

**Система автоматизации библиотек ИРБИС:  
расширение возможностей полнотекстовых баз данных,  
интеграция имидж-каталога и другие новые решения**  
**ALIS IRBIS: Extension of Opportunities for Full-text Databases,  
Integration of Image Catalog and other New Solutions**

**Система автоматизації бібліотек ІРБІС:  
розширення можливостей повнотекстових баз даних,  
інтеграція імідж-каталогів та інші новітні вирішення**

*А. И. Бродовский*

*Государственная публичная научно-техническая библиотека России, Москва, Россия*

*Alexander Brodovsky*

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

*О. Й. Бродовський*

*Державна публічна науково-технічна бібліотека Росії, Москва, Росія*

В докладе дается общая характеристика системы автоматизации библиотек ИРБИС и обзор новых решений, характеризующих нынешний этап ее развития.

The paper offers the general description of ALIS IRBIS and a review of new solutions at the present stage.

У доповіді представлено загальну характеристику системи автоматизації бібліотек ІРБІС та огляд нових рішень, які характеризують нинішній етап її розвитку.

Система автоматизации библиотек ИРБИС как типовое комплексное решение для автоматизации библиотечных технологий – являясь фактически ровесницей Крымской конференции – существует уже почти 15 лет. За это время система прошла через четыре поколения развития, превратившись из программы с минимальными функциями, рассчитанную на операционную систему MS-DOS, в большой программный комплекс, охватывающий практически все библиотечные технологии, удовлетворяющий мировым и отечественным стандартам, функционирующий в больших корпоративных и глобальных сетях. За этот период пользователями системы стало более тысячи библиотек из России и стран СНГ. Среди них библиотеки разного масштаба и профиля, от сельских до национальных, от школьных до вузовских, от отраслевых до массовых. Среди пользователей системы музеи, архивы, другие предприятия и организации, занимающиеся автоматизацией информационных технологий.

Разумеется, в коротком выступлении нельзя представить систему ИРБИС в полном объеме. Я остановлюсь лишь на основных характеристиках системы, дающих о ней общее представление. Расскажу о составе системы, т. е. тех продуктах, которые входят в семейство ИРБИС, а также дам короткий обзор новых решений, появившихся в последнее время.

Говоря об основных характеристиках, прежде всего следует сказать о том, что система полностью интегрируется в современные корпоративные системы и технологии на основе, во-первых, полной совместимости с Российским коммуникативным форматом RUSMARC и международными библиографическими форматами UNIMARC и MARC21, и во-вторых, на основе поддержки Интернет-технологий, а точнее, Web-технологии и протокола Z39.50. Именно благодаря этим характеристикам пользователи системы имеют возможность представлять свои ресурсы в Интернет, включать их в распределенные корпоративные ресурсы, осуществлять непосредственное онлайн-вое заимствование готовых библиографических описаний из электронных каталогов других библиотек и корпоративных систем.

Система предлагает все средства для автоматизации традиционных библиотечных технологий, направленных на создание и ведение электронного каталога, и в то же время предлагает средства

для новых технологий, отвечающих требованиям уже электронной библиотеки, а именно – средства для создания и ведения полнотекстовых баз данных.

Следуя требованиям библиотек, которые находятся на начальных этапах автоматизации, система ИРБИС скрупулезно поддерживает «бумажные» технологии – от печати листов книги суммарного учета до всех видов каталожных карточек, и в то же время система предлагает средства для новых, сугубо электронных технологий, основанных на штрих-кодах и радиометках, в частности с использованием этих средств предлагается реализовывать современную технологию книговыдачи.

Система позволяет создавать и поддерживать любое количество баз данных, составляющих Электронный каталог, представляющих собой проблемно-ориентированные библиографические или полнотекстовые базы данных.

Система предлагает технологию автоматического формирования словарей, на основе которых реализуется быстрый поиск по любым элементам описания и их сочетаниям.

Система предлагает средства для ведения и использования таких информационно-лингвистических ресурсов как базы данных основных классификаторов (УДК/ББК), Рубрикатор ГРНТИ, Тезаурус, Авторитетные файлы Индивидуальных авторов, Коллективов, Издательств и Предметных заголовков.

Система включает средства, которые позволяют использовать в качестве иллюстративного материала любые внешние по отношению к библиографическому документу объекты, такие как полные тексты, графика, таблицы, аудио- и видеоматериалы, а также ресурсы Интернет.

Система поддерживает многоязычие на основе UNICODE, т. е. обеспечивается ввод данных на любых языках.

Система предлагает большой набор сервисных средств, обеспечивающих удобство и наглядность пользовательских интерфейсов, упрощающих процесс ввода, исключая ошибки и дублирование информации. В частности – в системе работает орфографический контроль.

Система располагает широкими возможностями для адаптации к условиям работы конкретной библиотеки, включая средства создания уникальных рабочих профилей для всех категорий пользователей.

Система является открытой – что позволяет пользователю самостоятельно вносить изменения в широких пределах: от изменения входных и выходных форм до включения собственных оригинальных приложений.

В настоящее время система ИРБИС представлена несколькими самостоятельными продуктами:

Прежде всего это два поколения собственно систем автоматизации библиотек – ИРБИС32 и ИРБИС64. С точки зрения функциональности эти две системы в основном совпадают. Основное отличие между ними – внутренняя архитектура. ИРБИС32 – это файл-серверная система, а ИРБИС64 – клиент-серверная. С этим связан разный масштаб библиотек, на которые они ориентированы. ИРБИС32 в большей степени ориентирована на малые и средние библиотеки, а ИРБИС64 – на средние и крупные. (Под масштабом библиотеки здесь понимается объем электронного каталога и количество пользователей, одновременно с ним работающих).

Третьим продуктом семейства ИРБИС является ИРБИС64 ПОЛНОТЕКСТОВЫЕ БД (ИРБИС64 ПБД) – система, предназначенная для создания и использования полнотекстовых баз данных.

Также к продуктам ИРБИС можно отнести выпускаемые ГПНТБ России базы данных на компакт-дисках – это БД УДК, ББК, ГРНТИ и Российский сводный каталог по научно-технической литературе.

Теперь о новых решениях, которые появились в системе в последнее время.

Важнейшими новыми решениями в системе ИРБИС можно считать новые режимы в АРМЕ (автоматизированное рабочее место) Каталогизатор для реализации **технологии заимствования**. Предлагается три режима:

**ИМПОРТ ИЗ ЛИБНЕТ** – обеспечивает непосредственное заимствование документов из ресурсов ЛИБНЕТ в соответствии с авторизацией пользователя (включая поиск и просмотр в БД).

**ИМПОРТ ИЗ Web-ресурсов ИРБИС** – дает возможность вести заимствование из ресурсов, доступных на основе Web-ИРБИС, т. е. любая группа пользователей системы ИРБИС – будь то две библиотеки, библиотеки региона или вообще все пользователи ИРБИС – может организовать

корпоративную работу между собой, т. е. вести непосредственное (онлайновое) заимствование документов друг у друга. Используя эту технологию, ГПНТБ России, как пользователь системы ИРБИС, предоставляет всем остальным пользователям ИРБИС возможность вести свободное, т. е. бесплатное, заимствование из всех своих ресурсов, выставленных в Интернет, – в общей сложности это более миллиона документов.

**ИМПОРТ ИЗ Z-ресурсов** – дает возможность вести заимствование из любых мировых библиографических ресурсов, доступных по протоколу Z39.50 – а это электронные каталоги всех крупнейших библиотек России: РГБ, ГПНТБ, ВГБИЛ, БЕН и другие. И все – за редким исключением – бесплатно.

В последнюю версию системы (2007.2) мы включили режимы, которые позволяют выполнять действия обратные заимствованию – а именно, передачу документов в корпоративные ресурсы по протоколу Z39.50 и через WebИРБИС. Т. е. непосредственно из ARMa Каталогизатор ИРБИС можно передавать созданные документы в корпоративные (сводные) ресурсы, доступные по Z39.50 или по Web-технологии ИРБИС.

Очевидно, что все эти режимы заимствования, позволяют существенно повысить эффективность работы каталогизаторов.

**Также очень важными и принципиальными мы считаем новые решения, касающиеся полнотекстовых БД, т. е. ИРБИС64 ПОЛНОТЕКСТОВЫЕ БД:**

- расширение номенклатуры используемых исходных текстов – дополнительно к текстовым файлам типа DOC, PDF, RTF, TXT, HTML, PPT, XLT добавлены файлы типа DJVU;
- обеспечение возможности включения в полнотекстовые БД – наряду с собственно текстовыми документами – графических образов текстов на основе OCR-технологии, т. е. путем распознавания текстов «на лету» на основе технологии фирмы АBBYY;
- применение для поиска в полнотекстовых БД вербальных классификационных систем или точнее – тематических навигаторов;
- автоматическое разбиение многостраничных текстов – в результате каждая страница виртуально функционирует в базе данных как отдельный документ.

На основе ИРБИС64 ПБД с возможностью автоматического распознавания текстов создан новый продукт ИРБИС – **ИРБИС ИМИДЖ-КАТАЛОГ**.

Имидж-каталоги – одно из наиболее эффективных средств ретроконверсии карточных каталогов для библиотек с большим ретрофондом (т. е. тем фондом, который не отражен в полноценном электронном каталоге)

Решения по созданию имидж-каталогов, которые предлагаются в настоящее время большинством фирм, не включают автоматического распознавания текстов карточек и строятся на основе:

- индексирования карточных разделителей;
- строгого сохранения порядка следования отсканированных образов карточек, т. е. соответствующие файлы должны размещаться в строгом соответствии с порядком установки карточек в традиционном каталоге.

В таких имидж-каталогах **ТЕМАТИЧЕСКИЙ** поиск (со стороны конечного пользователя – читателя) практически невозможен. Они приемлемы лишь для профессионалов, выполняющих **БИБЛИОГРАФИЧЕСКИЙ** поиск.

Решение, которое мы предлагаем, строится прежде всего на основе автоматического («на лету») распознавания текстов каталожных карточек, за счет чего:

Во-первых, реализуется полнотекстовый поиск (с ранжированием результатов, с учетом близости слов, морфологии и др.), который, включая возможности библиографического поиска для профессионалов, предлагает конечному пользователю возможности полноценного тематического поиска;

Во-вторых, не предъявляются жесткие требования к порядку сканирования и размещения образов карточек. В случае же соблюдения этого порядка полностью реализуется модель «бумажного» поиска по ящикам и разделителям.

Но, пожалуй, главной отличительной чертой нашего решения **ИМИДЖ-КАТАЛОГА** является то, что база данных **ИМИДЖ-КАТАЛОГА** функционирует **НЕ** отдельно, **НЕ** как автономная система, а интегрируется в автоматизированные технологии полноценного электронного каталога,

а именно – в технологии АРМов Каталогизатор и Книговыдача ИРБИС64, что позволяет осуществлять ввод структурированных данных (например, производить роспись экземпляров ретрофонда), корректировку распознанных данных, удаление записей, списание и проверку фонда, бронирование, выдачу, возврат, продление и пр.

Решение включает три модуля:

- Администратор – для создания собственно базы данных ИМИДЖ-КАТАЛОГА на основе сканированных образов каталожных карточек (пакетная операция), именно в него встраивается «движок» АВВУУ;
- оригинальный клиент для конечного пользователя ИМИДЖ-КАТАЛОГА;
- Web-шлюз для доступа к БД ИМИДЖ-КАТАЛОГА.

И дополнительно могут использоваться АРМы Каталогизатор, Комплектатор и Книговыдача ИРБИС64.

Важные функциональные возможности и технологии созданы в АРМе **Книговыдача**.

Это прежде всего технология **книговыдачи на основе RFID-технологии**, в том числе с использованием противокражной подписи и группового считывания радиометок. (в качестве примера полной реализации технологии книговыдачи на основе RFID-меток можно указать библиотеку факультета менеджмента СПбГУ, где также реализована оригинальная технология самовыдачи).

Также создана новая технология, связанная с предварительным заказом и бронированием изданий, т. н. технология **БРОНЕПОЛКИ**, в которой реализованы четыре новые операции: собственно бронирование, т. е. перевод экземпляров на бронеполку, выдача с бронеполки, возврат на бронеполку и возврат с бронеполки.

Для случаев сложной топографии книговыдачи, когда существует несколько мест (кафедр) выдачи и мест хранения, имеющих неоднозначную связь, разработаны технологии контроля соответствия мест выдачи и мест хранения (например, чтобы книга возвращалась там же, где она выдавалась, чтобы препятствовать выдаче экземпляра из хранения, не связанного с данным местом выдачи и т. п.)

Также предлагаются дополнительные решения, связанные с книговыдачей **многоэкземплярной** литературой, что особенно критично для вузовских библиотек.

Для анализа **статистики книговыдачи** пользователю предлагаются разнообразные новые статистические формы, в которых дается распределение выдач, возвратов, посещений, должников по датам, местам выдачи, категориям читателей и т. п. При этом пользователю предлагается аппарат формирования собственных статистических форм, отвечающих его специфическим требованиям.

В очередной версии системы в АРМе Книговыдача появится новая подсистема, реализующая технологию **платных услуг**. Данная технология предлагает возможность ведения для каждого читателя (получателя услуг) истории предоставленных ему УСЛУГ и полученных от него ВЗНОСОВ на основе фиксированного перечня услуг библиотеки, включающего их описание и расценки. В качестве УСЛУГ учитываются как собственно выполненные услуги, так и заказы на их выполнение, а под ВЗНОСАМИ понимаются как непосредственная плата за выполненные услуги, так и произвольные взносы на личный счет. При учете услуг и взносов конкретного читателя фиксируются такие данные, как вид услуги, сумма взноса, дата/время, место обслуживания, ответственное лицо и др. Автоматически в режиме реального времени фиксируются (вычисляются «на лету»): объем выполненных услуг, объем взносов и остаток на счете текущего читателя, а также – общий объем выполненных услуг по установленному месту обслуживания (выдачи) на текущий день (а именно – общее количество выполненных услуг, количество обслуженных читателей, общая сумма выполненных услуг). Предлагается выходная форма исчерпывающего статистического отчета, в котором отражается распределение выполненных услуг – по их видам – по дням заданного месяца с подсчетом итоговых значений: количество услуг, количество обслуженных людей, сумма.

Весьма интенсивно развивается АРМ **Книгообеспеченность**, востребованный вузовскими библиотеками (особенно при их лицензировании). Совершенствование его ведется в двух направлениях: это с одной стороны – создание новых сервисов для ведения связанных баз данных электронного каталога, читателей и учебных дисциплин и с другой – создание новых выходных форм. Хотя второе направление можно считать попыткой объять необъятное: сколько вузовских библиотек,

столько разнообразных требований по выходным формам, т. е. помимо типовых форм Минобразования пользователи хотят получать большое количество специфичных только для них форм.

Говоря о новых решениях в части информационно-технологического обеспечения АРМов Комплектатор и Каталогизатор, следует отметить следующие новшества:

- с целью упрощения и повышения наглядности существенно доработаны технологии **списания литературы и проверки фонда**, в частности, усовершенствован аппарат пересчета цен и создан специальный «мастер» списания;
- расширена технология **оформления подшивок**, в частности, обеспечена возможность включения в одну подшивку выпуски разных изданий;
- разработан **алгоритм автоматического расчета тиража каталожных карточек** (основных, добавочных, систематических и оборота); при этом предлагаются средства настройки для уточнения алгоритма в соответствии с требованиями конкретного пользователя;
- разработана технология заказа книг с использованием **прайс-листов издательств** (получаемых в формате Excel) с подготовкой специальной выходной формы **«Реестр на закупку товаров»** для объявления открытых аукционов, тендеров, конкурсов и котировок (ФЗ № 94);
- разработана специальная база данных, а также соответствующие входные и выходные формы для описания и ведения документов **технического архива предприятия** с возможностью фиксирования древовидных (иерархических) связей между ними (таких как «имеет приложение...», «входит в состав...» и пр.).

Важной особенностью нынешнего этапа развития системы ИРБИС является подключение к процессу разработок самих пользователей. И речь идет не о каких-то частных решениях или тем более о создании новых входных и выходных формах представления данных – что является вполне рядовым явлением, – а о законченных функциональных разработках, имеющих типовой характер. Так, на основе творчества наших пользователей сейчас создается полнофункциональная система доступа к базам данных на РНР (разработчик Михайленко Илья, Омский государственный университет), что открывает возможность к многоплатформенности ИРБИС, т. е. к использованию вместо платформы Windows, которая сейчас является единственно возможной для ИРБИС, например, Linux или Unix. Также разработкой нашего пользователя является новый модуль системы ИРБИС т. н. J-ИРБИС, представляющий собой интеграцию WebИРБИС в систему управления контентом (CMS) Joomla и являющийся мощным инструментом для создания сайта вузовской библиотеки (разработчик Соколинский Кирилл, Северо-западный заочный технический университет, Санкт-Петербург).