



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 2

СУДАК
(основная программа)
Симферополь, Феодосия, Старый Крым, Коктебель
(выездные заседания)

Автономная Республика Крым, Украина

5-13 июня 1999

Интеграция Z39.50 и CDS/ISIS: состояние и перспективы развития

Z39.50 and CDS/ISIS Integration: Current State and Prospects for Development

Інтеграція Z39.50 та CDS/ISIS: стан і перспективи розвитку

Мазов Н.А., Жижимов О.Л.

Объединенный институт геологии, геофизики и минералогии СО РАН, Новосибирск, Россия

Nikolai A. Mazov, Oleg L. Zhizhimov

Joint Institute of Geology, Geophysics and Mineralogy of the Russian Academy of Sciences' Siberian Branch, Novosibirsk, Russia

Мазов М.А., Жижимов О.Л.

Об'єднаний інститут геології, геофізики та мінералогії СВ РАН, Новосибірськ, Росія

Описываются пути интеграции СУБД CDS/ISIS с протоколом Z39.50. В качестве примера реализации рассматривается провайдер данных CDS/ISIS для сервера Z39.50, созданного в ОИГГиМ СО РАН. Обсуждаются пути преодоления ограничений, присущих базовому CDS/ISIS в свете этой интеграции.

The ways of integration of CDS/ISIS database management system with Z39.50 protocol are described. CDS/ISIS data provider for Z39.50 server, created by the Institute is considered as an example of realization. Discussed are the ways to overcome the inherent CDS/ISIS limitations by way of this integration.

В настоящее время, как в России, так и за рубежом существует значительное количество библиографических баз данных (БД), поддерживаемых программным обеспечением CDS/ISIS, разработанным в ЮНЕСКО [1]. В то время как различные СУБД реляционного типа ориентированы на работу с БД, имеющих табличные формы, CDS/ISIS в своей основе имеет свободно-текстовый формат, наиболее адекватно представляющий структурированные нечисловые данные в форме записей произвольной переменной длины. Библиографическая информация может быть представлена наилучшим образом именно в такой форме. Достаточно сослаться на рекомендации стандарта ISO-2709, предназначенного для обмена библиографической информацией на магнитных носителях. При этом необходимо заметить, что CDS/ISIS в своей основе построен на использовании ISO-2709 (внутренняя организация основного файла записей, режимы импорта и экспорта в формате ISO-2709). Поэтому традиционно сложилось так, что БД используемые в СО РАН (ГПНТБ СО РАН, ОИГГиМ СО РАН, ИК СО РАН и др.) ведутся под управлением CDS/ISIS.

Ряд организаций, располагающих такими БД, работают в настоящее время над тем, чтобы обеспечить унифицированный сетевой доступ к БД CDS/ISIS по стандартным протоколам для более широкого обращения к ним ученых, исследователей или обычных граждан [2].

Все вышесказанное, а также предшествующий опыт работы авторов доклада по разработке программного обеспечения доступа к БД CDS/ISIS [3,4], явилось толчком для создания специализированного провайдера данных CDS/ISIS (далее: провайдер CDS/ISIS) в рамках разрабатываемой в ОИГГиМ СО РАН информационной системы на основе протокола Z39.50 [5]. Насколько известно авторам, настоящая разработка является уникальной и актуальной не только для организаций СО РАН, но и для других организаций России, использующих в своей работе БД под управлением CDS/ISIS.

При построении провайдера CDS/ISIS принимались во внимание следующие необходимые условия:

- соответствие спецификациям программного интерфейса абстрактного провайдера данных сервера Z39.50;
- эффективные по времени и потребляемым системным ресурсам способы обработки информации в многозадачном окружении;
- переносимость программного кода провайдера на различные аппаратные платформы в средах операционных систем Windows NT, UNIX;
- низкая себестоимость разработки.

Более подробного обсуждения из вышеперечисленных условий заслуживает первое, которое в свою очередь предполагает выполнение минимального набора таких требований как:

- обработка запросов и атрибутов;
- отображение записей БД CDS/ISIS на абстрактную схему данных Z39.50;
- обеспечение доступа к словарию БД CDS/ISIS;
- отображение записей БД и элементов абстрактной схемы данных в стандартных форматах.

В соответствии с вышеизложенными требованиями авторами был реализован провайдер CDS/ISIS для сервера Z39.50 (версий: 1 – Windows NT, 2 – Windows NT и UNIX). Как уже было отмечено выше,

этому предшествовал большой опыт разработок программного обеспечения по доступу к БД CDS/ISIS, что во многом послужило основой при создании провайдера CDS/ISIS.

Провайдер CDS/ISIS для сервера Z39.50 версии 1 реализован в виде набора функций, вызываемых из динамически загружаемой библиотеки (DLL) для Windows NT. Основной задачей функций этой библиотеки является отображение запросов на обработку данных Z39.50 в логику запросов CDS/ISIS и отображение ответов CDS/ISIS в логику ответов для Z39.50. В качестве поисковой машины и основного средства доступа к БД были использованы функции библиотеки WWWISIS.DLL (Бразилия, BIREME/ПАНО/WHO, версии 3.0) [6].

Провайдер CDS/ISIS для сервера Z39.50 версии 2 организован как встраиваемый серверный модуль сервера. Этот модуль является оригинальной разработкой авторов настоящего доклада. В отличие от версии 1 сервера Z39.50, для провайдера CDS/ISIS версии 2 были реализованы следующие аппаратно-независимые компоненты:

- модуль извлечения записей БД CDS/ISIS;
- модуль доступа к словарю БД CDS/ISIS;
- уникальная поисковая машина;
- интерпретатор потоковой обработки данных для внутреннего форматирования.

Адаптация конкретной БД CDS/ISIS к работе по протоколу Z39.50 для описываемого провайдера заключается в корректном построении ряда служебных таблиц, которые не предусматриваются при работе базового CDS/ISIS:

- таблицы конфигурации БД CDS/ISIS (имя БД, физическое расположение основных файлов БД, кодовая таблица и т.д.);
- таблицы отображения индексируемых полей БД CDS/ISIS на наборы поисковых атрибутов Z39.50;
- таблицы отображения множества меток реальной БД CDS/ISIS во внутренние метки абстрактной схемы данных сервера;
- таблиц описания абстрактной схемы данных (иерархическая схема, набор меток, набор элементов, набор вариантов, наборы мапирования на другие схемы данных).

Программно-технические характеристики провайдера CDS/ISIS:

- Поддерживаемые платформы:
 - Windows NT 4.0 SP3 и выше (Intel) (версии 1 и 2);
 - Solaris 2.* (Intel, Sparc) (версия 2);
 - FreeBSD (Intel) (версия 2);
 - AIX (PowerPC) (версия 2).
- Система запросов – Type-0 (версия 1), Type-1 (версии 1 и 2), количество поисковых термов – неограничено, правостороннее усечение, логика И, ИЛИ, И-НЕ;
- Форматы внешнего представления (RecordSyntax);
 - SUTRS;
 - GRS-1;
 - ISO-2709 MARC форматы (RUSmarc, USmarc, Unimarc).
- Наборы элементов (B-Brief, F-Full, S-Spisok);
- Временные дисковые файлы не используются.

Описанные провайдеры CDS/ISIS прошли апробацию в составе серверов Z39.50, функционирующих в СО РАН (ОИГГиМ СО РАН, ГПНТБ СО РАН, ИВТ СО РАН). Эксплуатация провайдеров CDS/ISIS показала высокую эффективность при поиске, извлечении и предоставлении информации из БД CDS/ISIS, а также высокую степень надежности при нештатных ситуациях. Сервер Z39.50 (z3950://geolibr.uiggm.nsc.ru:210), в составе которого функционирует провайдер, прошел успешное тестирование и сертификацию на соответствие стандарту Z39.50, проводимую компанией IndexData [7] и включен в список серверов Z39.50 доступных в мире.

Описанная выше технология включения БД CDS/ISIS в информационную систему на базе протокола Z39.50 обеспечивает, помимо всего прочего, расширение функциональных характеристик базового CDS/ISIS, а именно:

- возможность работы с БД CDS/ISIS в режиме “клиент-сервер”;
- реализация мультибазовой концепции работы с данными CDS/ISIS;
- возможность доступа к данным CDS/ISIS по стандартному открытому протоколу;
- возможность интеграции БД CDS/ISIS с другими системами в гетерогенной информационной системе;

В заключение отметим, что описанный провайдер CDS/ISIS позволяет снять ряд ограничений, существующих в CDS/ISIS, с которыми сталкиваются пользователи при работе с базовым программным обеспечением CDS/ISIS.

В настоящее время работа по совершенствованию провайдера данных CDS/ISIS продолжается. В частности, прорабатывается технология ведения и модификации данных CDS/ISIS в рамках расширенного сервиса Z39.50.

Литература

1. Шрайберг Я.Л. Современное состояние развития библиотечных технологий и место CDS/ISIS. НТБ, 1995, Спец.выпуск, стр.5-8.
2. Шрайберг Я.Л., Воройский Ф.С. Автоматизированные библиотечно-информационные системы России. М.: "Либерия", 1996, 272 с.
3. Мазов Н.А., Жижимов О.Л., Поморцев А.Ю. Организация доступа к базам данных CDS/ISIS-M из Internet. Матер.3 Междунар.Конф. "Крым-96", т.2, 1996, С.177-179.
4. Жижимов О.Л., Мазов Н.А., Фролов А.С. Доступ к базам данных ISIS из Internet и построение распределенной информационной системы. Вычислительные технологии, т.2, №3, 1997, стр.45-50.
5. ANSI/NISO Z39.50-1995. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification. Z39.50 Maintenance Agency Official Text for Z39.50-1995, July 1995.
6. BIREME. WWWISIS: a word-wide web server for ISIS-databases. Version 3.0. Sao Paulo, Brazilian, October 1997, 57 p.
7. Index Data.