



# Шестая Международная Конференция “Крым 99”

Конференция проводится в рамках мероприятий ИФЛА 1999 г.

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

*Материалы конференции*

**Том 1**

**СУДАК**

(основная программа)

Симферополь, Феодосия, Старый Крым, Коктебель

(выездные заседания)

Автономная Республика Крым, Украина

5-13 июня 1999

**WEB-ориентированная информационно-поисковая система  
для доступа к базам данных CDS/ISIS**

**WEB-oriented Information Retrieval System for Access to CDS/ISIS Databases**

**WEB-орієнтована інформаційно-пошукова система для доступу  
до баз даних CDS/ISIS**

*Мазов Н.А., Малицкий Н.А., Баженов С.Р., Баженов И.С., Жижимов О.Л.*

*Объединенный Институт Геологии, Геофизики и Минералогии СО РАН, Государственная Публичная Научно-техническая Библиотека СО РАН,  
Новосибирск, Россия*

*N.A. Mazov, N.A. Malitsky, S.R. Bazhenov, I.S. Bazhenov, O.L. Zhizhimov  
Joint Institute of Geology, Geophysics and Mineralogy of the Russian Academy of Sciences'  
Siberian Branch, Novosibirsk, Russia*

*Мазов М.А., Малицкий Н.А., Баженов С.Р., Баженов И.С., Жижимов О.Л.  
Об'єднаний інститут геології, геофізики та мінералогії СВ РАН, Державна публічна  
науково-технічна бібліотека СВ РАН, Новосибірськ, Росія*

В докладе описується Web-орієнтована інформаційна система, що дозволяє здійснювати доступ за допомогою WWW к бібліографічним базам даних, підтримуваною інформаційно-пошуковою системою CDS/ISIS. Основа при функціонуванні системи є програма WWW-ISIS, розроблена в BIREME/PAHO/WHO, Бразилія. Система передбачає реалізацію повного спектра функцій, необхідних при роботі з бібліографічними базами даних.

The paper describes Web-oriented information system, which provides access to bibliographic databases, supported by CDS/ISIS information retrieval system. The system is based on the WWW-ISIS software designed in BIREME/PAHO/WHO, Brasil. The system provides a full spectrum of functions necessary for bibliographic databases operation.

В Сибірському Відділенні РАН історично склалося так, що його бібліотечно-інформаційні ресурси велись і ведуться під управлінням СУБД CDS/ISIS [1]. На теперішній момент часу накоплено достатньо великих інформаційних масивів під управлінням CDS/ISIS, основними держателями яких є ГПНТБ СО РАН і інформаційно-бібліотечні центри інститутів Сибірського Відділення, одним з яких є Інформаційно-бібліотечний Центр ОИГГиМ СО РАН. Вибір даної СУБД був обумовлений в своє час наступними факторами:

- відносно висока продуктивність обробки даних, не вимагаючи значительних обчислювальних ресурсів (технічна оснащеність бібліотек і інформаційних центрів завжди залишала бажати кращого);
- специфіка структури бібліографічних даних (вільно-текстова нечислова структурована інформація, що відповідає рекомендаціям стандарту ISO-2709) вимагає спеціалізованого програмного забезпечення, яким і є СУБД CDS/ISIS;
- CDS/ISIS є комерційним продуктом.

Для роботи з БД в стандартному CDS/ISIS користувач повинен мати певні знання мови запитів при проведенні пошуку в БД, а також навички роботи в середі СУБД CDS/ISIS. Це створювало певні труднощі при роботі з БД як наукового співробітника, так і рядового читача, не завжди добре підготовленого до роботи з інформаційно-пошуковою системою. Тому при створенні принципово нової схеми доступу до бібліотечно-інформаційних ресурсів СО РАН необхідно було вирішити цю основну проблему.

В теперішній час використання сервісних можливостей, надаваних мережею Internet, все глибше проникає в повсякденну діяльність людини. Це здебільшого пов'язано з її відкритістю і простотою використання. Користувач швидко оволодіває і звикає до використання уніфікованих інтерфейсів, побудованих на механізмі гіперссылки і пошукових форм. При цьому використовуються всі переваги технології "клієнт-сервер", що зменшує вимоги до користувальницького (клієнтського) програмного забезпечення. Для роботи потрібно, щоб комп'ютер користувача був підключений до мережі Internet і на ньому був встановлений один з стандартних Web браузерів для перегляду інформації. Найбільш відомими браузерами в теперішній час є "Microsoft Internet Explorer" і "Netscape navigator", без яких сьогодні вже важко уявити комп'ютер, підключений до мережі Internet. Тому, враховуючи вище сказане, на наш погляд, цілком природно було б надати користувачам можливість роботи з інформаційними ресурсами в мережі Internet.

Ідеальним рішенням в даній ситуації є використання Internet технологій на основі протоколу HTTP. При цьому ми як би переходимо до технології "клієнт-сервер", яка позбавлена недоліків технології "файл-сервер", т.е. досягається зменшення мережевого трафіку, зменшення вимог до обчислювальних ресурсів робочої станції користувача, відсутність дубльованих ресурсів (формально вони розташовані в одному місці). Схеми взаємодії при цьому виглядає наступним чином: клієнт (користувач) надсилає запит серверу, який йому (сервером) виконується, після чого результат запиту передається назад клієнту. Клієнт в даному випадку - це робоча станція з встановленим Web браузером. Сервер - станція з значительними обчислювальними і дисковими ресурсами (для зберігання БД) з встановленим Web сервером (для взаємодії з Web браузером клієнта), який через сервер баз даних CDS/ISIS отримує інформацію з БД в відповідь на запит клієнта і надсилає її назад клієнту. Користувач може використовувати будь-яку ОС (Windows, Unix, Mac і т.д.) і Web браузер (Internet Explorer, Netscape Navigator і др.). При цьому передбачається, що інтерфейс роботи з БД буде всюди мати один і той же вигляд. Таким чином досягається уніфікація користувальницького інтерфейсу системи. Набір елементів екранної форми при цьому повинен бути жорстко типізований і простий. Це дозволяє спростити вимоги до підготовки користувача в плані роботи з БД. Їм достатньо уміти працювати з Web браузером і вивчити довідковий матеріал системи про правила складання пошукового запиту.

Ключовим моментом при створенні такої системи доступу до ресурсів CDS/ISIS був вибір пошукової машини, яка б забезпечувала повну функціональність і високу швидкість. Під функціональністю в даному випадку розуміється:

- можливість взаємодії з Web сервером для обміну інформацією при обробці запитів клієнта до БД і видачі результатів пошуку;
- найбільш повний набір функцій для обробки даних БД CDS/ISIS (пошук, робота зі словником, вивід результатів пошуку, введення, оновлення і др.).

В теперішній час існує кілька варіантів серверів БД CDS/ISIS (в т.ч. власної розробки), на яких ми не будемо зосереджувати увагу. Однак було прийнято рішення вибрати програму WWWISIS (розробка BIREME/PAHO/WHO, Бразилія) [2]. Варто зауважити, що на даний момент BIREME/PAHO/WHO є одним з провідних центрів з розробки програмного забезпечення для БД під управлінням CDS/ISIS. Одним з його напрямків є розробка сервера БД CDS/ISIS WWWISIS для WWW Windows NT. Існують також версії сервера WWWISIS для найбільш поширених апаратних платформ. Сервер WWWISIS взаємодіє з Web сервером через інтерфейс CGI, який на сьогодні підтримується всіма існуючими Web серверами. Використання Win32 API в версії WWWISIS для MS Windows NT, дозволяє суттєво підвищити швидкість одночасної (багатокористувальницької) обробки запитів до БД CDS/ISIS (багатопоточність при виконанні запиту) порівняно з швидкістю традиційної СУБД CDS/ISIS (ЮНЕСКО). Варто також зауважити, що поряд з сервером WWWISIS виробник пропонує документацію [3], що містить детальне описання всіх функціональних можливостей сервера, спектр яких достатньо широкий. Достатньо зауважити, що сервер WWWISIS реалізує наступні функції:

- взаимодействия с Web сервером,
- выполнения поискового запроса к БД,
- работы с словарем поисковых терминов,
- вывода результатов поиска,
- ввода новых записей в БД,
- редактирования и удаления существующих записей БД,
- сортировки результатов поиска по различным полям БД,
- блокирования данных,
- экспорта/импорта записей БД в формате ISO-2709, а также набор функций для диагностики работы сервера в процессе обработки запросов пользователя.

При этом производитель постоянно работает над развитием функциональности сервера в его последующих версиях. Немаловажным фактором также явилась возможность бесплатного приобретения данного программного продукта.

Далее мы приведем ряд требований к интерфейсу создаваемой системы, которые на наш взгляд являлись наиболее актуальными:

- обеспечение возможности работы пользователя с системой, с учетом используемого им клиентского программного обеспечения (возможность выбора кодовой таблицы (Windows, Unix, Mac, Dos);
- безкоризненная работа системы с доминирующими Web браузерами (IE 3.0 и Netscape 3.0 и выше);
- выбор оптимального разрешения экрана (800x600);
- использование стандартных унифицированных средств (HTML, JavaScript) для отображения запрашиваемой информации);
- полностью текстовый интерфейс экранных форм и как следствие минимизация времени доступа к информации (исключение графических изображений, что существенно позволяет снизить сетевой трафик);
- обработка полноты и корректности данных экранных форм на стороне клиента (возможностями используемого Web браузера) без участия сервера (уменьшения количества транзакций “клиент-сервер”);
- эргономичность интерфейса (использование спокойных тонов, расположение кнопок и окон для ввода, сохранение значений экранной формы при возвращении к ней, прозрачность значений элементов экранных форм, возможность выбора пользователем режима поиска с учетом своих требований и квалификации, наличие краткой и конкретной справочной информации по работе с системой);
- полная функциональность в рамках решаемой задачи поиска информации (несколько режимов поиска, наличие разнообразных сервисных возможностей: выбор поисковых терминов из словаря БД; выбор формата, порций выдаваемой информации; навигация внутри выданных результатов поиска (переход к любой записи из найденного диапазона); подсветка поисковых ключей в тексте найденных записей;
- сохранение результатов поиска в файл);
- наличие контактного адреса службы администрирования БД для осуществления обратной связи с пользователями (внесения замечаний и предложений по дальнейшему развитию системы).

Исходя из вышеизложенных требований авторами в СО РАН (ГПНТБ, ОИГГиМ) была разработана Web информационно-поисковая система, которая сегодня успешно функционирует в трех организациях Сибирского Отделения РАН. Ниже приведем основные свойства реализованной системы.

Доступ к ресурсам той или иной организации может обеспечиваться как для пользователей внутренней сети, например доступ к БД Информационно-библиотечного Центра ОИГГиМ СО РАН (в этом случае в системе устанавливается ограничение прав доступа только по IP-адресам пользователей), так и для внешних пользователей, например доступ к БД ГПНТБ СО РАН (в этом случае в системе предусмотрена полная аутентификация пользователей).

БД в описываемой системе включены в группы для структуризации информации. Пользователь может выбрать для работы одну или несколько групп БД. В этом случае будет сформирован объединенный список БД из этих групп. Это особенно важно при проведении одновременного поиска в БД, выбранных пользователем. В зависимости от сложности поискового запроса пользователь может выбрать любой (по возрастанию сложности) из четырех режимов поиска: *Простой*; *Стандарт*; *Эксперт*; *Мультипоиск*.

Все режимы кроме первого дают возможность выбора поискового термина из словаря текущей БД и переноса его в формируемый запрос (режимы *Эксперт* и *Мультипоиск*, в отличие от режима *Стандарт*, предлагают, помимо выбора одного значения поискового термина словаря БД, осуществлять их групповой выбор). Режим *Мультипоиск* предназначен для проведения одновременного поиска в одной или нескольких базах данных. Выдача результатов поиска происходит в выбранном формате выдачи и порциями, определяемыми пользователем. При выдаче результатов поиска предлагается следующий сервис:

- оперативная смена формата выдачи;
- прямой переход к любой странице или записи из списка найденных записей;
- подсветка значений поисковых терминов в тексте выдаваемых записей;
- сохранение текущей записи или списка отмеченных записей в ASCII файл на компьютере клиента;
- оперативное проведение повторного уточняющего поиска по заранее определенным полям (например, полю *Автор* и др.) с возможностью возврата к основному запросу.

Для всех режимов поиска существует подробная справка по формированию поискового запроса. Это особенно важно при использовании режимов *Эксперт* и *Мультипоиск*, т.к. в этом случае пользователь должен в достаточной форме владеть языком составления запросов для БД CDS/ISIS. Пользователь в любой момент работы с БД может перейти из текущего в любой другой режим поиска, сменить текущую БД или группу БД. Имеется возможность получить краткое описание текущей БД.

В процессе эксплуатации описываемой системы доступа к БД была разработана и внедрена система сбора и обработки статистики обращений к БД организации. В реализованной версии системы предусмотрен сбор следующей статистической информации: информация, идентифицирующая клиента; используемые кодировка и Web браузер; режим поиска и источник информации; дата и время проведения поиска; трафик. Эти данные в свою очередь позволяют получать отчеты об использовании и актуальности предоставляемых ресурсов (степени спроса на них); проводить анализ дальнейшего развития системы в целом; выявлять приоритетные научные направления, а также решать ряд других задач, связанных с использованием информационных ресурсов различными группами пользователей.

В заключение отметим, что описанная система находится в эксплуатации более года 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://info.spsl.nsc.ru> – ГПНТБ СО РАН, <http://geolibr.uiggm.nsc.ru> – ОИГГиМ СО РАН).

#### *Литература*

1. Пакет прикладных программ CDS/ISIS/M версия 2.3.: Методические материалы и документация по пакетам прикладных программ. Вып. 70 - М.: МЦНТИ, 1991.
2. <http://www.bireme.br/wwwisis.htm> BIREME/PAHO/WHO. WWWISIS: a word-wide web server for ISIS-databases. Ver. 3.0. Sao Paulo, Brazilian, Oct. 1997.