

## СЕКЦИЯ «КОРПОРАТИВНЫЕ ПРОЕКТЫ»

### **Создание научно-образовательных ресурсов в Интернете библиотеками РАН (базы данных, электронные коллекции и библиотеки)**

*Госина Людмила Игоревна, вед. н. с., д-р филол. наук,  
Библиотека по естественным наукам РАН, филиал в  
Математическом институте им. В.А. Стеклова РАН,  
Москва*

*Королькова И.И., гл. библиотекарь, Библиотека по  
естественным наукам РАН, филиал в Математическом  
институте им. В.А. Стеклова РАН, Москва*

*Погорелко К.П., вед.н.с., канд. техн. наук, Библиотека по  
естественным наукам РАН, филиал в Математическом  
институте им. В.А. Стеклова РАН, Москва*

*В докладе рассказывается об академических (БЕН РАН) электронных ресурсах и проектах, в принципе совпадающих с проектами АРБИКОНа «МАРС» и «ЭПОС», но предусматривающих свободный доступ к ресурсам. Эти проекты направлены на создание универсальных по тематике или отраслевых (математика) библиографических БД, полнотекстовых коллекций и электронных библиотек, доступных всем пользователям в Интернете. Работы финансируются по научным программам РАН, грантам РФФИ и РГНФ. Предлагается расширить проект «КОРСАР» и создать информационную страницу, рассказывающую об издательской продукции библиотек, в т.ч. электронной, и позволяющую сделать заказ.*

Проект «МАРС», результатом которого является политематическая библиографическая БД отечественных журнальных статей, достаточно точно отвечает общественным потребностям в «ключе» к отечественным публикациям. Эта потребность очень высока и в научной среде. Поэтому в Библиотеке по естественным наукам (БЕН РАН) несколько библиотек и научных центров ее сети уже свыше 15 лет создают и поддерживают корпоративную библиографическую БД «Наука России». Эта система поддерживается РФФИ. Она включает библиографические описания отечественных статей по биологии, химии, частично физике, прикладной математике, библиотековедению и информатике. Ее объем в настоящее время превышает 30 тыс. записей и она доступна в Интернете на сайте БЕН РАН ([www.benran.ru](http://www.benran.ru)). Но принцип ее формирования – иной: БД строится на основе персоналий, т.е. публикаций академических ученых. Не ведется

сквозная роспись отечественных журналов и поэтому не собирается полный массив библиографических данных.

В структуре БЕН РАН уже более семи лет реализуется проект создания отраслевой информационной системы по математике. Библиотека Математического института им. В.А. Стеклова РАН, которую мы представляем, формирует массив данных о журнальных статьях по физико-математическим наукам. Основные отечественные профильные журналы распределены между тремя академическими партнерами, из которых 2 библиотеки (Математического института им. В.А. Стеклова РАН и ВЦ РАН) являются филиалами БЕН, а библиотека С.-Петербургского отделения Математического института им. В.А. Стеклова входит в систему БАН. Библиографические описания журнальных публикации вводятся в единую БД и отражаются в журнальных каталогах библиотек, где появляется пометка, что номер журнала расписан и можно посмотреть оглавление.

Этот проект задумывался в начале 2000-х гг. и основывался на идее создания отраслевых информационных систем в едином информационном пространстве. Мы полагали, что отраслевая библиографическая БД с высокой степенью надежности удовлетворит часть информационных потребностей и позволит в дальнейшем создать полнотекстовую БД, если редколлегия изданий захотят это сделать. БД создается профессиональными библиотекарями, оперативно пополняется (пока журналы экспонируются на выставке новых поступлений), соответствует нормам библиографического описания статей, включает адекватные описания формул, находится в свободном доступе.

Мы понимали, что малое количество участников длительный период времени будет накапливать информацию, а ученым для работы нужен массив библиографических записей глубиной хотя бы 10 лет. Поэтому начальным информационным ядром, с которым мы в дальнейшем работали, стал фрагмент БД Всероссийской книжной палаты, который мы приобрели в 2003-04 гг. Этот массив включал оглавления 35 математических журналов за различные периоды времени, но не ранее 1990 г.

Важными вопросами являются полнота БД и качество библиографических записей. Первые проверки качества информации не выявили грубых ошибок. Но когда мы начали проверять полноту оглавлений конкретных номеров и журнальных комплектов, полноту авторских коллективов в статьях с несколькими соавторами, отображения формул в названиях статей, - стало понятно, что весь массив записей (40-45 тыс.) требует редакции «de visu», чем мы продолжаем заниматься и в настоящее время. Наиболее распространенными ошибками, помимо многочисленных опечаток, был пропуск соавторов, слов в названиях публикаций, отсутствие формул и математических символов в заглавиях работ, пропуск отдельных статей и наличие не расписанных номеров почти во всех журналах. Особенно низким было качество библиографической росписи такого уважаемого и востребованного политематического журнала как «Доклады Академии наук». Учитывая высокий научный авторитет и заинтересованность очень широкого

круга ученых в этом издании, для него были приняты решения: во-первых, не выделять математическую тематику, а расписать его полностью; и, во-вторых, ввести формулы в названия статей по физике и химии. Это трудоемко, но удобно и академическим пользователям, и высшей школе.

Список журналов распределен между библиотеками-партнерами, и сотрудники не только осуществляют ввод новой информации, но и выверку более старых выпусков. Около 70% ввода осуществляет наш коллектив, располагающий большим штатом и квалифицированным персоналом. Руководит работой этого небольшого консорциума заведующий нашей библиотекой и специалист по информационным системам и базам данных к.т.н. К.П. Погорелко. Он же является разработчиком программного обеспечения для наших проектов.

На первом этапе редактирования приобретенной БД были проверены более новые публикации (от настоящего времени - к 2000 г.), сейчас проверяются и редактируются описания статей 1990-2000 гг. По нашим оценкам до 30% записей были отредактированы или введены заново. Для перевода описаний использовалась БД «МАРС», пока она была доступна в Интернете (копировались описания статей, что ускоряло работу). К настоящему времени наша БД включает свыше 100 тыс. статей из 40 отечественных физико-математических научных журналов, в т.ч. 16 названий не расписываются участниками консорциума АРБИКОН. Свободный доступ к ресурсу осуществляется с сайта нашей библиотеки ([libserv.mi.ras.ru](http://libserv.mi.ras.ru), кнопка «Межбиблиотечный каталог», затем «Статьи в отечественных математических журналах» - и, нажав кнопку «Выбрать журнал», там можно увидеть список расписываемых журналов), а также наших партнеров - академический математический портал ([www.mathnet.ru](http://www.mathnet.ru)).

Третьим важным вопросом являются поисковые возможности БД и сервис для пользователей. Мы подразделяем поисковые возможности на технические и содержательные. К техническим возможностям относятся набор поисковых полей и возможность поиска по любому элементу (автор, заглавие, год, название журнала) или фрагменту библиографического описания публикации. К содержательным относятся дополнительные возможности, отсутствующие при поиске по оглавлению отдельно взятого номера журнала. Эти возможности появляются при некотором дополнительном вводе информации, основанном на осмыслении логики поиска сведений о персоналиях, конкретных работах и иного фактического материала.

Во-первых, полностью раскрывается авторский коллектив, вне зависимости от количества участников. Во-вторых, даются сведения о переименовании или продолжении какой-либо работы, поправках и исправлениях в расчетах, опечатках к статьям, которые часто публикуются через несколько номеров после выхода основного текста, но содержат важную информацию, неразрывно связанную с основным текстом. В-третьих, учитываются все мелкие рубрики в научных журналах, включающие письма авторов и читателей, материалы о юбилеях, наградах, премиях,

дополнительно раскрываются разделы с обсуждениями докладов и иных научных работ (в оглавлениях присутствует рубрика без фамилий участников). В-четвертых, сведения о переименовании или расширенной публикации какой-либо работы (выход части 2, 3-й и т.д. вносится в оглавление номера, в котором вышла первая часть). Поэтому при поиске по фамилии автора будет получена более полная информация, часто необходимая его коллегам, оппонентам, ученикам и нам, библиотечным и информационным работникам.

Существенно для поисковых возможностей БД, что современные издательские технологии имеют не только плюсы, но и минусы. Например, редакция Журнала Экспериментальной и Теоретической Физики (ЖЭТФ) имеет возможность бесплатно рассылать оглавления выпусков журнала по электронной почте еще до выхода в свет печатного издания. Такая опережающая информация интересна многим ученым. Но если в ходе редактирования меняется название работы (а такие случаи есть), после выпуска номера существенно затруднен, а порой невозможен для пользователя поиск этих статей в Интернете. Наши сотрудники стараются корректировать все информационные расхождения путем отсылок и примечаний. Таким образом, качество создаваемой БД и ее поисковые возможности возрастают и круг пользователей постоянно растет.

Следующий из проектов АРБИКОНа, имеющий аналоги в библиотеках РАН – это «ЭПОС», дающий доступ к использованию полнотекстовых документов партнеров и позволяющий объединять разнородные коллекции в единое собрание электронных ресурсов. Библиотеки РАН основным направлением в информационной деятельности считают создание универсальных по тематике или отраслевых порталов с различными возможностями поиска нужных записей, текстов, фактов. Производство электронных полнотекстовых БД долгое время не считалось задачей библиотек, поскольку за рубежом этим направлением занимаются издательства. Однако уже два года БЕН РАН, ряд библиотек ее сети, БАН и библиотеки ее сети, Архив РАН, ИНИОН и Суперкомпьютерный центр при Президиуме РАН совместно работают над созданием электронной политематической библиотеки «Научное наследие России». Хранители сканируют литературу и важные исторические документы в библиотеках и архиве, группа разработчиков работает над созданием программного обеспечения для архивирования, презентации, поиска и пользования электронными текстами. Научная программа предполагает, что труды российских ученых XVIII-начала XX вв., многие из которых давно стали библиографической редкостью, в результате этой работы будут доступны всему научному сообществу, а также преподавателям и учащимся высшей школы в Интернете. В этом году будет открыт доступ к первому кругу сканированных текстов на сайте БЕН РАН. Если проект будет продолжаться, в итоге мы получим фундаментальную общедоступную ЭБ, интересную для многих университетов.

В научно-образовательных целях в библиотеке Математического института им. В.А. Стеклова РАН создается электронная коллекция «История Российской академии наук в документах» (проект поддержан РФНФ). Оцифровано 10-томное собрание «Материалов для истории Императорской Академии Наук», включающее документы первой половины XVIII в., сейчас ведется работа с двухтомной «Историей» П. Пекарского и 8-томной «Историей Российской Академии» М.И. Сухомлинова. Эти издания XIX-го века давно стали библиографической редкостью. Собрание будет дополнено рядом отчетов по физико-математическому и филологическому отделениям начала XX в. (дореволюционный период). Сканированные тексты редактируются с применением программных средств, разработанных К.П. Погорелко. В итоге электронная коллекция будет содержать интересный и разнообразный исторический материал, включающий сведения о выдающихся деятелях российской науки и культуры, который будет полезен и преподавателям, и учащимся, и ученым. Коллекция будет доступна в Интернете на сайтах библиотеки Института, БЕНа и ряда других академических учреждений. Также предполагается выпуск этой коллекции на CD-ROM и свободное распространение записанной версии по центральным, региональным и университетским библиотекам.

В этой связи хотелось бы перейти к издательскому проекту АРБИКОНа «КОРСАР», который нацелен на пополнение библиотечных фондов традиционными изданиями путем замены ветхих экземпляров новым переизданием спрашиваемого произведения. Это хороший проект, и он решает конкретную проблему определения списка спрашиваемых изданий и их переиздания для библиотек.

Однако, с нашей точки зрения, этот проект должен быть расширен или дополнен самостоятельным направлением, помогающим библиотекам узнать об издательских проектах друг друга и записаться на получение нужных публикаций. Это могут быть электронные издания на CD-ROM, подобные нашей тематической коллекции, а могут быть и печатные издания, репринты, библиографическая продукция, создаваемая библиотеками на гранты или государственные средства. О ее появлении не известно широкому кругу библиотек, т.к. нет общеизвестного сайта, оповещающего о выпуске подобных материалов, нет удобных механизмов заказа и получения этих изданий. Между тем, многие библиотеки подают заявки на создание электронных коллекций, создание разнообразных библиографических указателей и получают гранты. По условиям фондов, опубликованные на их средства материалы должны широко распространяться по библиотекам, но выбираются эти библиотеки в значительной мере случайно. Так, в наш адрес неоднократно присылали совершенно непрофильные материалы как авторы, так и сотрудники фондов. Возможно, эти книги были нужны другим учреждениям. Поэтому, как мы полагаем, на сайте АРБИКОНа целесообразно открыть отдельную страницу, чтобы библиотеки-издатели могли размещать информацию о своей продукции, а заказчики оставили

заявку на желаемое количество экземпляров. Эта страница должна быть доступна всем желающим.