



# **Восьмая Международная Конференция «Крым 2001»**

*Конференция проводится в рамках мероприятий ИФЛА 2001 г.*

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

***Тема 2001 года:***

***Производители и пользователи  
печатной и электронной информации  
на пути к информационному  
обществу***

***Труды конференции  
Том 2***

**СУДАК**

**(основная программа)**

**Ялта, Алушта, Феодосия, Керчь, Старый Крым**

**(выездные заседания)**

**Автономная Республика Крым, Украина**

**9-17 июня 2001**

**Издательство ГПНТБ России  
Москва, 2001**

**Перспективы применения протокола Z39.50 в информационном сообществе России**  
**Advanced Applications of the Z39.50 Protocol in Russian Information Community**  
**Перспективи застосування протоколу Z39.50 в інформаційному співтоваристві Росії**

*Жижимов О.Л., Мазов Н.А.*

*Объединенный Институт Геологии, Геофизики и Минералогии СО РАН, Новосибирск, Россия*

*Oleg L. Zhizhimov and Nikolai A. Mazov*

*Joint Institute of Geology, Geophysics and Mineralogy,  
Siberian Division of the Russian Academy of Sciences, Novosibirsk, Russia*

*О.Л. Жижимов, М.А. Мазов*

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

В настоящем докладе рассмотрены возможные области применения протокола Z39.50 в информационном сообществе России. Основной акцент сделан на построение распределенных информационных систем. Такие системы могут интегрировать различную информацию: библиографическую, метаданные общего характера, метаданные по электронным коллекциям, музейную информацию, данные геоинформационных систем, тезаурусы и классификационные схемы и т.п. В качестве примера рассмотрена распределенная информационная система СО РАН.

The possibilities of the Z39.50 protocol use in Russian information community of are discussed. The basic accent is made distributed information systems creation. Such systems can integrate various types of information: bibliographic information, general metadata, metadata on electronic collections, museum information, data from geoinformation systems, thesauri and classification systems, etc. As an example the distributed information system SD RAS is analyzed.

У даній доповіді розглядаються можливі галузі застосування протоколу Z39.50 в інформаційному співтоваристві Росії. Основний акцент зроблено на побудову розподілених інформаційних систем. Такі системи можуть інтегрувати різну інформацію: бібліографічну, метадані загального характеру, метадані по електронних колекціях, музейну інформацію, дані геоінформаційних систем, тезауруси і класифікаційні схеми і т.п. Як приклад розглянуто розподілену інформаційну систему СВ РАН

Широкое распространение локальных и глобальных сетей в России сегодня делает актуальным вопрос ( построении информационных систем в «правильной» идеологии и архитектуре. Это связано с потребностях распределенного хранения и обработки данных, с необходимостью интеграции различных информационных систем и унификации пользовательских интерфейсов. В конечном счете, речь идет о стандартизации правил работы с данными в сетевом окружении. Именно для этих целей был создан протокол ANSI/NISO-Z39.50 [1] получивший международное признание (IS023950).

В последние год-два в России наметился явный прогресс во внедрении протокола Z39.50 [2]. Это связано как с объективными (существует потребность информационного сообщества в создании интероперабельных систем), так и с субъективными причинами -поощрение применения Z39.50 через гранты РФФИ и ИОО «Фонд Сороса». Гранты последней организации сыграли ведущую роль в создании инфраструктуры поддержки протокола Z39.50 в библиотечном сообществе России.

### **Программное обеспечение**

На сегодняшний день в России существуют официально зарегистрированные группы, разрабатывающие серверное и клиентское программное обеспечение (ПО) Z39.50:

(ID=148) Государственный Технический Университет (г.Санкт-Петербург);

(ID=155) Объединенный Институт Геологии, Геофизики и Минералогии СО РАН (ОИГГМ СО РАН (г. Новосибирск);

(ID=163) Сибирский Технический Центр (г. Томск);

(ID=172) GIPER (г. Москва).

(ID=175) Междисциплинарный Центр Обучения Государственного Университета (г.Санкт-Петербург);

Кроме официально зарегистрированных групп разработчиков существуют незарегистрированные группы. Насколько известно авторам, это:

ЗАО «LIBER» (г. Москва);

Московский Государственный Университет (г. Москва).

Эти группы показали свою способность обеспечить российские потребности в серверном ПО Z39.50, во всяком случае, для библиотечных систем. Разрабатываемое ПО поддерживает российскую специфику представления данных с множеством кодовых таблиц и российские национальные форматы (RUSMARC) в рамках существующего стандарта Z39.50.

Следует заметить, что на сегодняшний день основной упор разработчиков сделан на серверное ПО. Это естественно, т.к. без наличия надежных серверов Z39.50 невозможно развивать эту технологию. При этом, как правило, в серверах реализованы только функции поиска и извлечения данных, реже встречаются функции просмотра словарей и расширенного сервиса. Практически отсутствует поддержка Explain. Также ни один сервер не поддерживает сегментацию. Что касается поиска и извлечения данных, все серверы поддерживают набор поисковых атрибутов Bib-1 (поиск в библиографических базах данных) и простые форматы представления (SUTRS, RUSMARC, иногда USMARC). Эти функциональные возможности серверов можно признать удовлетворительными для обеспечения доступа к каталогам библиотек, но явно недостаточными для обеспечения доступа к другим информационным ресурсам, например, к аналитической библиографической информации, к полным текстам, к мультимедиа ресурсам и т.п.

Такая же явная направленность ПО Z39.50 на библиотечные системы наблюдается и для клиентского ПО и для ПО промежуточного слоя (шлюзы Z39.50-HTTP). На сегодняшний день только шлюзы, установленные в СО РАН, позволяют выйти за рамки Bib-1. Все остальные шлюзы ориентированы на очень узкую задачу — обеспечение доступа к электронным каталогам библиотек.

Гораздо хуже в России обстоит дело с автономными программами-клиентами Z39.50. В настоящее время авторам неизвестна ни одна российская готовая (законченная) программа-клиент Z39.50, хотя их разработка ведется в ОИГГМ СО РАН (г. Новосибирск), в Центре «Открытые Библиотечные Системы» (г. Санкт-Петербург), в ЗАО «LIBER» (г. Москва), в компании «GIPER» (г. Москва), в Сибирском Техническом Центре (г. Томск) и ГПНТБ России (г. Москва). Наверное, следует надеяться, что такие продукты скоро появятся хотя бы для библиографических данных.

### **Библиотечные системы**

Как видно из вышеизложенного, основное применение протокол Z39.50 в России нашел в библиотечных системах. При этом уровень использования его в этих системах различен: от применения во всех внутренних библиотечных технологических процессах, до ограничения областью поиска и извлечения информации из электронных каталогов.

Можно предположить, что дальнейшее внедрение новейших информационных технологий в библиотечные технологические процессы, развитие распределенных информационных систем на основе стандартных протоколов и унифицированных механизмов, а также усилия отдельных энтузиастов и влиятельных организаций типа Института «Открытое Общество», РФФИ и LibWeb продолжат внедрение Z39.50 в России. Во всяком случае, на сегодняшний день библиотечные информационные системы -наиболее подготовленная почва для внедрения и использования Z39.50.

### **Метаданные и информационные ресурсы 635**

Наряду с доступом к библиографическим данным, Z39.50 содержит механизмы обеспечения доступа к распределенным базам данных иного наполнения. Основная проблема при этом — унификация и глобальная стандартизация поисковых атрибутов и схем данных для конкретной области применения. Однако многие области проработаны уже в мире достаточно хорошо. В частности, хорошо проработано описание информационных ресурсов общего характера на основе GILS (Global Information Locator Service) (<http://www.gils.nct/>). Механизмы, заложенные в Z39.50, и детальная проработка схемы метаданных GILS, на наш взгляд, заманчивы для организации распределенных информационных систем. В этих системах поиск информации может происходить по распределенным базам метаданных в стандарте GILS. При этом нет нужды скрывать описание ресурсов MARC-форматами, которые ориентированы в основном на библиографические описания. Наличие специальных тэгов в GILS-записях позволяет легко связать, в том числе и через гипертекстовую ссылку, описание ресурса с самим ресурсом, предоставив тем самым пользователям доступ к полным текстам, мультимедиа и другим базам данных через однотипные интерфейсы со сквозной поисковой системой, одинаковой для всего мира.

Прототип подобной реализованной системы подробно описан в работе [3].

### **Тезаурусы и рубрикаторы**

С развитием технологий построения больших распределенных информационных систем, включающих в себя множество различных баз данных (БД), достаточно актуальным становится вопрос поиска информации в БД с использованием тезаурусов и классификационных схем. Более того, в распределенной информационной

системе логично обеспечить доступ к БДТК (базы данных тезаурусов и классификационных схем) в той же самой технологии, в которой осуществляется доступ к БД, т.е. в технологии «клиент-сервер» с использованием единого протокола Z39.50 [4].

Однако, следует отметить, что сам по себе сетевой доступ к БДТК хотя и предоставляет интерес, но, на наш взгляд, не несет особого смысла без возможности одновременного выхода в поисковую систему по БД. Иными словами, просматривая статьи тезауруса или классификационной схемы, хотелось бы проводить одновременный поиск в БД по соответствующим ключевым словам или кодам рубрик. Здесь Z39.50, ввиду стандартизации поискового механизма, дает уникальную возможность подключать к параллельному поиску совершенно различные БД.

В качестве иллюстрации применения Z39.50 для доступа к БДТК можно сослаться на соответствующую подсистему распределенной информационной системы ОИГГМ СО РАН [5]. При этом доступ к БДТК осуществлялся через шлюз Z-CGI с одновременным выходом на поиск в библиографических базах данных, в которых коды рубрикатора являются поисковыми признаками.

### **Координатная привязка и ГИС-технологии**

Специальным типом метаданных является географическая информация о положении и размерах описываемых объектов. Для четкой идентификации объекта и его визуализации на карте необходимо использовать данные ГИС систем. Это могут быть контуры материков, океаны, реки, озера, а также специфическая информация, такая как геологические складки, формации с их пространственными данными - размерами, площадями, мощностями. Актуальна и атрибутивная информация об объектах — названия, геологический возраст и т.п. Несмотря на то, что некоторая географическая информация заложена в схемы метаданных GILS, CIMI, MARC21, полное описание ГИС-объекта возможно только в специальных схемах данных. В качестве такой специальной схемы в Z39.50 принято считать схему GEO [6], основанную на стандарте метаданных FGDC (<http://www.fgdc.gov>).

Нам кажется, что актуальность поисковых систем, поддерживающих специфические для ГИС-систем модели поиска, со временем будет возрастать. И здесь, несомненно, будет очень полезен опыт работы как со схемами метаданных типа GEO, так и схемами метаданных типа GILS, которые содержат минимальную географическую информацию, но позволяют построить информационную систему с поиском по координатам. В качестве примера можно сослаться на работу [3], где интеграция поисковой подсистемы Z39.50, элементов GILS и ГИС-информации, позволила построить пользовательские интерфейсы с реализацией визуального поиска по масштабируемой карте и отображением на ней записей из баз метаданных.

### **Музеи и доступ к ресурсам по культурному наследию**

Если в первые годы своего существования протокол Z39.50 использовался преимущественно для организации доступа к библиографическим ресурсам, то на сегодняшний день область его применения существенно расширена, он используется для доступа к научно-технической информации, к геоинформационным ресурсам, к глобальным базам метаданных и др. Существует также возможность предоставления доступа по Z39.50 к музейной информации и цифровым коллекциям. Правила, которые регламентируют работу с музейной информацией в Z39.50, сформулированы в документах CIMI (Consortium for the Computer Interchange of Museum Information) и составляют профиль CIMI [7]. Элементы этого профиля имеют глобальные идентификаторы и являются частью международного стандарта ISO-23950.

Полное описание профиля CIMI достаточно объемно и здесь приводиться не будет. Ниже будут отмечены только некоторые аспекты, которые на наш взгляд характерны для этого профиля [8]. Во-первых, все серверы CIMI обязаны поддерживать запросы типа 1 (RPN) с набором поисковых атрибутов CIMI-1, который содержит как атрибуты общего вида (Title, Author, Abstract и т.п.), так и атрибуты, специфические для музейных коллекций (repositoryPlace, dateOfOrigin, placeOfOrigin, dateCollected и т.п.). Кроме этого набор атрибутов CIMI содержит в себе поисковые атрибуты Dublin Core (<http://dublincore.org/>). Во-вторых, схема данных и абстрактная структура записи CIMI основана на иерархической модели с использованием схем Digital Collections и Dublin Core. Элементы последней явно присутствуют в CIMI на первом уровне вложения. Собственно структура записи CIMI вложена в элемент actualDO записи Digital Collection [7].

Схема CIMI позволяет представить в стандартном виде описательную информацию о музейных ресурсах, т.е. является схемой метаданных. Связь с изображениями и полными текстами осуществляется через URL в элементе `mrObject/rendition/resource`.

В ситуации, когда музейное информационное сообщество России только начинает создавать информационные системы с предоставлением доступа к ним из Интернет, нам кажется очень актуальным вопрос о выборе правильной технологии и правильной схемы метаданных. Несомненно, предпочтения

должны отдаваться решениям, которые опираются на международные стандарты, на технологии, которые прошли проверку временем. Технология Z39.50 здесь является одним из самых привлекательных вариантов.

### Распределенные информационные системы

На наш взгляд, информационную систему можно считать распределенной, если она удовлетворяет, как минимум, следующим требованиям:

- информация хранится на физически различных серверах;
- существует возможность сквозного поиска информации одновременно на всех серверах;
- существует возможность извлечения найденной информации.

Это, безусловно, самые минимальные требования, без выполнения которых информационную систему нельзя считать распределенной. Интересно то, что указанным выше критериям удовлетворяют сегодня следующие системы:

Корпоративная Библиотечная Система Новосибирска (<http://z3950-2.uiggm.nsc.ru/zgwc/>);

Корпоративная Сеть публичных библиотек Москвы (<http://corporate.gpntb.ru/>);

Корпоративная Библиотечная Система Нижнего Новгорода (<http://www1.nounb.sci-nnov.ru/gateway.html>);

Распределенная Информационная Система СО РАН (<http://z3950-2.uiggm.nsc.ru/zgwk/>);

Распределенная Информационная Система LIBWEB (<http://ben.irex.ru/Zgw>, <http://www.cnshb.ru/zgw/>);

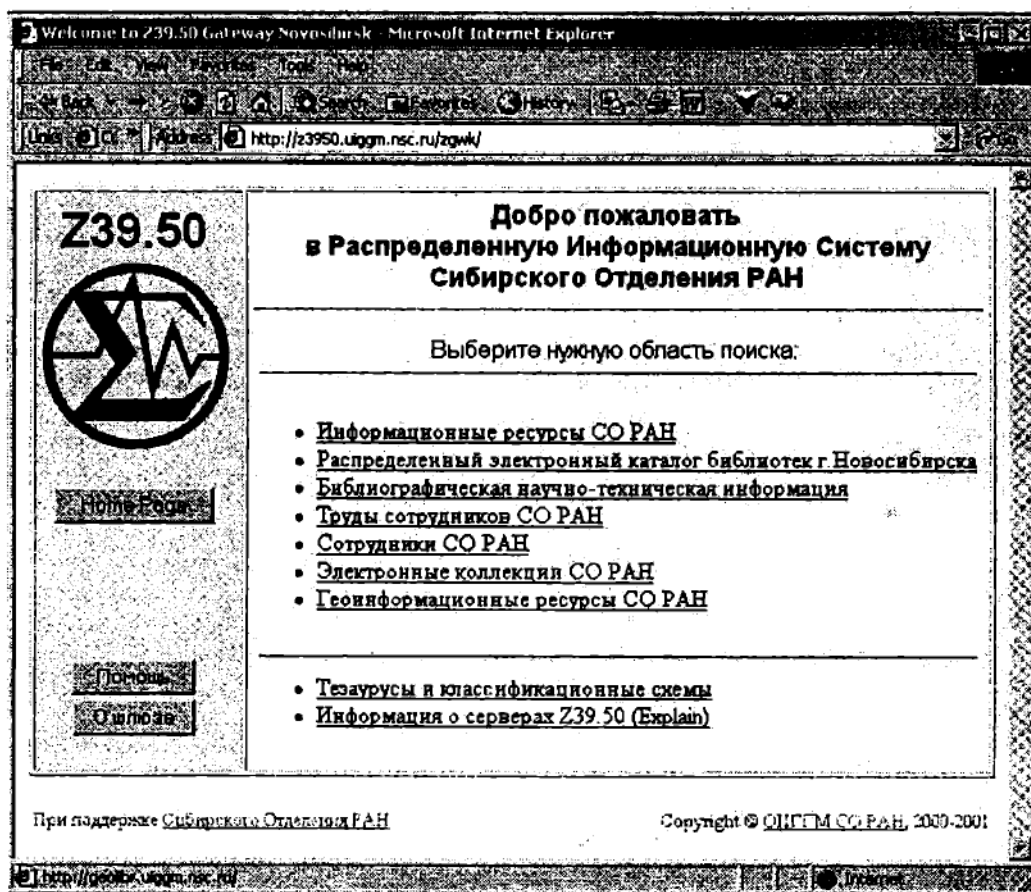


Рис. 1. Доступ к распределенной информационной системе СО РАН через шлюз Z-CGI

Все остальные информационные системы, в том числе библиотечные, распределенными не являются.

Поскольку технологии Z39.50 предоставляют прекрасную возможность для построения распределенных информационных систем, мы надеемся, что интеграция информационных ресурсов будет продолжаться. При этом направлением интеграции будет не столько создание огромных хранилищ данных, сколько объединение в рамках единой информационной системы множества различных серверов, рассредоточенных по всей России и, может быть, по всему миру.

В заключение следует заметить, что протокол Z39.50 в последние годы, благодаря различным программам (ИОО, РФФИ) находит все большее применение в информационно-библиотечном сообществе

России при построении распределенных библиотечных систем. Однако уровень использования его в этих системах в основном ограничен лишь областью поиска и извлечения информации из электронных каталогов, что резко сужает возможности применения протокола. Это обуславливается, на наш взгляд, узкой направленностью финансирования программ в сторону библиотек.

Дальнейшее развитие и внедрение новых информационных технологий в других областях (музеи, архивы, электронные коллекции и др.) и дополнительное финансирование могут привести к расширению использования технологий Z39.50. Для этого есть все предпосылки.

### Литература

1. 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.
2. V. Baranov, A. Plennek, N. Sokolova et al. Review of Z39.50 servers and Z39.50 environment in Russia//Library Hi Tech.-2000.-Vol. 18.-N. 4.-p. 304-314. (<http://www.unilib.nova.ru/olsc/publications/lht2000.html>)
3. Жижимов О.Л., Коджесян В.С., Мазов Н.А. Пример распределенной информационной системы на основе метаданных и международных стандартов // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Вторая Всероссийская научная конференция, Протвино, 26-28 сентября 2000 г.: Сб. докл., Протвино, ГНЦ ИФВЗ, 2000, 102-106, ISBN 5-88738-029-2.
4. Mike Taylor. Zthes: A Z39.50 Profile for Thesaurus Navigation. Version 0.3b. <http://lcweb.loc.gov/z3950/agency/profiles/zthes-03>
5. Жижимов О.Л., Мазов Н.А. Тезаурусы и классификационные схемы в распределенных информационных системах: проблемы и решения. Международная конференция EVA-2000 «Электронные изображения и визуальные искусства», Москва, Государственная Третьяковская галерея, 30 октября — 3 ноября 2000 г. (<http://www.artinfo.ru/cva/EVA2000M/cva-papers/200008/Zhizhimov2-R.htm>)
6. Z39.50 Application Profile for Geospatial Metadata — Geo Profile; Version 2.2; 1999-04-19 (<http://www.blueangeltch.com/standards/GeoProfile/geo22.htm>).
7. The CIMI Profile Release 1 .OH A Z39.50 Profile for Cultural Heritage Information [http://www.cimi.org/public\\_docs/HarmonizedProfile/HarmonProfile1.htm](http://www.cimi.org/public_docs/HarmonizedProfile/HarmonProfile1.htm)
8. Мазов Н.А., Жижимов О.Л. Применение протокола Z39.50 для работы с тезаурусами и классификационными схемами. Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: 7-я междунар. конф. «Крым 2000» (3-11 июня 2000 г., г.Судак): Материалы конф., Т. 1, Симферополь, Таврида, 2000, С. 402-405
9. (<http://www.artinfo.ru/cva/EVA2000M/eva-papers/200008/Zhizhimov-R.htm>).
10. Жижимов О.Л., Мазов Н.А. Модель распределенной информационной системы Сибирского Отделения РАН на базе протокола Z39.50. Электронные библиотеки, 1999, т. 2, вып. 2, ISSN 1234-5678.