



Седьмая Международная Конференция «Крым 2000»

Конференция проводится в рамках мероприятий ИФЛА 2000 г.

***Библиотеки и ассоциации в меняющемся мире:
новые технологии и новые формы сотрудничества***

***Тема 2000 года:
Библиотеки, издательства,
книгораспространение и образование
в едином информационном и
социокультурном пространстве***

***Труды конференции
Том 2***

СУДАК
(основная программа)
Ялта, Алушта, Феодосия, Симферополь,
Коктебель, Старый Крым
(выездные заседания)
Автономная Республика Крым, Украина
3-11 июня 2000

Текущее состояние Web-ориентированной системы управления библиографическими базами данных CDS/ISIS СО РАН

Web-oriented Bibliographic Database Management System: State of the Art

Поточний стан Web-орієнтованої системи управління бібліографічними базами даних CDS/ISIS СВ РАН

*Малицкий Н.А., Баженов И.С., Мазов Н.А., Баженов С.Р.
Объединенный Институт Геологии, Геофизики и Минералогии СО РАН,
Государственная Публичная Научно-техническая Библиотека СО РАН, Новосибирск, Россия*

*Nikolai A. Malitsky, I. S. Bazhenov, Nikolai A. Mazov, Sergei R. Bazhenov
Joint Institute of Geology, Geophysics and Mineralogy of the Russian Academy of Sciences' Siberian Division;
Russian Academy of Sciences' State Public Library for Science and Technology, Novosibirsk, Russia*

*Малицький М.О., Баженов І.С., Мазов М.О., Баженов С.Р.
Об'єднаний Інститут Геології, Геофізики і Мінералогії СВ РАН,
Державна Публічна Науково-технічна Бібліотека СВ РАН, Новосибірськ, Росія*

Рассматривается текущее состояние развития Web-ориентированной системы управления библиографическими базами данных СО РАН. Обсуждаются условия обеспечения полной функциональности систем подобного класса.

Considered is state of the art of the Web-oriented bibliographic database administration system developed by the Russian Academy of Sciences' Siberian Division. Discussed are conditions necessary for perfect operation of such systems.

Розглядається поточний стан розвитку Web-орієнтованої системи управління бібліографічними базами даних СВ РАН. Обговорюються умови забезпечення повної функціональності систем подібного класу.

На сегодняшний момент времени существует множество вариантов систем on-line доступа к информационно-библиотечным ресурсам в сети Internet. Все они обладают более или менее развитыми средствами поиска. Созданная в ГПНТБ СО РАН и ИБЦ ОИГТИМ СО РАН система доступа к библиографическим базам данных показала в эксплуатации свою эффективность [1,2,3]. Она обладает развитыми средствами поиска, предлагает в зависимости от степени подготовленности пользователя несколько режимов составления поискового запроса: ПРОСТОЙ, СТАНДАРТ, ЭКСПЕРТ, МУЛЬТИ. Для проведения одновременного поиска в нескольких однотипных по поисковым терминам базах данных используется режим мультипоиска. При составлении поискового запроса можно воспользоваться средством работы со словарем базы данных. В процессе просмотра словаря требуемые поисковые термины могут быть выбраны для включения в текущий поисковый запрос, причем для режимов ЭКСПЕРТ и МУЛЬТИ возможен групповой выбор сразу нескольких близлежащих поисковых терминов. Формой выдачи результатов поиска можно управлять через установку таких параметров как формат и порция выдачи записей. Если количество найденных записей превышает порцию выдачи, то пользователь имеет возможность последовательно или произвольно перемещаться в их диапазоне и изменять формат выдачи. При просмотре результатов поиска возможно проведение уточняющего поиска по некоторым, определенным для этого, полям с возможностью отката к основному поиску. К примеру, этими полями могут быть такие поля как «Автор», «Рубрика НТИ» и т.п. Найденные в результате поиска записи могут быть выборочно или все полностью сохранены пользователем в текстовый файл на его локальном компьютере. Также в рамках данной системы в ГПНТБ СО РАН реализована работа с полнотекстовыми базами данных. После проведения поиска в этих базах полные тексты, хранящиеся в формате PDF, могут быть загружены пользователем на его собственный компьютер.

Важной составной частью системы доступа является ведение базы данных статистики обращений пользователей к выставленным информационным ресурсам. Это позволяет оперативно выявлять степень использования предлагаемых сервисных возможностей системы, определять степень актуальности выставленных информационных ресурсов и, в конечном итоге, способствовать выработке дальнейшего направления развития функциональных возможностей данной системы.

Система доступа имеет своей целью обеспечить оперативный доступ к выставленным информационно-библиотечным ресурсам из любой точки мира. Поэтому важным моментом является наличие возможности работы с ней любому пользователю, как знающему русский язык, так и не знакомого с ним. Система ориентирована на работу с основными, используемыми в Internet, кириллическими кодировками (Windows-1251, KOI-8, ISO-8859-5, IBM-866/DOS, Mac). Это позволяет отечественному пользователю работать с системой практически с любого типа компьютеров и в любой из существующих операционных систем. Также система имеет англоязычный вариант интерфейса. Одной из актуальных потребностей стала необходимость обеспечить возможность работы с отечественными библиографическими базами данных для иностранных пользователей системы. Суть проблемы заключается в отсутствии установленных кириллических кодировок на компьютере иностранного пользователя, что не позволяет правильно формировать поисковые запросы и отображать результаты поиска в русскоязычных базах данных. Говорить о создании, по крайней мере, английского варианта для каждой базы данных лишено смысла в рамках существующей технологии обработки документов и отсутствия требуемых для этого ресурсов. Единственным решением в данной ситуации является использование транслитерации русских символов. Причем существовало два варианта реализации транслитерации: либо выставлять для каждой базы данных ее транслитерированный вариант, либо проводить транслитерацию «на лету», используя единственный, оригинальный экземпляр базы данных. В первом случае пришлось бы сопровождать два варианта базы данных (две ее копии), что потребовало еще более значительных технических ресурсов (дискового пространства для хранения, вычислительных мощностей для проведения транслитерации базы данных при каждом ее обновлении). Во втором случае дублирование исключается и транслитерация происходит «на лету» - когда происходит отработка поискового запроса и выдача результатов поиска. Для однозначного соответствия конвертирования символов была выбран немецкий вариант транслитерации русских символов. Ознакомиться с правилами транслитерации пользователь может, воспользовавшись on-line подсказкой, где приведены таблица соответствия русских и латинских символов, а также простые примеры. Транслитерация реализована как в клиентской части системы доступа при работе со словарем базы данных, формировании и выполнении поискового запроса, так и в серверной части при выдаче результатов поиска. Пользователю при составлении запросов следует явно указывать, что вводимые им отдельные символы или целые слова являются транслитерированной формой соответствующих кириллических символов. Это связано с тем, что при выполнении поискового запроса системе необходимо провести обратное конвертирование транслитерированных символов в соответствующие кириллические символы. Поэтому для однозначного определения являются ли текущие, введенные пользователем, символы или слова транслитерированными их следует заключать в фигурные скобки. Например, поисковый термин «ВОДА» в транслитерированной форме имеет вид «VODA». Чтобы система при обратном конвертировании знала, что слово является транслитерацией слова «ВОДА» и снова получила этот термин в оригинальной форме пользователю следует ввести его как «{VODA}». Если пользователь введет термин, не заключая его в фигурные скобки, то система при выполнении поискового выражения оставит его без изменения, считая его обычным иностранным словом. Важно отметить то, что транслитерация является составной частью пользовательского интерфейса системы (его англоязычного варианта) и не привязана к конкретной базе данных. Это позволяет легко и гибко в дальнейшем вносить изменения в программную логику текущей реализации транслитерации. Использование в рассматриваемой системе динамической транслитерации дало возможность, не нарушая традиционной технологии обработки баз данных в ГИПТБ СО РАН и ИБЦ ОИГГиМ СО РАН, предоставить доступ к выставленным русскоязычным библиографическим базам данных для любого пользователя Интернет.

Наличие широкого спектра выставленных информационно-библиотечных ресурсов потребовало создания в рамках данной системы подсистемы регистрации и управления правами доступа пользователей к этим ресурсам. Ввиду того, что используемый в системе поиска модуль WWWISIS [4] обладает ограниченными возможностями для проведения изменений в базах данных CDS/ISIS по сравнению с динамической библиотекой ISIS32.DLL [5] было принято решение выбрать эту библиотеку в качестве основного механизма при создании инструментария системного администрирования пользователей системы и выставленных информационно-библиотечных ресурсов ГИПТБ СО РАН и ИБЦ ОИГГиМ СО РАН. Созданная подсистема администрирования была написана в средах Borland Delphi и C++ Builder с использованием ISIS32.DLL. Однако в дальнейшем выяснилось, что данная библиотека некорректно работает с поисковыми выражениями: происходит потеря номера каждой 63 найденной записи, зависание при работе с большими базами данных или при нахождении большого количества записей и т.д. В связи с этим была разработана альтернативная динамическая библиотека, в которой были корректно реализованы все основные функции работы с поисковыми выражениями. При разработке библиотеки был использован модуль WWWISIS. Помимо этого в ней, в отличие от ISIS32.DLL, была реализована стандартная для CDS/ISIS возможность использования пользовательских форматных выходов (подпрограмм) в форматах вывода и таблицах индексации словаря поисковых терминов. Это особенно важно ввиду большого количества созданных ранее пользовательских форматных выходов в рамках существующей в ГИПТБ СО РАН и ИБЦ ОИГГиМ СО РАН технологии обработки информации. Разработанная библиотека была применена в подсистеме администрирования

пользователей и выставленных информационно-библиотечных ресурсов, при разработке режима работы с выпусками журналов и т.д. В частности, при работе с последними удалось легко реализовать удобную навигацию по выпускам журналов посредством выбора года или порции выпусков, которые выдаются по 15 выпусков. С использованием данной библиотеки разработана система формирования сводного каталога, позволяющая при работе с базами данных книг ввести новое описание издания, при работе с периодическими изданиями – новое описание журнала, и новый выпуск журнала, при этом автоматически заносится сигла текущего пользователя, которая заранее прописывается в базе данных пользователей при регистрации.

С развитием технологий построения больших информационных систем, включающих в себя множество различных баз данных, достаточно актуальным становится вопрос поиска информации в БД с использованием тезаурусов и классификационных схем. Поэтому значительным шагом в развитии стало введение в эксплуатацию режима работы с тезаурусами и рубрикаторами в рамках данной системы, использование которых позволяет пользователю систематизировать проводимый поиск и сделать его более эффективным. Причем существенным плюсом стал тот факт, что этот режим был реализован в общей идеологии обеспечения доступа к выставленным ресурсам, т.е. тезаурусы и рубрикаторы представляют собой обычные базы данных с возможностью поиска в них, обеспечивая при этом дополнительную возможность произвольного перемещения по дереву уровней.

Говорить о создании Web-ориентированной системы управления библиографическими базами данных можно в том случае, когда она помимо наличия развитого поискового инструментария будет включать в себя механизм ведения данных. Это предполагает добавление, изменения и удаление записей. Поэтому следующим шагом в развитии системы доступа стала разработка программного модуля on-line ведения библиографических баз данных. Закономерно, что в качестве инструмента реализации ввода данных был выбран обычный Web-навигатор. Это связано с тем, что, во-первых, Web-навигатор, имеет стандартный, унифицированный пользовательский интерфейс, состоящий из ограниченного набора элементов, что позволяет пользователю быстро освоиться с интерфейсом; во-вторых, отпадает необходимость сопровождения клиентской части системы при внесении изменений в программный код, который загружается Web-навигатор с сервера; в-третьих, не имеет значения, где расположен компьютер пользователя – он может находиться как во внутренней Интранет сети организации, так и в любом другом месте глобальной сети Интернет. Обеспечение возможности изменения данных в on-line режиме особенно остро ставит вопрос, связанный с правами доступа пользователей на редактирование содержимого выставленных баз данных. Поэтому при реализации системы on-line ведения была решена задача аутентификации пользователей с предоставлением им возможности, в соответствии с их правами, производить редактирование в выставленных базах данных. Созданная система on-line ведения представляет собой удобный для пользователя инструмент ввода, который обладает максимально возможной функциональностью и, в то же время, легок для понимания. Интеграция данной системы ведения с уже существующей информационно-поисковой системой позволяет говорить о появлении полнофункциональной Web ориентированной системы управления библиографическими базами данных. Это является законченным решением для ведения библиографических баз данных в Интернет характерными чертами которого являются:

- прозрачный в понимании, дружелюбный пользовательский интерфейс;
- богатый спектр возможностей при формировании поискового запроса и выдачи результатов поиска;
- легкость в администрировании выставленных информационных ресурсов и пользователей системы;
- широкий диапазон решаемых информационно-библиотечных задач (заказ литературы по МБА, формирование сводного каталога книг и журналов, работа с полными текстами, сохранение результатов поиска);
- высокая скорость поиска в больших базах данных;
- возможность проводить удаленное on-line редактирование в доступных базах данных;
- заложенный потенциал для дальнейшего развития.

В заключении следует отметить, что описанная система находится в эксплуатации более двух лет и за это время показала свою жизнеспособность. Тем не менее, она постоянно находится в развитии, что естественно для системы такого класса. Более подробно ознакомиться с работой системы можно по следующим адресам (<http://www.rsl.psc.ru> – ГПНТБ СО РАН, <http://geolibr.nisgml.nsc.ru> - ОИГТИМ СО РАН).

Литература

1. Баженов С.Р., Мазов Н.А., Малицкий Н.А., Баженов И.С. Создание программного комплекса доступа из Интернет к базам данных на основе WWW-ISIS // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы конф. "Крым 98". - М., 1998. - Т. 2. - С. 581 - 584.
2. Мазов Н.А., Малицкий Н.А., Баженов С.Р., Баженов И.С., Жижимов О.Л. WEB-ориентированная информационно-поисковая система для доступа к базам данных CDS/ISIS // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы 6 Междунар. конф. "Крым 99". - Судак, 1999. - Т. 1. - С. 113 - 115.
3. Баженов С.Р., Мазов Н.А., Малицкий Н.А., Баженов И.С. Создание программного комплекса доступа из Интернет к базам данных на основе WWW-ISIS// Научные и технические библиотеки. - 1999.- N2.- С. 47 - 52.

4. BIREME/PAHO/WHO. WWWISIS: a word-wide web server for ISIS-databases. Ver. 3.0. Sao Paulo, Brazilian, Oct. 1997.
5. BIREME/PAHO/WHO. ISIS Application Program Interface: ISIS_DLL User's Manual. Sao Paulo, Brazilian, Aug. 1997.