суперкомпьютерного Центра РАН - филиала ФГУ ФНЦ НИИСИ РАН, Москва

Соболевская Ирина Николаевна – кандидат физико-математических наук, научный сотрудник Межведомственного суперкомпьютерного центра РАН – филиала ФГУ ФНЦ НИИСИ РАН, Москва

Сотников Александр Николаевич – доктор физико-математических наук, профессор, заместитель директора по научной работе Межведомственного суперкомпьютерного центра РАН – филиала ФГУ ФНЦ НИИСИ РАН, Москва

Аннотация. Рассматривается текущий статус электронной библиотеки «Научное наследие России» как с точки зрения содержания информационных фондов, так и состояния ее технологической компоненты. Формулируются основные принципы наполнения цифровых фондов электронной библиотеки, обеспечения связности представляемых в ней материалов. Обсуждаются вопросы комплектования центров сканирования и особенностей используемого в их составе оборудования.

Ключевые слова: электронная библиотека, интернет-портал, электронные фонды, цифровые копии.

В настоящее время цифровое информационное пространство развивается быстрыми темпами. Электронные коллекции библиотечных, архивных и музейных фондов представляют собой новый класс информационных ресурсов и играют важную роль в сохранении, развитии и предоставлении научно-образовательной и культурно-просветительской информации. На официальных сайтах различных библиотек, архивов и музеев размещены оцифрованные копии печатных изданий, архивных документов, музейных предметов, аудио- и видеоматериалов.

Многие библиотеки, архивы и музеи, располагающие коллекциями особо ценных книг, редкими изданиями — памятниками истории и культуры, или коллекциями музейных предметов, например, хранящихся в запасниках музея, сталкиваются с задачами по обеспечению сохранности и, в тоже время, доступности этих фондов широкому кругу пользователей.

Достичь этой цели можно с помощью современных технологий по созданию, сопровождению электронных фондов и их размещения в среде интернет.

Электронная библиотека "Научное наследие России" (ЭБ ННР) [1] представляет собой информационно-технологическую платформу для проектов интеграции электронных коллекций научно-образовательной и культурно-просветительской информации, прежде всего создаваемых в академической среде. Она обеспечивает широкому кругу пользователей интернет информацию о выдающихся российских, а также работавших в России зарубежных ученых, внесших значительный вклад в развитие фундаментальных, естественных и гуманитарных наук, содержит полные тексты опубликованных ими наиболее значительных работ, архивные материалы и музейные коллекции, связанные с их жизнью и деятельностью.

Основными целями ЭБ ННР являются:

- сохранение в цифровой форме научного наследия российских ученых, внесших значительный вклад в развитие науки;
- интеграция цифровых информационных фондов (библиотек, архивов, музеев) в единый информационный ресурс;
- популяризация научного знания.

При этом решаются следующие задачи:

- формирование, комплектование и хранение в электронном виде объектов библиотечного, архивного, музейного хранения;
- обеспечение локального и удаленного доступа различным категориям пользователей, реализация принципа «единого окна доступа» к распределенным информационным ресурсам;
- создание и реализация справочно-поискового аппарата по всему фонду электронной библиотеки с возможностью многоаспектного поиска.
- координация деятельности по созданию электронных библиотек, отражающих научное наследие, исключение дублирования в процессе оцифровки фондов;
- интеграция в единое российское электронное пространство знаний Национальной электронной библиотеки (НЭБ);
- следование положениям об охране авторских прав.

С архитектурной точки зрения ЭБ ННР является распределенной информационной системой, обеспечивающей развитие информационных фондов участников проекта на основе согласованных правил.

Реализованный программно-аппаратный комплекс библиотеки обеспечивает:

- сохранность уникальных информационных фондов и их доступность в среде Интернет;
- независимое развитие электронных библиотек организаций-участников с одновременной интеграцией фондов в единое информационное пространство;
- равные права участников проекта в вопросах формирования и использования фондов ЭБ ННР;
- децентрализованную подготовку электронных копий с централизованной диспетчеризацией;
- централизованное формирование и хранение множества метаданных;
- накопление, хранение и использование электронных документов;
- информационное обеспечение различных областей науки, образования и культуры.

В настоящее время ЭБ ННР развивается в нескольких ключевых направлениях.

В первую очередь это наращивание фондов за счет увеличения количества оцифрованных объектов, расширения номенклатуры информационного контента, включением в него объектов библиотечного, архивного, музейного хранения, а также их интеграции в единый тематический информационный ресурс. Для решения этой задачи необходимо, в том числе, использовать устройства оцифровки, позволяющие увеличить скорость сканирования, но при этом, сохранять максимальное качество получаемого цифрового объекта. Кроме того, в ЭБ ННР используются специально разработанные программные средства для более быстрого формирования электронной копии печатного издания и их интеграции.

Характеристика фондов электронной библиотеки «Научное наследие России»

Формирование фондов осуществляется под руководством Совета ЭБ ННР, определяющего основные приоритеты в отборе объектов для оцифровки и погружения в фонды электронной библиотеки. Эти приоритеты и правила являются обязательными для поставщиков цифрового контента, которыми являются научные библиотеки, архивы, музеи и научные институты РАН, а ныне ФАНО России [1].

В ЭБ ННР содержатся цифровые материалы, отражающие развитие российской науки, начиная с XVIII века.

Значительное внимание при выборе объектов для оцифровки и размещения в фондах ЭБ ННР уделяется безусловному исполнению Закона о защите авторских прав. На портале размещены копии изданий, которые не подпадают под действие закона о защите авторских прав, в том числе результаты научных исследований, опубликованные на иностранных языках.

Издания, если они защищены законодательством, имеются письменные согласия правообладателей.

Особенностью проекта является требование о создании развернутой биографической справки по всем авторам, произведения которых включены в ЭБ ННР. Алгоритм формирования объектов электронной библиотеки следующий: выбирается ученый, создается биографическая справка, затем оцифровываются и помещаются в ЭБ ННР его научные труды, архивные данные, сведения о посвящённых ему музейных коллекциях, а также другая информация.

В настоящее время (по состоянию на 17.07.2017) в технологической базе электронной библиотеки содержатся метаданные, описывающие более 5000 персон, из них сведения о 4553 ученых загружены на сайт.

Одновременно с вводом в электронный каталог биографических данных ведется работа по поиску внешних источников знаний по авторам. Электронная библиотека содержит ссылки на сайты музеев, архивов, а также различные электронные страницы в интернете, которые содержат значимую информацию о жизни выдающихся ученых, представленных в ЭБ ННР.

Технологически фонды электронной библиотеки «Научное наследие России» состоят из двух частей: опубликованных цифровых объектов и находящихся в работе. На текущий момент в технологический электронный каталог были введены метаданные на более чем 20000 изданий, из них прошли полную проверку и загружены на сайт 19486 изданий (3 644 108 страниц).

Распределение книг по языку издания. Как видно из таблицы 1 большая часть публикаций издана на русском языке. Кроме русского языка в библиотеке широко представлены книги на латинском языке, как правило, это российские академические издания XVIII века, и на французском языке, который преимущественно являлся языком научного общения в XIX веке. Издания на английском и немецком языках равномерно представляют XVIII, XIX и XX века.

| Язык | % |
|-------------|-------|
| Русский | 84,96 |
| Французский | 4,64 |
| Латинский | 4,09 |
| Английский | 2,11 |
| Немецкий | 2,65 |
| Прочие | 1,51 |

Таблица 1. Распределение публикаций по языку издания

При анализе содержания таблицы 1 следует учитывать, что одно издание может содержать статьи на разных языках.

Распределение книг по времени издания. В электронной библиотеке «Научное наследие России» широко представлены издания, отражающие все периоды развития научной мысли со дня основания Академии наук.

Представленные в таблице 2 данные характеризуют, в частности, развитие научной активности в России в различные периоды. Следует отметить, что резкое сокращение размещенных в фондах электронной библиотеки работ, со второй половины XX века по настоящее время, обусловлено нормами закона об авторском праве, которым руководствуются крупные библиотеки при поставке материалов для электронной библиотеки. Впрочем, этот период (вторая половина XX века -2017 год) вполне компенсируется изданиями научных институтов, поставляющих в электронную библиотеку работы ученых, созданных в рамках служебной деятельности в этих учреждениях, и с согласия автора или наследников.

| Период | % |
|-----------|------|
| 1700-1799 | 6,7 |
| 1800-1899 | 24,3 |
| 1900-1949 | 57,5 |
| 1950-1999 | 10,1 |
| 2000-2017 | 1,4 |

Таблица 2. Распределение публикаций по времени издания

Распределение публикаций по типу издания. Приведенные в таблице 3 сведения подчеркивают преобладание монографий как основного типа изданий в фондах ЭБ ННР.

| Тип издания | % |
|------------------------------------|----|
| Монография | 42 |
| Том многотомника | 15 |
| Статья из журнала, сборника трудов | 14 |
| Отдельный оттиск | 4 |
| Выпуск сериального издания | 4 |
| Сборник статей | 4 |
| Прочие виды публикаций | 17 |

Таблица 3. Распределение публикаций по типу издания

Распределение публикаций по рубрикам ГРНТИ. Для систематизации и поиска изданий в электронном каталоге ЭБ ННР применяется Государственный рубрикатор научно-технической информации (ГРНТИ).

При анализе таблицы 4 следует учитывать, что одно издание может иметь несколько значений рубрик ГРНТИ из разных тематических разделов, так и в рамках одного тематического раздела.

Пример: Жиров, Н.Ф. Люминофоры: (Светящиеся твердые составы)/ Н.Ф. Жиров; Под ред. С.И. Вавилова и Б.Я. Свешникова. - М. : Гос. изд-во оборон. пром-сти, 1940. Это издание определяется следующими рубриками ГРНТИ: 29.31.23; 61.69.39; 61.31.53; 47.09.43; 38.41.19

Особенно тщательно в нашем проекте (две и более рубрики ГРНТИ на одно издание) были рубрицированы издания по истории, философии и теологии, что иллюстрируется значениями, приведенными в таблице 4 .

| Рубрика ГРНТИ | Кол-во рубрик в базе технологического сервера метаданных в % |
|---|--|
| Другие | 5,15 |
| Теология. Атеизм | 12,56 |
| Исторические науки | 9,43 |
| Литературоведение. | 9,23 |
| Философия | 8,04 |
| Физика | 7,24 |
| Математика | 5,76 |
| Экономика. Экономические науки | 6,47 |
| Геология | 5,88 |
| Биология | 4,99 |
| Химия | 4,64 |
| Языкознание | 4,43 |
| Государство и право. Юридические науки | 4,02 |
| Сельское и лесное хозяйство | 2,52 |
| Механика | 2,19 |
| Астрономия | 1,39 |
| Политика. Политические науки | 1,31 |
| География | 1,03 |
| Культура. Культурология | 1,04 |
| Геофизика | 1,09 |
| Науковедение | 1,48 |

Таблица 4. Распределение публикаций по рубрикам ГРНТИ

В рамках проекта электронной библиотеки так же разрабатываются проблемно-ориентированные базы данных, в которых размещены материалы по отдельным направлениям науки, например, геологии, биологии, антропологии, этнографии и пр. В них предоставлены изображения, фотографии и видеоматериалы, входящие в различные коллекции институтов ФАНО России.

Анализ востребованности фондов ЭБ ННР [2] показывает растущий интерес научно-образовательного сообщества к данному ресурсу. Достаточно заметить, что пользователями библиотеки являются представители 132 стран, в год читатели просматривают более 2,5 млн. страниц.

Развитие технологической платформы проекта

Для проведения работ ЭБ ННР располагает необходимым специализированным оборудованием и сопутствующей инфраструктурой.

Комплекс аппаратных и программных средств, состоящий из рабочих станций и серверов обработки, позволяет выполнять весь спектр задач по формированию сопровождению электронной библиотеки, а именно отбор материалов для оцифровки, изготовление цифровой копии объекта, формирование библиографического описания, обеспечения сохранности цифровых объектов и их публикации на сайте библиотеки [3].

Сервер интернет-портала (<http://e-heritage.ru>), обеспечивающий функционирование ЭБ ННР, расположен в МСЦ РАН. Прикладные программные средства для загрузки, поиска и просмотра данных на сервере разработаны специалистами МСЦ РАН, ВЦ РАН и БЕН РАН.

Для создания электронной библиотеки печатная информация (например, страницы книги) переводится в цифровой вид с использованием различных типов сканеров отличающихся разрешающей способностью, возможностью сканирования различного типа изображений, скоростью сканирования, размерами обрабатываемых объектов и другими свойствами. В связи с этим в нашем проекте сформировались 5 основных типов комплексов сканирования.

1. Комплекс сканирования на основе планетарного сканера MINOLTA стал базовым на первом этапе становления проекта ЭБ ННР в 2006 году. Оператор сканера Minolta PS7000 оцифровывает черно-белые книги формата от А5-А3 с разрешением 600 dpi. Большие издания, формата до А2, сканируются с разрешением 400 dpi. Отсканировав книгу, оператор переходит к проверке и первичному редактированию файлов на прикрепленной к сканеру рабочей станции. При этом он использует встроенные функции программы SRZ Proscan, такие как: автоматическое или ручное выравнивание страницы по горизонтали, автоматическое удаление мелких (1-3 пикселя) черных и белых точек, обрезку страницы по шаблону типичной страницы.

2. Комплекс сканирования на основе цветного планетарного сканера ПланСкан С2-ЦА-600 постепенно заменяет устаревший комплекс на основе сканера Minolta ps 7000. Комплекс сканирования на основе цветного

планетарного сканера ПланСкан С2-ЦА-600 обладает широкими возможностями по формированию цифровых копий и обеспечивает удобства оператору-сканировщику. В частности, имеется возможность сканировать любые типы публикаций до формата А2 (в цветном, сером и черно-белом режиме). Книги формата А5-А3 сканируются с разрешением до 600 dpi. Издания формата А2 сканируются с разрешением до 400 dpi. Встроенная автоматическая книжная колыбель позволяет сканировать оригиналы толщиной до 10 см и весом до 6 кг. Имеется прижимное стекло с системой автоматического подъема. Программное обеспечение, встроенное в сканер, управляет процессом сканирования и обладает функцией коррекции изображения. Внедрение данной модели сканера позволило улучшить качество сканирования и снизить трудоемкость выполнения работ по получению сканов с различного вида печатных работ. В настоящее время этими сканерами оснащены участки сканирования МСЦ РАН, БЕН РАН, АРАН, ИЭА РАН, ИНИОН РАН в Москве, а также два центра сканирования в Санкт-Петербурге.

3. *Комплекс планетарного цветного сканирования ПланСкан А2-VC-B* по основным характеристикам аналогичен ПланСкан С2-ЦА, однако он более компактен и оборудован встроенной светодиодной системой подсветки. Комплекс ПланСкан А2-VC-B позволяет сканировать оригиналы до формата А2+ с разрешением до 600 dpi, обеспечивает следующие режимы сканирования: черно-белый - 2 бит, серый – 8 бит, цветной - 24 (30) bit/pix. Он имеет встроенную V-образную книжную колыбель с возможностью раскрытия до 180 градусов, сенсорную русскоязычную панель управления с графическим меню. Комплекс может использоваться без управляющего компьютера с просмотром отсканированного изображения на встроенном мониторе.

4. *Комплекс сканирования на основе планетарного сканера Пауэрскан Д14000 А0-20/25.* Основной особенностью работы данного комплекса является возможность сканировать материалы особо большого формата, с высокой разрешающей способностью, например, карты, плакаты, картины. Комплекс Пауэрскан Д14000 А0-20/25 с режимами сканирования - цветной, серый, черно-белый и максимальным форматом сканирования до А0,- позволяет сканировать, например, особо ветхие издания. Для этого комплекс имеет эргономичный стол сканирования, книжную колыбель для книг до 20 см толщиной и весом до 25 кг. А так же систему холодной подсветки, сканирующую головку ССD 3x14000 pix.

5. *Комплекс сканирования на основе сканера микрофиш Kodak ABR 3000 DSV* служит для оцифровки материалов, хранящихся на микрофишах и микроплентке. Сканер Kodak ABR 3000 DSV имеет объектив «8731408» с увеличением 7,5х; объектив «1677293» с увеличением 13-27х; ручную универсальную каретку «3846201» UC-5 для микрофиш, микроплентки 16/35 мм и апертурных карт с ручной загрузкой и позиционированием микрофильмов.

Применение указанных комплексов сканирования позволяет выполнять максимально точный перевод публикаций в цифровую форму.

Еще одно важное направление развития ЭБ ННР – это интеграция с другими информационными ресурсами [4]. ЭБ ННР предоставляет ряд сервисов, предусматривающих интеграцию, как непосредственно с «непечатным» контентом (аудио-, видеоматериалы, музейные предметы, архивные документы), так и с информационными ресурсами других информационных систем. На сегодняшний день, одной из приоритетных задач развития ЭБ ННР является интеграция с Национальной электронной библиотекой (НЭБ) в направлении построения единого российского электронного пространства знаний.

Заключение

Сегодня Электронная библиотека «Научное наследие России» представляет собой интегрированный информационный ресурс, включающий не только биографические сведения о персоналии, печатные и архивные документы (как самих ученых, так и документы, посвященные научным проектам, научным событиям, научным школам и т.д.), но и связанные с ними цифровые коллекции музейных предметов, а также других внешних источников. Средствами электронной библиотеки обеспечена интеграция разнородных материалов, в том числе электронных копий объектов библиотечного, архивного и музейного хранения, представленных в виде текстов, графических образов, аудио-видео объектов.

Пользователи имеют возможность:

- найти электронную копию печатного издания, архивного документа, объекта музейного хранения, связанного с той или иной персоной;
- прочитать полный текст найденной работы;
- познакомиться с цифровыми копиями некоторых музейных коллекций;
- посмотреть фото- видео- материалы, представленные на виртуальных выставках, созданных в рамках совместных проектов с государственными естественнонаучными музеями и архивами;
- использовать сервисы, предоставляемые ЭБ ННР на платформах Android и iOS.

Как уже отмечалось, важным направлением развития ЭБ ННР является сотрудничество с Национальной электронной библиотекой в качестве интегратора научно-образовательных информационных ресурсов организаций информационной сферы, подведомственных ФАНО России.

Данная работа выполнена в рамках фундаментальных научных исследований по теме (проекту) 0065-2014-0103 «Исследование и разработка технологий, методов и средств формирования и использования, интегрированных электронных информационных ресурсов».

Литература

1. Н.Е. Калёнов, Г.И. Савин, В.А. Серебряков, А.Н. Сотников. Принципы построения и формирования электронной библиотеки "Научное наследие России" // Программные продукты и системы, 2012. - Т. 4, № 100. - С. 30-40.
2. К.П. Погорелко. Анализ востребованности электронной библиотеки «Научное наследие России» // Информационное обеспечение науки: новые технологии: Сб-к научных трудов / Каленов Н.Е., Цветкова В.А. (ред.) -М.: БЕН РАН, 2015. - С. 191-199.
3. С.А. Кириллов, А.Н.Сотников. Подготовка электронных изданий для электронной библиотеки «Научное наследие России» // Информационное обеспечение науки: новые технологии. Сборник научных трудов / Каленов Н.Е., Цветкова В.А. (ред.) -М.: БЕН РАН. 2015. - С. 178-190.
4. Н.Е. Калёнов, И.Н. Соболевская, А.Н. Сотников. Интеграция музейной информации в электронную библиотеку «Научное наследие России» // Роль музеев в информационном обеспечении исторической науки. Сборник научных трудов, авт.-сост. Е.А. Воронцова; отв. ред. Л.И. Бородкин, А.Д. Яновский. М. Этерна, с.445-450, 2015.