

## Информационные ресурсы АРБИКОН - пора подумать об их качестве

Автор: Воройский Ф. С., Острая С. А.

Воройский Ф. С., Острая С. А.

ГПНТБ России

За два года существования Ассоциации региональных библиотечных консорциумов (АРБИКОН) с учетом предшествующих двух лет становления вступивших в нее двенадцати региональных корпоративных библиотечно-информационных систем (КБИС) создан весьма представительный информационный ресурс и предоставлен в широкий доступ через Интернет разным категориям пользователей Российской Федерации, ближнего и дальнего зарубежья. Сегодня электронные каталоги участников АРБИКОН содержат десятки миллионов библиографических записей (БЗ), которые (пока еще только потенциально) могут удовлетворить информационные потребности пользователей, представляющих самые разные слои населения и сферы деятельности.

Для примера приведем некоторые количественные данные по одной из участниц АРБИКОН - Корпоративной сети московских библиотек (КСМБ).

1. Объем сводного каталога по категориям ресурсов (вне зависимости от способа организации каталога - централизованный или распределенный):

ЭК неперIODических изданий - 2463,72 тыс. БЗ,

ЭК периодических изданий - 194,43 тыс. БЗ,

ЭК аналит. росписи статей - 930,46 тыс. БЗ,

ЭК аудиозаписей - 2,91 тыс. БЗ,

ЭК мультимедиа - 2,79 тыс. БЗ,

ЭК киноvideоматериалов - 3,73 тыс. БЗ,

ЭК нот - 22,91 тыс. БЗ,

ЭК, прочие типы документов - 6,55 тыс. БЗ,

*Всего - 3703 тыс. БЗ.*

2. Объем каталогов, доступных по Z39.50 в сводных каталогах регионального консорциума - 1926,96 тыс. БЗ.

3. Объем ЭК участников консорциума, доступных в Интернете по веб - 2637,58 тыс. БЗ.

4. Объем полнотекстовых электронных ресурсов участников КСМБ, доступных в библиотеках:

\* неперiodические издания и серая литература - 8728 БЗ, в том числе учебники, учебно-методические пособия - 2897 БЗ;

\* авторефераты диссертаций - 5158 БЗ;

\* статьи, доклады - 12135 БЗ;

\* справочники, энциклопедии, словари - 459 БЗ;

\* прочие типы документов (собрания сочинений писателей-классиков художественной литературы, проектная и организационно-технологическая документация КСМБ, редкие издания, электронные версии сборников, журналов и др.) - 114 БЗ.

### **Информационные ресурсы АРБИКОН глазами пользователей**

*В соответствии с Уставом АРБИКОН [1] целью его создания и функционирования является "оказание содействия российским библиотекам в предоставлении более качественного информационного обеспечения пользователей на основе внедрения современных библиотечно-информационных технологий, модернизации управления библиотечными ресурсами, в первую очередь при обработке поступающей литературы, а также при оперативном использовании ресурсов удаленных библиотек страны на основе доступа через Интернет".*

*Ничуть не стремясь хоть как-то принизить большие усилия и немалые успехи, достигнутые руководством АРБИКОН и его участниками в обеспечении информационной, программной и технической совместимости всех входящих в него автоматизированных библиотечных систем, посмотрим, что может получить пользователь при поиске в среде электронных каталогов АРБИКОН и входящих в его состав КБИС.*

*В ходе долговременной работы с автоматизированными библиотечно-информационными системами нам не встречались их конечные пользователи, которые производили бы поиск по классификационным индексам разделов рубрикаторов. Как правило, поиск производится с использованием ключевых слов и фамилий авторов, иногда дополняемых другими признаками, которые могут входить в библиографическое описание искомого документа (например год издания).*

*Заголовки документов и сами рубрики весьма поверхностно отражают содержание документа. Приведем только один достаточно типичный пример составления БЗ,*

наполняющих ЭК участников АРБИКОН:

*Автор: Пикуль В. С.*

*Заглавие: Исторические миниатюры: В 2 т.*

*Издано: М.: Молодая гвардия, 1991*

**ББК: Р2**

*Аннотация: В настоящий сборник включены исторические миниатюры, в основном не опубликованные ранее.*

*На самом деле сборник содержит интереснейшие и малоизвестные факты из жизни более чем 100 российских (и не только!) выдающихся деятелей XVIII-XIX вв., оказавших существенное влияние на самые разные стороны развития (государственного, политического, социального, экономического, военного, культурного и т.д. и т.п.) России и ее отношений со странами Западной Европы и Азии. Причем сведения эти почерпнуты из архивных документов, личных писем, дневников, воспоминаний и других свидетельств как самих героев, так и их современников. В случае полноценного индексирования этой книгой заинтересовались бы тысячи самых разных пользователей, а не случайно вышедшие на нее читатели (подобно авторам этого доклада).*

*Попытка найти документы по тематическому запросу к распределенным ресурсам Ассоциации через портал АРБИКОН. Тема: "Технологии очистки сточных вод". Поисковое предписание (ПП) - ко всем полям, язык любой, в каждой строке ПП по одному слову: очистк\*, вод\*, сточн\*, логические операторы - "И". Результат ждем 21 минуту - система "молчит". Повторяем тот же запрос, убрав из ПП слово сточн\*. Результат - система предлагает произвести поиск в БД аналитических росписей журналов и газет РКП. Обращаемся в РКП - разные варианты поиска (по БЗ, аннотациям и т.п.) дают нулевую выдачу. Возвращаемся в режим поиска по ресурсам АРБИКОН (сделать теперь это можно только через главное меню с потерей установок предыдущего ПП). Формулируем запрос: экология (условия: везде, русский, по всем полям). Система практически мгновенно снова рекомендует провести поиск в БД РКП. Там результат поиска - 0. Общие затраты времени на поиск с нулевым результатом составили 2,5 часа.*

*Многokrатно проведенный поиск через портал АРБИКОН показал, что время реакции системы от момента нажатия клавиши "поиск" до выдачи на экран его результатов составляет в среднем ~5 мин. Сказанное относится к обработке ею простейших запросов, обращенных только к одной КБИС - участнице Ассоциации. Для ознакомления с содержанием большого числа БЗ и отбора из них релевантных, как правило, необходимо значительное время, тогда как система через ~10 мин. отключает выполнение запроса и предлагает произвести новое соединение. После этого снова приходится вводить поисковое предписание и далее все повторяется. Причем в результате повторения запроса система может выдать количество записей,*

*существенно отличающееся от предыдущей выдачи (иногда в несколько раз).*

### *Состояние технологии индексирования документов в АРБИКОН*

*Когда пользователь обращается в ЭК локальной или корпоративной библиотечной системы через Интернет, ему нужны конкретные данные, соответствующие его информационной потребности. При этом первостепенное значение имеют точность и полнота поиска при минимальных затратах сил и времени на получение и отбор БЗ документов, содержащих нужную информацию.*

*Здесь уместно кратко напомнить некоторые вопросы теории поиска, проверенные долговременной и многосторонней практикой.*

*Для обеспечения высоких показателей работы всех видов информационно-поисковых систем современная лингвистика предлагает две категории средств индексирования документов: предкоординацию и посткоординацию, которые соответственно построены на информационно-поисковых языках классификационного и вербального типа.*

*Предкоординация предполагает деление предметной области на связанные в иерархическом плане составные разделы разного уровня глубины, однако, как правило, не достигающей до частных (или конкретных) наименований объектов, которые их наполняют. Кроме того, она построена на использовании достаточно сложных словарных статей, описывающих содержание конкретных разделов (рубрик) предметной области. Словарными средствами предкоординации являются традиционно используемые в библиотечной и информационной практике рубрикаторы (УДК, ГРНТИ, Дьюи, МКИ и др.).*

*Указанные средства, безусловно, необходимы и прежде всего - для решения ряда внутрибиблиотечных и внутрикорпоративных задач. Однако для конечных пользователей они обладают низкой разрешающей способностью и не обеспечивают высоких показателей поиска. В условиях локальной АБИС пользователи вынуждены затрачивать много времени на анализ документов, взятых "с полки", до которых в условиях удаленного поиска через Интернет/Интранет, в частности по ЭК участников АРБИКОН, добраться сложно, а скорее - просто невозможно.*

*Посткоординация предполагает оперирование конкретными терминами, определяющими предметы содержания описания документов и информационных потребностей пользователей. Она также основана на делении сложных понятий на короткие (более адресный поиск по объектам, персоналиям и т.п.). Существуют два способа ее использования: нормирование лексики (через словарные средства - тезаурусы) и применение ненормированной терминологии - "ключевых слов", т.е. отдельных слов и коротких словосочетаний, которые максимально точно характеризуют искомый объект. Это обеспечивает высокую точность поиска, однако не позволяет добиться приемлемой его полноты из-за синонимии и омонимии терминов в поисковых образах документов и запросов. Для устранения этого недостатка в публичных библиотеках КСМБ, а также в ряде других библиотек, например в Екатеринбургской, Ярославской,*

*Нижегородской КБИС, ОЭБ Томска, применяются принципы "мягкого нормирования" [2, 3].*

*Принципы и средства посткоординации широко используются при индексировании документов и запросов для проведения детального или так называемого тематического поиска в Интернете, автоматизированных системах научно-технической информации и АСУ. В последние 5 - 6 лет они начали применяться и в библиотеках, в том числе участницах КБИС, а также в локальных АБИС.*

*Посткоординация не конфликтует с предкоординацией. Оба принципа должны совместно использоваться, взаимодействовать, поскольку у лингвистических средств, построенных по тому и другому принципу, имеются свои преимущества и недостатки. Это же относится к выполняемым этими средствами функциям, поэтому следует грамотно и добросовестно их использовать. Последнее, в частности, предполагает корректное и полное описание содержания индексируемых документов с использованием рубрикаторов и ключевых слов.*

*Результаты анализа использования средств лингвистического обеспечения в библиотеках КБИС - участницах АРБИКОН показали, что при аналитической обработке документов библиотеки используют сочетание одного-двух рубрикаторов и ключевые слова или только рубрикаторы, либо только ключевые слова. Так, по результатам выполненного в марте 2004 г. тестирования, из 31 московской библиотеки 3 используют только рубрикаторы, 16 - одновременно рубрикаторы и ключевые слова, 12 - только ключевые слова. Есть основание полагать, что подобная ситуация имеет место и в других КБИС, входящих в АРБИКОН.*

*В библиотеках, входящих в Ассоциацию, используются (по отдельности или в сочетании с другими) разные виды рубрикаторов:*

- \* УДК (в большинстве библиотек, исключая значительное число публичных),*
- \* ББК для массовых библиотек (практически во всех публичных библиотеках),*
- \* ББК для научных библиотек (практически во всех вузовских библиотеках),*
- \* ГРНТИ (во всех крупных НТБ, а также значительной части вузовских библиотек),*
- \* Десятичная классификация Дьюи (пока используется только в ГПНТБ России),*
- \* отраслевые рубрикаторы, в частности MESH (в отдельных вузовских, а также в медицинских библиотеках).*

*Само по себе использование разнородных средств предкоординатного индексирования в рамках одной ассоциации не очень привлекательно, но неизбежно. Это определяется как различиями их функционального назначения, так и условиями ведомственной принадлежности библиотек. К тому же, по нашему мнению, с позиции большинства*

*конечных пользователей АРБИКОН отсутствие единого рубрикатора для всех его членов не существенно по следующим причинам:*

*\* основным средством поиска в ЭК через Интернет являются ключевые слова (в том числе выбираемые из элементов БЗ) и фамилии авторов;*

*\* обращение пользователей к ЭК участников КБИС и АРБИКОН с запросами может иметь избирательный характер, учитывающий вероятность совпадения тематики информационных потребностей пользователей и комплектования конкретной библиотеки или группы библиотек в АРБИКОН.*

*В настоящее время по всему массиву БЗ накопленных в ЭК библиотек, действительной проблемой является низкое качество предкоординатного индексирования документов - их систематизация и предметизация.*

*Анализ представительных массивов БЗ, взятых из ЭК разных библиотек, показал, что как правило простановка индексов производится в рамках одной тематической рубрики, что не соответствует многоаспектному содержанию документов. В частности, при проведении количественного определения тематического состава документов в ряде сводных каталогов суммарное их число по всем тематическим рубрикам верхнего уровня всего на 10 - 15% превышает число БЗ в этих ЭК. Анализ содержания большинства представленных этими БЗ документов показывает, что на самом деле оно относится к нескольким (двум, трем и более) тематическим разделам используемого рубрикатора.*

*В указанных условиях, если в БЗ отсутствует поисковый образ документа, отражающий его содержание ключевыми словами, заинтересованный в нем пользователь никогда не выйдет на этот документ, поскольку заглавие или фамилия автора могут быть для него не информативными.*

*При использовании ключевых слов достаточно распространены следующие основные недостатки и ошибки индексирования документов:*

*\* неполное отражение содержания документов в ПОД;*

*\* наличие смысловых ошибок индексирования;*

*\* включение в ПОД ключевых слов, соответствующих малоинформативным аспектам содержания документов (например, когда о каком-либо объекте или факте в документе только упоминается);*

*\* использование случайных, малоупотребительных терминов;*

*\* неправильное построение ключевых слов, в частности использование неоправданно длинных словосочетаний и т.д.*

*Наличие этих ошибок приводит к увеличению уровня шума и/или сокращению полноты*

*результатов поисков, производимых по запросам пользователей. Другими словами, увеличиваются затраты времени и сил пользователей на нахождение документов и/или возможность их потери.*

*В силу всего вышеизложенного многие нужные пользователям документы остаются полностью или частично невостребованными.*

*Очевидно, что для повышения качества индексирования документов с целью обеспечения продекларированного Уставом АРБИКОН "...предоставления более качественного информационного обеспечения пользователей..." необходима разработка и реализация определенного комплекса мер. Понятно также, что переработка уже накопленных ресурсов БЗ в ЭК всех участников АРБИКОН нереальна. Однако оставлять без внимания сложившуюся ситуацию недопустимо, поскольку со временем проблемы с "потребительской" ценностью информационных ресурсов АРБИКОН будут только расти.*

*Что требуется сделать?*

*По нашему мнению, чтобы повысить качество информационных ресурсов АРБИКОН, необходимо прежде всего улучшить качество библиографического и аналитического описания первичных документов в ЭК всех участников Ассоциации. Для этого должен быть разработан, утвержден и внедрен в практику "Паспорт качества библиографического и аналитического описания документов в ЭК участников АРБИКОН".*

*Для создания "Паспорта качества..." требуется:*

*\* разработать и утвердить на уровне Совета АРБИКОН общую для всех его участников инструктивно-методическую документацию по аналитической обработке документов;*

*\* сформировать систему контроля качества ЭК участников АРБИКОН;*

*\* подготовить и реализовать систему стимулов, способствующих скорейшему и эффективному внедрению в библиотеках "Паспорта качества..." после его принятия (в частности, такими стимулами могут служить публикации результатов контроля качества ЭК библиотек-участниц на сайтах АРБИКОН и входящих в его состав консорциумов);*

*\* в региональных консорциумах организовать постоянно действующие курсы повышения квалификации с согласованными программами обучения, ориентированными на задачи повышения качества электронных ресурсов членов АРБИКОН.*

*Функции научно-методического обеспечения внедрения указанных требований в консорциумах, а также контроля за их выполнением должен взять на себя НТС 2 - "Информационные ресурсы и обслуживание".*

*Поскольку повышение качества аналитической и библиографической обработки текущего потока документов и ретроспективы связано с увеличением затрат времени персонала подразделений комплектования и обработки библиотек, следует облегчить условия заимствования готовых БЗ из ЭК участников АРБИКОН. В частности, предложить НТС-3 "Программно-технические средства и технологии" разработать схему размещения распределенных и сводных ЭК АРБИКОН в сети для обеспечения возможности ускоренного доступа к ним через ограниченное число региональных зеркальных серверов (в настоящее время доступ к ним и поиск проходят очень медленно).*

*Должна поощряться наибольшая активность деятельности отдельных участников АРБИКОН в качественном создании информационных ресурсов общего пользования, в разработке программно-технических комплексов и их предоставлении участникам Ассоциации на льготных условиях, в проведении консультаций, стажировок и других видов активной профессиональной помощи, в частности путем снижения размеров годовых членских взносов для таких организаций по решению Совета или Общего собрания АРБИКОН.*

#### *Список источников*

- 1. Устав Некоммерческого Партнерства "Ассоциация региональных библиотечных консорциумов" <http://www.arbicon.ru/docs/>.*
- 2. Воройский Ф. С. Индексирование документов в АБИС // Библиотека. - 1996. - N 9. - С. 42 - 44.*
- 3. Создание Корпоративной сети публичных библиотек Москвы: Рабочий проект. Т.1 / Науч. руководитель д-р техн. наук Я. Л. Шрайберг, отв. исполн. канд. техн. наук, проф. Ф. С. Воройский. - М.: ГПНТБ России, 2001. - 85 с. На правах рукописи <http://corporate.gpntb.ru/documents.php>.*