



**5-я Юбилейная
Международная Конференция
"Крым 98"**

Конференция проводится в рамках мероприятий ИФЛА 1998 г.

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

***Материалы конференции
Том 2***

Судак, Автономная Республика Крым, Украина
6-14 июня 1998

Создание программного комплекса доступа из Интернет к базам данных на основе WWW-ISIS

Development of WWW-ISIS-based Software for Accessing Databases via Internet

Створення програмного комплексу доступу з Інтернет до баз даних на основі WWW-ISIS

*Баженов С.Р., Мазов Н.А., Малицкий Н.А., Баженов И.С.
ГПНТБ СО РАН, Новосибирск, Россия*

*Bazhenov S.R., Mazov N.A., Malitsky N.A., Bazhenov I.S.
SPLS&T SB RAS, Novosibirsk, Russia*

*Баженов С.Р., Мазов М.А., Малицький Н.А., Баженов І.С.
ДПНТБ СВ РАН, Новосибірськ, Росія*

В докладе описывается создание программной системы, осуществляющей доступ средствами Интернет к библиографическим базам данных, поддерживаемых информационно-поисковой системой CDS/ISIS. В качестве основы при функционировании системы используется программа WWW-ISIS, разработанная в BIREME, Бразилия.

Described is the software system providing Internet access to bibliographic databases supported by CDS/ISIS information retrieval system. WWW-ISIS program developed at BIREME, Brazil, is used as a basis for the system operation.

В доповіді описано створення програмної системи, що забезпечує доступ засобами Інтернет до бібліографічних баз даних, які підтримуються інформаційно-пошуковою системою CDS/ISIS. В якості основи при функціонуванні системи використовується програма WWW-ISIS, що розроблена в BIREME, Бразилія.

Описываемая система предполагает реализацию полного спектра функций, необходимых при работе с библиографическими базами данных. Она обеспечивает поиск как по простым запросам, так и запросам любой сложности с применением логических операторов и скобок. При формулировании запроса имеется возможность использования словарей баз данных для любого индексированного поля. В описываемой системе реализовано ведение базы данных CDS/ISIS (запись и обновление информации) в среде Интернет, что позволяет решать практически весь комплекс информационно-библиотечных задач в среде Интернет.

На предыдущих конференциях авторами был представлен ряд докладов, в которых приводилось описание системы доступа к базам данных CDS/ISIS/M, реализованных на основе серверов баз данных CDS/ISIS собственной разработки для платформ Unix и Windows-NT [1,2]. Эти сервера позволяли решать задачи поиска и отображения найденной информации, однако в них не были реализованы функции ведения баз данных, что ограничивало их применение.

В 1997 году в свободном доступе появился WWW-сервер для баз данных, разработанный в BIREME совместно с UNESCO, который реализует полный спектр функций работы с базами данных. Он и был принят в качестве базового средства доступа к базам данных CDS/ISIS после тщательного и всестороннего тестирования.

Клиентская часть системы реализована стандартными средствами WWW, а именно с применением языков HTML и JavaScript. Для обработки данных, получаемых со стороны пользователя используются CGI-скрипты (CGI-scripts), реализованные на языке Си. Ввиду того, что в качестве базовой операционной системы сервера используется Windows NT, необходимо, чтобы используемые модули работали как приложения Win32.

Анализ существующих систем

При проектировании настоящей системы поиска в библиографических базах данных был проведен анализ аналогичных систем, используемых в иностранных и отечественных библиотеках. При этом преследовались следующие цели:

- выявить характерные черты интерфейсов, используемых в этих системах;
- оценить функциональные возможности этих систем;
- используя данные анализа, определить требования к разрабатываемой системе.

В ходе анализа были рассмотрены, как наиболее типовые, некоторые системы поиска в библиографических базах данных, в частности, зарубежные - OPAC-97, библиотеки Конгресса США, системы SUBITO Германии и отечественные - ГПНТБ России, БЕН, Интегрум-Техно и другие.

Отметим, что о достоинствах и недостатках рассмотренных систем можно говорить исходя из трех точек зрения:

- удобство работы с системой для пользователя;
- поддержка системных возможностей пользователя;
- принципы построения и функционирования системы, возможности ее дальнейшего развития как программного продукта.

К недостаткам, отмеченным нами, можно отнести следующие:

- отсутствие справочной информации о базе данных, о функциях, реализуемых кнопками, о возможностях поиска и т.д. (при наличии же такой информации - она либо слишком обрывочна, либо недостаточно понятна неподготовленному пользователю, либо (в России!) приведена на английском языке);

- обеспечение нормальной работы системы только с одним определенным браузером (browser) вследствие чего отсекаются от доступа пользователи, не имеющие нужного браузера, либо не привыкшие с ним работать;

- отсутствие поддержки нескольких кодировок, в которых может работать клиент (при наличии же такой поддержки пользователя заставляют подбирать необходимую кодировку из предлагаемых не всем понятных слов типа Win, KOI8, ISO, ALT, Mac и др.);

- отсутствие возможности работы со сложными запросами - режим ЭКСПЕРТ;

- отсутствие возможности работы с простыми запросами неподготовленному пользователю;

- отсутствие возможности работы со словарем;

- неудачное расположение кнопок, в частности, отсутствие нужной кнопки на экране и необходимость пролистывания текста до тех пор, пока она не появится;

- отсутствие выделения терминов из запроса в найденных документах;

- невозможность управления объемом выдачи;

- ограничения на возможности листания выдачи вперед/назад, начало/конец;

- засилье "картинок" замедляющих работу;

- и, наконец, плохо ищущие, зависающие, а то и просто пока не работающие, но уже выставленные в Интернет системы.

При разработке настоящей системы поиска в библиографических базах данных авторами была предпринята попытка учесть замеченные недостатки, а также использовать некоторые красивые решения проанализированных систем.

На основе анализа рассмотренных систем, были определены требования, которым должна соответствовать разрабатываемая система поиска в библиографических базах данных:

- система должна учитывать различные символьные кодировки, применяемые в различных операционных системах;

- интерфейс системы должен быть максимально понятен пользователю.

Это достигается за счет правильного построения экранной формы поискового запроса, а также наличия краткого, но при этом информационно емкого описания работы с системой;

- система должна предлагать различные варианты работы с ней в соответствии с подготовкой пользователя;

- будучи простой в использовании, система поиска должна обладать максимальным числом поисковых инструментов, одним из которых является словарь поисковых терминов;

- как программный продукт данная система должна обладать должной гибкостью, которая позволяет легко ее модифицировать в соответствии с текущими требованиями к ней.

Описание системы

Система поиска в базах данных ГПНТБ СО РАН представляет собой программно-информационный комплекс, основным назначением которого является предоставление возможности работы конечного пользователя сети Интернет с информационными ресурсами библиотеки.

Характерными чертами данной системы являются:

- простота в использовании;

- дружественный интерфейс;
- возможность регистрации пользователя с дальнейшей установкой индивидуальных прав работы в системе;
- возможность выбора режима поиска в зависимости от уровня одготовки пользователя;
- функциональная полнота с возможностью дальнейшего развития.

Данная система позволяет пользователю в ходе интерактивного диалога провести информационный поиск в выбранной им базе данных. В системе три режима формирования поискового запроса: режим НОВИЧОК, режим СТАНДАРТ и режим ЭКСПЕРТ. При использовании режима НОВИЧОК используется простейшая формулировка запросов; в режиме СТАНДАРТ пользователь более гибко может задавать поисковый запрос, а также может пользоваться словарем поисковых терминов для выбора значения поискового поля при формировании запроса к текущей базе данных; в режиме ЭКСПЕРТ пользователь должен хорошо знать поисковый язык, используемый в СУБД CDS/ISIS/M. Также во всех режимах пользователь может управлять видом выдаваемой ему информации через задание таких параметров выдачи, как форма и порция выдачи. Он также может свободно перемещаться в диапазоне найденной информации. При возникновении вопросов по использованию данной системы пользователь всегда может воспользоваться

справочной информацией.

Описываемая система эксплуатируется в ГПНТБ СО РАН по Интернет-адресу: <http://rust.spsl.nsc.ru>.

Для функционирования вышеуказанных программных средств необходимы 32-разрядная операционная система Windows NT с установленным в качестве WWW-сервера Интернет Information Server. Это требования к серверной части общесистемного программного обеспечения.

Использование при разработке системы языка JavaScript требует, чтобы конечные пользователи данной системы использовали Web-браузеры (Web-browsers), которые поддерживали бы данный язык. Тестирование и отладка системы проводилась для двух браузеров: Internet Explorer 3.0 и Netscape Navigator 3.0. Поэтому рекомендуется использовать данные браузеры версий не ниже указанных. Это является требованием к клиентской части общесистемного программного обеспечения.

Описание работы пользователя с системой

Сеанс работы пользователя с системой начинается с выбора символьной кодировки, в которой работает его браузер. В настоящий момент система поддерживает следующий набор кодовых таблиц: Win-1251, KOI8-R, ISO-8859-5, ALT-866 и Mac. Также существует поддержка работы на английском языке. Затем пользователю предлагается способ входа в систему. Он может войти как анонимный пользователь, или, если он ранее прошел регистрацию и получил идентификационный код, как зарегистрированный пользователь. Также, для получения идентификационного кода он может пройти регистрацию, чтобы в дальнейшем работать с системой как зарегистрированный пользователь. Регистрация дает пользователю возможность работать с большим перечнем баз данных и доступ к определенным задачам. Информация о зарегистрированных пользователях, учет их работы также ведется в базах данных CDS/ISIS/M.

Далее система предлагает выбрать пользователю режим работы с системой. Ему предлагаются следующие варианты: Новичок; Стандарт; Эксперт. В дальнейшем, в ходе работы с системой пользователь может в любой момент сменить текущий вариант поиска на любой другой из вышеуказанного списка. В следующем меню пользователю предлагается выбрать из списка доступных ему баз данных ту, с которой он намерен работать. После выбора базы данных пользователь переходит в меню формирования поискового запроса к текущей базе данных. Рассмотрим подробнее как формируются запросы в каждом из существующих режимов работы с системой.

Режим работы Новичок.

В данном режиме возможности поиска ограничены тем набором поисковых полей, который представлен в меню. Все поля связываются между собой в запросе через логическую связку И. Внутри каждого поля возможен ввод нескольких значений через символ пробела. Связь между этими значениями определяется выбором типа логической связки из списка связка, который находится в данном меню. Введенные пользователем значения поисковых полей усекаются справа, т.е. при вводе в поле автор значения Иванов система найдет как, например, значение <Иванов И.К.>, так и <Ивановский А.С.> Если значение поля представляет собой фразу (т.е. содержит несколько слов разделенных пробелами), то ее необходимо заключить в двойные кавычки (например, "Иванов А.А."). Необязательно заполнение всех поисковых полей для проведения поиска. Как и в любом другом режиме работы системы, пользователь может сменить текущую базу данных, режим работы системы, установить параметры вывода результатов проведенного поиска (форма выдачи и порция выдачи).

Режим работы Стандарт.

В данном режиме работы возможен ввод значений любых пяти поисковых полей текущей базы данных. Выбор названия поля осуществляется из списка ПОЛЕ, который содержит набор поисковых

полей определенных для текущей базы данных. После этого выбора в соответствующем поле ЗНАЧЕНИЕ вводится единственное значение поискового поля, усеченное справа. Для связи в запросе любого поискового поля с другими используется значение логической связки из соответствующего списка связка (расположенного справа от поля значение). Не обязательно заполнение всех поисковых полей для проведения поиска. Связки действуют только между непустыми поисковыми полями.

Существенным отличием данного режима работы системы от ранее описанного является присутствие в меню словаря поисковых терминов текущей базы данных. Для каждого из пяти поисковых полей в меню определена соответствующая кнопка Термины. Ее нажатие приводит к поиску значения соответствующего поля из словаря поисковых терминов текущей базы данных. При этом содержание поля ЗНАЧЕНИЕ выступает в качестве начального значения ключа словаря.

Проводя поиск значения поля в словаре, пользователь может продвигаться в словаре последовательно вперед, а также переходить в любую точку словаря, задав новое значение ключа. При нахождении в словаре нужного значения данного поля пользователь может вернуться в меню формирования поискового запроса вместе с найденным значением, которое помещается в поле ЗНАЧЕНИЕ. Если это значение не найдено, то возможно возвращение в режим формирования поискового запроса без выбора какого-либо значения.

После того как запрос полностью сформирован, пользователь активирует поиск в базе данных. Выдача результатов поиска происходит порциями (значение которых определено в меню формирования поискового запроса) с указанием информации о текущем положении в найденном диапазоне документов. Пользователь может перемещаться в данном диапазоне в любом направлении. После просмотра результатов поиска пользователь возвращается в меню формирования поискового запроса. При этом ранее сформированный запрос полностью сохраняется для дальнейшего использования.

Режим работы Эксперт.

Данный режим работы предполагает, что пользователь знаком с синтаксисом поискового языка CDS/ISIS/M.

Аналогично вышеописанным режимам пользователь может устанавливать параметры выдачи найденных в ходе запроса документов. Запрос может содержать логические связки, скобки. Значению каждого поискового поля должен предшествовать соответствующий префикс, получить который можно при активации ссылки ПОЛЕ.

Аналогично режиму Стандарт, в режиме Эксперт также существует возможность выбора значения поля из индексированного словаря текущей базы данных. При выборе нового значения из словаря оно добавляется в конец текущего поискового запроса. Пользователю лишь необходимо указать вид логической связи между этим значением и ранее сформированным запросом. Получение и просмотр результатов полностью соответствует режиму Стандарт.

Заключение

Описанная выше система находится в стадии опытно-промышленной эксплуатации. Реализованы основные задачи для работы с системой поиска в базах данных ГПНТБ СО РАН. Далее планируется реализовать комплекс крупных и мелких задач, которые превратят систему в универсальный и многофункциональный комплекс, позволяющий в полном объеме реализовать возможности Интернет как основной информационной системы настоящего и будущего.

Данная система может гибко модифицироваться и неограниченно дополняться, что важно при мощном развитии Интернет и возложении на нее все большего числа функций обычных средств связи.

Литература

1. Мазов Н.А., Жижимов О.Л., Поморцев А.Ю. Организация доступа к базам данных CDS/ISIS-M из Internet // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: 3-я международная конференция "Крым-97": Материалы конференции. - Форос, 1996. - Т. 2. - С. 177.

2. Мазов Н.А., Жижимов О.Л., Фролов А.С., Баженов С.Р. Создание системы открытого публичного доступа из Интернет к распределенным библиографическим базам данных CDS/ISIS // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: 4-я международная конференция "Крым-97": Материалы конференции. - Судак, 1997. - Т. 2. - С. 482-483.