

Российская академия наук
Сибирское отделение
Объединенный институт геологии, геофизики и минералогии
им. академика А.А.Трофимука

Жижимов О.Л., Мазов Н.А.

**Принципы построения
распределенных информационных
систем
на основе протокола Z39.50**

Новосибирск - 2004

УДК 681.324
ББК 32.973.202
Ж70

Жижимов О.Л., Мазов Н.А.

Ж70 Принципы построения распределенных информационных систем на основе протокола Z39.50. - ОИГГМ СО РАН, Новосибирск: Изд во ИВТ СО РАН, 2004. - 361 с.: ил., табл.

ISBN 5-9554-0017-6

doi: 10.18303/B5-9554-0017-6

В книге рассматриваются различные аспекты применения протокола *Z39.50 (ISO 23950)* для построения распределенных информационных систем (РИС). Утверждается, что технологии, основанные на использовании протокола *Z39.50*, являются наиболее пригодными для организации доступа к разнородным базам данных в РИС. Показывается, что, несмотря на некоторые недостатки, технологии *Z39.50* с успехом могут быть использованы для создания эффективных РИС. Приводится список необходимых расширений *Z39.50* для РИС. В книге детально описывается серверное программное обеспечение *ZooPARK*, предназначенное для работы с различными базами данных в РИС. Монография снабжена обширным иллюстративным и табличным материалом, а также большим списком литературы по рассматриваемой проблеме.

Книга содержит не только теоретические разделы, но и существенную практическую составляющую, базирующуюся на собственном опыте и программных разработках авторов. Она представляет интерес для специалистов, работающих над созданием распределенных информационных систем различного назначения, для преподавателей, аспирантов и студентов образовательных учреждений информационного профиля.

УДК 681.324
ББК 32.973.202

ISBN 5-9554-0017-6

© Жижимов О.Л., Мазов Н.А., 2004
© ОИГГМ СО РАН, 2004

Оглавление

Предисловие ответственного редактора	9
Предисловие	13
1 Введение	15
1.1 История протокола <i>Z39.50</i>	15
1.2 Область действия протокола <i>Z39.50</i>	16
1.3 Модель взаимодействия	16
1.4 Сеанс связи	19
1.5 Модель поиска	20
1.6 Модель извлечения данных	21
1.7 Дополнительные функции	21
1.8 Сервис <i>Explain</i>	23
2 <i>Z39.50</i> как основа РИС	25
2.1 Что такое РИС	25
2.2 Требования к РИС	26
2.3 Технологии построения РИС	27
2.4 <i>Z39.50</i> как технология РИС	29
2.5 Недостатки <i>Z39.50</i> как технологии РИС	29
2.6 Расширения <i>Z39.50</i> для построения РИС	30
2.6.1 Локальные идентификаторы объектов	31
2.6.2 Использование специальных полей в <i>APDU</i>	32
2.6.3 Использование внешних определений	32
2.7 <i>CIP</i> – попытка модернизации <i>Z39.50</i> для РИС	32
2.8 Выводы	34
3 Архитектура РИС на основе <i>Z39.50</i>	35
3.1 Основные типы ИС	35
3.2 Основные блоки РИС на основе <i>Z39.50</i>	37
3.3 Способы и механизмы ”распределенности”	40
3.3.1 Объединение баз данных	40
3.3.2 Переадресация запросов	41
3.3.3 Контроль маршрутизации запросов	42
3.3.4 Построение распределенной БД <i>Explain</i>	42
3.4 Оптимизация распределенных запросов	43
3.5 Администрирование распределенных систем	46
3.5.1 Пополнение, модификация и удаление информации	46
3.5.2 Учет ресурсов РИС	47

3.5.3	Сбор статистической информации о работе РИС	47
3.5.4	Аутентификация пользователей РИС	48
3.5.5	Ограничение доступа к ресурсам РИС	49
3.5.6	Мониторинг целостности РИС	49
3.6	Выводы	49
4	Состав РИС на основе Z39.50	51
4.1	Метаданные как основа РИС	51
4.2	Особенности метаданных в Z39.50	52
4.3	Типы метаданных	52
4.3.1	MARC	52
4.3.2	GILS	54
4.3.3	Dublin Core	55
4.3.4	GEO	56
4.3.5	Digital Collections	57
4.3.6	CIMI	59
4.4	Тезаурусы и рубрикаторы в РИС	60
4.4.1	Архитектура клиент-сервер	61
4.4.2	Тезаурусы и рубрикаторы в технологии Z39.50	61
4.5	Z+SQL – доступ к реляционным СУБД	63
4.6	Библиографические информационные системы	65
4.6.1	Специфика РИБС	66
4.6.2	О программном обеспечении	66
4.6.3	Модели интеграции ресурсов	69
4.7	Доступ к первичной информации	69
5	Сервер ZooPARK в РИС	71
5.1	История создания сервера ZooPARK	71
5.2	Основные характеристики	72
5.3	Архитектура сервера ZooPARK	72
5.4	Модульность и расширяемость	73
5.5	Расширения Z39.50 сервера ZooPARK	75
5.6	Обеспечение "распределенности" сервером ZooPARK	76
5.6.1	Объединение баз данных	76
5.6.2	Переадресация запросов	76
5.6.3	Контроль маршрутизации запросов	76
5.6.4	Модификация информации	76
5.6.5	Учет ресурсов РИС	77
5.6.6	Сбор статистической информации о работе РИС	77
5.6.7	Аутентификация пользователей РИС	78
5.6.8	Ограничение доступа к ресурсам РИС	78
5.6.9	Мониторинг целостности РИС	78
5.7	Оптимизация распределенных запросов	78
5.7.1	Распараллеливание поисковых запросов	78
5.7.2	Обеспечение режимов переадресации	80
5.7.3	Контроль петель переадресации	80
5.8	Работа сервера ZooPARK в РИС	80
5.9	Сервер ZooPARK и метаданные	81
5.10	Поддержка Z-SQL и SQL-RS	81
5.11	Сервер ZooPARK и другие разработки	84
5.12	Выводы	86

6	Установка и настройка сервера <i>ZooPARK</i>	87
6.1	Как установить сервер	87
6.1.1	Состав дистрибутива	87
6.1.2	Установка сервера для <i>Windows NT</i>	89
6.1.3	Установка сервера для <i>UNIX</i>	90
6.1.4	Настройка конфигурационных файлов	91
6.1.5	Подготовка тестовых баз данных	92
6.1.6	Запуск сервера	92
6.1.7	Проверка работоспособности сервера <i>Z39.50</i>	93
6.1.8	Проверка работоспособности сервера <i>WEB</i>	93
6.1.9	Проверка <i>PHP</i> -модуля сервера <i>WEB</i>	93
6.1.10	Проверка <i>PERL</i> -модуля сервера <i>WEB</i>	94
6.1.11	Проверка сервера <i>SRW/SRU</i>	95
6.1.12	Проверка шлюза <i>Z39.50-HTTP</i>	95
6.2	Конфигурирование и настройка	96
6.2.1	Конфигурационный файл сервера	96
6.2.2	Ограничение прав доступа	99
6.2.3	Аутентификация	99
6.2.4	Настройка провайдеров данных	101
6.2.5	Параметры запуска сервера	123
6.2.6	Сбор статистической информации	124
6.3	<i>Explain</i>	128
6.3.1	Категории	129
6.3.2	Как создать базу данных <i>IR-Explain-1</i>	129
6.3.3	Пример базы данных	130
6.3.4	Категории <i>Explain</i> для сервера <i>ZooPARK</i>	132
6.3.5	Автоматически заполняемые поля	132
6.4	<i>WEB</i> сервер <i>ZooPARK</i>	133
6.4.1	Настройка <i>WEB</i> -сервера	133
6.4.2	Настройка внешних модулей <i>WEB</i> -сервера	134
6.5	Встроенный шлюз <i>Z-GW</i>	134
6.5.1	Основные характеристики шлюза <i>Z-GW</i>	135
6.5.2	Настройка шлюза <i>Z-GW</i>	136
6.5.3	Интерпретатор	139
6.6	Построение распределенных систем	144
6.6.1	Пример простой распределенной системы	144
6.6.2	Оптимизация распределенных систем	145
6.6.3	Некоторые рекомендации	148
7	Эксплуатация <i>ZooPARK</i> в РИС	149
7.1	Информационная система ОИГТМ СО РАН	149
7.2	Информационная система СО РАН	155
7.3	Информационная система <i>LibWeb</i>	159
7.4	Информационные системы корпоративных проектов ИОО	163
7.4.1	Корпоративная библиотечная РИС Новосибирска	164
7.4.2	Корпоративная библиотечная РИС Москва	166
7.4.3	Корпоративная библиотечная РИС Ярославль	168
7.4.4	Корпоративная библиотечная РИС Нижний Новгород	169
7.4.5	Корпоративная библиотечная РИС Омск	169
7.5	Информационная система по газгидратам	174
7.5.1	Особенности проекта	174

7.5.2	Структура информационной системы	175
7.5.3	<i>GILS</i> как основа информационной системы	177
7.5.4	Пространственная привязка и интерфейсы пользователя	179
7.5.5	Рубрикатор	180
8	Проекты и решения	183
8.1	Библиотечные проекты	183
8.1.1	<i>ZLOT</i>	188
8.1.2	<i>UNiverse</i>	188
8.1.3	<i>АРБИКОН</i>	192
8.2	<i>GILS</i>	194
8.3	Электронные библиотеки	196
8.3.1	<i>LAURIN</i>	196
8.3.2	<i>DNER</i>	197
8.3.3	<i>ELISE II</i>	203
8.3.4	<i>EULER</i>	206
8.4	Пректы по культурному наследию	209
8.4.1	<i>Aquarelle</i>	209
8.4.2	<i>Zavier</i>	209
8.4.3	<i>AHDS</i>	212
8.4.4	<i>ARTISTE</i>	213
8.5	Проекты по естественным наукам	214
8.5.1	<i>Species Analyst</i>	215
8.5.2	<i>FishNet</i>	216
8.5.3	<i>MaNIS</i>	217
8.5.4	<i>HerpNet</i>	218
8.6	Проекты по ГИС	218
	Заключение	223
A	Наборы атрибутов	259
A.1	Атрибуты Bib-1	259
A.1.1	Use attributes	259
A.1.2	Relation attributes	276
A.1.3	Position attributes	277
A.1.4	Structure attributes	277
A.1.5	Truncation attributes	278
A.1.6	Completeness attributes	279
A.2	Атрибуты Exp-1	279
A.2.1	Use attributes	279
A.3	Атрибуты GILS	280
A.3.1	Use attributes	280
A.4	Атрибуты Ext-1	282
A.4.1	Use attributes	282
A.4.2	Permission attributes	282
A.5	Атрибуты XD-1	283
A.5.1	Access Point (Use) attributes	283
A.6	Атрибуты Utility	283
A.6.1	Access Point (Use) attributes	283
A.7	Атрибуты Zthes-1	286
A.7.1	Access Point (Use) attributes	286
A.8	Атрибуты Collections-1	286

A.8.1	Access Point (Use) attributes	286
A.9	Атрибуты CIMI-1	287
A.9.1	Access Point (Use) attributes	287
A.9.2	Authority (Type-101) attributes	288
A.9.3	Charset (Type-103) attributes	290
A.10	Атрибуты GEO-1	292
A.10.1	Access Point (Use) attributes	292
B	Наборы меток (tagSet)	329
B.1	Metadata Tag Set (tagSet-M)	329
B.2	Generic Tag Set (tagSet-G)	330
B.3	GILS Tag Set	331
B.4	Zthes Tag Set	332
B.4.1	Collections Tag Set	333
B.5	CIMI Tag Set	334
B.6	UIGGM Tag Set	335
C	Схемы данных	339
C.1	Схема GILS	339
C.2	Схема Zthes	342
C.3	Схема Collections	342
C.4	Схема CIMI	346
C.5	Схема UIGGM	350
C.6	Схема GEO	352
D	Идентификаторы объектов сервера ZooPARK	357

Предисловие ответственного редактора

Предлагаемая Вашему вниманию монография посвящена описанию принципов построения распределенных информационных систем на основе протокола Z39.50.

Широкое использование современных средств доступа к информационно-вычислительным ресурсам открыло принципиально новые возможности использования информационных технологий и вычислительной техники. Современные информационные технологии, позволяющие создавать, хранить, перерабатывать и обеспечивать эффективные способы представления информационных ресурсов потребителю, стали важным фактором жизни общества и средством повышения эффективности управления всеми сферами общественной деятельности. Информация является важнейшим стратегическим ресурсом и наибольший экономический и социальный успех сегодня сопутствует тем странам, которые активно используют современные средства компьютерных коммуникаций и сетей, информационных технологий и систем управления информационными ресурсами.

Стремительное развитие глобальных информационных и вычислительных сетей привело к изменению фундаментальной парадигмы обработки данных, направление которой переместилось на использование распределенных информационно-вычислительных ресурсов и поддержку инфраструктуры для свободного доступа к ним. С одной стороны, наблюдается переход к исключительно распределенной схеме создания, поддержания и хранения ресурсов, а с другой — стремление к виртуальному единству посредством предоставления свободного доступа к любым ресурсам сети через ограниченное число «точек доступа». Таким образом, в современном информационном обществе на первое место выходят технологии использования распределенных информационно-вычислительных ресурсов. В западной литературе несколько лет назад появился даже новый термин «GRID-технологии» — технологии создания и использования распределенных информационно-вычислительных ресурсов.

На самом деле идея использования распределенных ресурсов далеко не нова. Она была сформулирована еще на заре компьютерной эры - здесь можно вспомнить Вэннивера Буша¹, который в своем труде “As We May Think” (1945) описал концепцию гипертекста. В дальнейшем идея распределенной обработки информации стала активно развиваться с появлением первых суперЭВМ второго

¹Вэннивер Буш (Vannevar Bush) (1890-1974) — основатель Национального Научного Фонда (NSF) США, создатель “дифференциального анализатора” (1930) — релейной машины, способной решать дифференциальные уравнения, руководитель работ по созданию первых ЭВМ.

поколения — это советская ЭВМ БЭСМ-6 и американская ILLIAC-IV. Именно на этих машинах впервые были поставлены эксперименты и создано соответствующее программное обеспечение по использованию распределенных вычислительных ресурсов. В дальнейшем технология массивового счета была перенесена и на “mainframe” третьего поколения². Недаром одним из лозунгов создателей технологий GRID является следующий: “*Forward-Back to Mainframe*” («Вперед (назад) к майнфреймам»).

Второй этап развития технологий использования распределенных ресурсов можно связать с появлением WWW-сервиса в сети Интернет. В 1989 году в CERN’e перед сотрудниками лаборатории ECP (Electronics & Computing for Physics) была поставлена задача разработать систему для обеспечения (и унификации) доступа к любым данным, содержащимся в сети, и создать универсальную технологию доступа к распределенной, разнородной информации. Таким образом, был создан WWW-сервис сети Интернет³. Дальнейшее развитие технологии WWW немного подкорректировало начальные установки, и из системы доступа к разнородным ресурсам он превратилась во всемирную “презентационную” систему.

Развитие концепций предоставления удобного и разветвленного доступа к информационно-вычислительным ресурсам, распределенным по сети Интернет, привело к созданию соответствующих протоколов и стандартов, определяющих основные механизмы информационного обмена (протоколы и стандарты OSI⁴). Сервис сети Интернет WAIS, замыкающий триаду технологических средств, которую возглавляют Gopher и WWW, принципиально обеспечивает функционально полный набор информационных инструментов поиска для Internet. Если Gopher (или LDAP) — это виртуальная файловая система, где в качестве элементов системы могут использоваться не только файлы различных форматов и каталоги, но и виртуальные объекты в виде поисковых критериев, а WWW — распределенная мультимедийная система, то WAIS реализует концепцию распределенной информационно-поисковой системы и распределенную систему доступа к ресурсам. Система базируется на универсальном протоколе, в основе которого лежит американский стандарт Z39.50.

Стандарт Z39.50 — это один из протоколов семейства OSI, который описывает прикладной уровень взаимодействия распределенных информационно-поисковых систем (ИПС)⁵. Протокол определяет механизм информационного обмена в процессе обработки поисковых запросов и протокол обмена данными в системах, осуществляющих поиск. Основная область применения протокола в настоящее время — это библиотечные системы и системы научно-технической

²Следует отметить, Россия (тогда СССР) в то время была страной наиболее мощного класса вычислительных систем — mainframe. За все время производства советских компьютеров типа БЭСМ-6 и ЕС ЭВМ, являющихся клоном системы IBM 360/370, на предприятия и в организации Советского Союза было поставлено десятки тысяч таких машин. Причем причиной их массового использования было не только и не столько отсутствие производства и закупок персональных компьютеров и UNIX или Wintel серверов, а необходимость решения вычислительных задач «глобального» характера и массового производственного счета. Даже при наличии последних — задачи, которые предполагалось решать с помощью вычислительной техники в СССР, были неподъемны для других платформ.

³WWW — технология “World Wide Web” (“Всемирная паутина”).

⁴OSI — Open System Interconnection — Взаимодействие Открытых Систем.

⁵В отечественной литературе за термином “Information Retrieval System” закреплено понятие “Информационно-поисковая система”. Стандарт Z39.50 посвящен организации механизма поиска в компьютерных сетях распределенных информационных ресурсов с использованием ключевых слов.

информации. Однако область применения протокола значительно шире перечисленных приложений — он может использоваться в информационно-поисковых системах общего назначения. При разработке протокола подразумевалось, что он будет описывать порядок обмена информацией между пользователями информационной системы и ее ядром через сеть передачи данных. При этом сами системы могут управлять данными, используя разные модели и различные языки манипулирования этими данными, будь это обычная файловая система или объектно-ориентированная СУБД.

Идея совместного использования информационно-вычислительных ресурсов была поддержана большим количеством фирм-производителей программного обеспечения. На сегодняшний день для разработки распределенных информационных систем было предложено большое количество различных технологий (например, RPC, DCOM, RMI, ODBC, JDBC, CORBA, SOAP, eXML, WSDL, WSFL, UDDI и др.). Однако большая часть этих технологий по своей сути ориентирована не на работу с информационными ресурсами, а на сетевое взаимодействие программ и распределенные вычисления в гетерогенных средах. Однако ни одна из упомянутых технологий не обеспечивает универсальных способов работы с информационными ресурсами, т.е. необходимого уровня абстрагирования от конкретных систем и платформ при сохранении высокой степени функциональности и предоставления механизма доступа к данным. Эти требования присущи технологии, в основе которой лежит протокол ANSI Z39.50 (стандарт ISO-23950).

Приведенные выше идеи создания единой системы доступа к информационно-вычислительным ресурсам у нас в Сибирском отделении сразу легли на подготовленную почву. Несмотря на некоторое затишье, вызванное годами перестройки, уже в середине девяностых годов в СО РАН стартовал ряд проектов по созданию и поддержке распределенных информационно-вычислительных ресурсов. В качестве первых из них следует упомянуть *“Электронный атлас биоразнообразия животного и растительного мира Сибири”* и *“Распределенный каталог библиотек Сибири”*. В качестве одного из важнейших результатов этих работ стала разработка и реализация проекта создания *“Электронной библиотеки Сибирского отделения РАН”*, который объединил сотрудников большого числа институтов СО РАН (ИБТ, ОИГГМ, ЦСБС, ИЦГ, ИК, ИЭОПП, ИВМ, ИДСТУ, ИОА, ГПНТБ и др.), включая авторов данной монографии.

Дальнейшее развитие работ по созданию *“Электронной библиотеки Сибирского отделения РАН”* привело к необходимости виртуальной интеграции создаваемых ресурсов в единую унифицированную систему: большая часть компонентов этой системы находится в разных местах и при выполнении иных задач может функционировать независимо, интероперабельность (унифицированность) достигается использованием согласованного набора стандартов (в том числе корпоративных), протоколов и сервисов, а виртуальное единство обеспечивается, за счет единого интерфейса и стандартизации структуры на основе единой схемы метаданных.

Создание интегрированной распределенной информационной системы СО РАН (ИРИС) опирается на идею электронных (цифровых) библиотек. Основная задача — это формирование в ресурсах “Сети” единого, математически однородного поля компьютерной информации, способного стать универсальным и машинезависимым носителем данных и глобально распределенных информационных ресурсов. В рамках этого подхода цифровые библиотеки рассматриваются как отдельная конкретная технология работы с информацией.

В основу создания ИРИС положен принцип информационных хранилищ,

с учетом поддержки уже функционирующих технологий, технологических решений используют принципы обработки метаописаний данных на основе серверов метаданных и протокола Z39.50. Разработана технология, которая обеспечивает виртуальную интеграцию разнородных информационных и вычислительных ресурсов, расположенных на серверах различных организаций, в единую систему с унифицированным доступом на основе открытых международных стандартов. Это позволит объединить в единое информационное пространство уже существующие многочисленные серверы организаций, входящих в научно-образовательную сеть, основанные на стандартных протоколах, организовать информационное обеспечение проведения исследований по фундаментальным и прикладным направлениям, проводимым в вузах, академических институтах.

По существу эта монография представляет собой первую попытку в России систематизировать и обобщить последние достижения мировой и отечественной мысли и практики в области основных принципов построения распределенных информационных систем, создании специального серверного программного обеспечения для доступа к распределенным информационным ресурсам с использованием метаданных и схем данных, основанных на открытых международных стандартах и протоколах, которые позволяли бы унифицировать сетевой доступ к любым базам данных, используя абстрагирование от структур конкретных баз данных и СУБД.

Данная монография позволяет на высоком профессиональном и в то же время доступном уровне ознакомиться с одним из старейших протоколов Z39.50, ставшим стандартом и успешно эксплуатирующемся уже многие годы за рубежом, а в последние годы, благодаря различным корпоративным проектам, и в России. Здесь отдельно необходимо отметить предложенную авторами совокупность требований к распределенной информационной системе, которые, на мой взгляд, являются универсальными, независимо от используемого в системе информационного ресурса.

Отдельный раздел в монографии посвящен разработке и реализации специального серверного программного обеспечения, способного обеспечить устойчивое функционирование распределенной информационной системы. Следует отметить, что разработанное авторами программное обеспечение для функционирования распределенных информационных систем эффективно эксплуатируется в настоящее время в Российском информационном сообществе, а для ряда проектов и программ является типовым программно-технологическим решением.

Глубокое знание информационных технологий и большой практический опыт авторов позволил построить одну из лучших в России распределенных информационных систем, о которой говорится в монографии. Большую ценность представляет также анализ эффективности применения различных конфигураций распределенных информационных систем. Авторами монографии дается достаточно обстоятельный обзор применения технологий Z39.50 в мировом информационном сообществе, различной видовой направленности.

Надеюсь, что научная информационная общественность и специалисты в области информационных технологий получат хорошую монографию по вопросам построения распределенных информационных систем и интеграции разнородных информационных ресурсов, сильной стороной которой является удачное сочетание глубокого рассмотрения теоретических вопросов с практикой и новыми возможностями информационных технологий и оборудования.

А.М.Федотов,
член-корресподент РАН

Предисловие

Сегодняшний мир переполнен информацией. Миллионы *WWW*-серверов всегда готовы поделиться своей информацией с пользователями сети Интернет. Можно найти все, что угодно. Однако, существует утверждение, что доступная сегодня через *WEB* информация составляет лишь единицы процентов от той, которой обладает человечество в машиночитаемом виде.

Основные информационные ресурсы находятся в базах данных, которые недоступны напрямую через *WEB*. Это различные реферативные и полнотекстовые библиографические базы данных, музейные коллекции, результаты наблюдений и экспериментов, цифровые фото-, аудио-, видео коллекции и многое-многое другое. Спектр этих ресурсов широк, но их объединяет одно свойство – все эти ресурсы упорядочены, хотя упорядочены по-разному.

Для обеспечения доступа к подобным ресурсам нужны специальные системы, для обеспечения унифицированного доступа к ресурсам, разнесенным географически, и управляемыми различными СУБД нужны распределенные информационные системы с реализацией сквозного поиска.

Итак, обеспечение универсальных способов работы с распределенными и разнородными данными, где заранее неизвестно, с какими видами объектов придется работать конечному пользователю, унификация представления этих данных – является главной задачей при интеграции информационных ресурсов в распределенных информационных системах.

Наряду с вышесказанным, увеличение производительности глобальных сетей на основе новых телекоммуникационных технологий, применение мощных каналов связи и стандартизация различных сетевых протоколов, сделало возможным построение реально действующих распределенных информационных систем, объединяющих разнородные электронные ресурсы многих локальных подсистем.

В появлении таких систем заинтересованы как владельцы электронных информационных ресурсов, например, библиотеки и информационные центры, музеи и архивы, создатели информационных порталов в Интернет и др., ведущие активные работы по предоставлению своих ресурсов конечным пользователям, так и различные сообщества, в том числе межведомственные, пытающиеся интегрировать разнородные ресурсы в рамках своих информационных систем.

Нельзя сказать, что человечество в этом направлении ничего не делает. Существует много проектов, основная цель которых – построение распределенных информационных систем, создание и внедрение технологий доступа к распределенным информационным ресурсам (коллекциям, базам данных и т.п.). В России также проявляются тенденции общего развития. Так, например, крупнейшими российскими корпоративными информационно-библиотечными образованиями (*LibWeb*, *ЛИБНЕТ*, *АРБИКОН*) предпринята успешная попытка создания

информационно-библиотечных распределенных систем, направленных на интеграцию ресурсов и технологий в библиотечной сфере.

Примечательно, что библиотечное сообщество одно из первых в России, сумело договориться о применении внутрикорпоративных правил обмена данными, на основе Российского коммуникативного формата *RUSmarc* и протокола *Z39.50*, что фактически позволило перейти к общей систематике управления информационно-библиотечными ресурсами на государственном уровне.

Предлагаемая Вашему вниманию книга посвящена принципам построения распределенных информационных систем на основе протокола *Z39.50 (ISO-23950 – ANSI/NISO-Z39.50)* с целью интеграции информационных ресурсов и унификации доступа к ним конечных пользователей.

В Введении вкратце изложены основы протокола *Z39.50*, далее, в первых трех главах, обсуждаются недостатки базового *Z39.50* с точки зрения построения распределенных систем, описываются возможные архитектуры и состав распределенных информационных систем на основе протокола *Z39.50*. Следующие три главы посвящены серверному программному обеспечению для работы по протоколу *Z39.50* в распределенных информационных системах – *ZooPARK*, вопросам его установки, настройки и эффективной эксплуатации. Последняя глава содержит обзор некоторых проектов по созданию информационных систем, в которых используется протокол *Z39.50*. Материал книги снабжен необходимым иллюстративным и справочно-ссылочным материалом.

Источником материалов для книги послужили многочисленные публикации авторов и дискуссии с коллегами на различных форумах по настоящей тематике, статьи, обзоры и монографии ведущих зарубежных и отечественных специалистов в области систем баз данных, различная специальная документация и информация, полученная из сети Интернет.

Мы прекрасно понимаем и отдаем себе отчет в том, насколько многогранна и обширна затронутая тематика, которая, конечно, не исчерпывается материалом, изложенным в настоящем издании.

Мы также выражаем надежду, что настоящая книга, отражающая результат многолетней работы с различными базами данных, создания эффективного программного обеспечения и систем для доступа к информационным ресурсам, будет Вам полезна и поможет Вам эффективно решать возникающие задачи при построении и эксплуатации распределенных информационных систем.

Мы выражаем благодарность всем, кто имел прямое или косвенное отношение к созданию этой книги. Особую благодарность мы хотим выразить нашим коллегам, с которыми чаще всего приходилось обсуждать проблемы построения распределенных информационных систем – Б. С. Елепову, А.М. Федотову, С. Р. Баженову, Малицкому Н.А., С. В. Скибину, И. В. Бычкову, М. А. Аветисову, М. В. Гончарову, О. С. Колобову, А. И. Племнеку.

Мы благодарны всем, кто принимал и принимает участие во внедрении в повседневную практику нашего программного обеспечения. Особая благодарность – Я. Л. Шрайбергу, Б. И. Маршаку, А. И. Бродовскому, Ю. Е. Хохлову, Л. В. Левовой, С. А. Амельченко, В. И. Стеллецкому, В. Н. Смирнову, Л. Г. Еремееву, Л. В. Мершиевой, Е. В. Мозину, Е. В. Ковязиной и многим-многим другим.

Глава 1

Введение

Прежде, чем обсуждать различные аспекты построения распределенных информационных систем, мы считаем необходимым привести некоторые базовые сведения о протоколе *Z39.50* [1].

1.1 История протокола *Z39.50*

Первая версия протокола, получившего наименование *Z39.50*, была подготовлена Комитетом по Национальным Информационным Стандартам США – *NISO* (*National Information Standards Organization*) и утверждена в 1988 г. как стандарт *Z39.50-1988* (см. [2]). Его действие распространялось на работу только с библиографической информацией. Именно эта версия протокола дала толчок к появлению известной службы Интернет *WAIS*.

В 1989 г. было организовано Агентство поддержки протокола *Z39.50* (*Maintenance Agency Z39.50*) под административным управлением лаборатории библиотеки Конгресса США, а в 1990 г. сформирована Группа исполнителей *Z39.50* – *ZIG* (*Z39.50 Implementors Group*). Ее членами стали производители, продавцы, консультанты, распространители разнородных видов информации (в т.ч. библиографической, текстовой, графической, финансовой, общественного назначения, химической и новостей), университетские сотрудники и т.д. Членство в Группе *ZIG* было открыто для всех заинтересованных сторон. Разработанная указанными организациями версия стандарта 1992 г. (*Z39.50-1992*) заменила собой стандарт 1988 г.

В ходе подготовки стандарта *Z39.50-1992* было сделано большое количество предложений для поддержки широкого диапазона возможностей информационного поиска. Однако реализация указанных предложений существенно задерживалась. Поэтому в качестве первоочередной задачи исполнителей была определена переработка версии 1988 г. для обеспечения его битовой совместимости с международными стандартами *ISO-10162/10163* (*SR* – *Search and Retrieve*). (*Z39.50-1992* заменил *Z39.50-1988* и *SR*).

В 1992 году Агентство провело формальный опрос среди разработчиков *Z39.50* с целью определить относительную значимость предложенных новых функций. Опрос должен был позволить сузить их список до разумных размеров, определить, адекватно ли формулируются и понимаются предложенные функции, и оценить их предполагаемую стоимость и сложность. Результаты опроса показали, что некоторые функции жизненно необходимы, а некоторые другие

можно полностью исключить из рассмотрения. По некоторым функциям опрос не дал однозначного результата, и по ним решение было принято на основании консенсуса.

Результаты этой работы учитывались в последующих версиях проекта стандарта. В апреле 1994 г. указанная работа была завершена выдачей черновой версии стандарта, которая получила временное наименование *Z39.50-1994*. Однако она не рассматривается как самостоятельная версия стандарта, а является исходной редакцией стандарта *Z39.50-1995*.

Версия 1992 года стала известной под названием "версия 2", а версия 1995 года – как "версия 3". Но, хотя эти обозначения версий имеют определенное протокольное значение, они никак не связаны с версиями стандарта. *Z39.50-1992* соответствует версии 2 протокола; *Z39.50-1995* соответствует версиям 2 и 3 протокола.

Хотя стандарт *Z39.50-1992* заменил *Z39.50-1988* (а стандарт *Z39.50-1988* устарел), отношение *Z39.50-1995* к *Z39.50-1992* совершенно иное: *Z39.50-1995* шире версии 1992 года и совместим с нею. В документе о *Z39.50-1995* можно найти полную информацию о версии 2 и создать на ее основе систему, совместимую с *Z39.50-1992*.

Указанное качество было достигнуто благодаря усилиям Агентства поддержки протокола *Z39.50-1995*, обеспечившего необходимые условия согласования обеих версий протоколов.

Агентство поддержки *Z39.50-1995* является постоянно действующим органом, производящим сопровождение эксплуатации протокола и его развитие. Сетевой адрес Агентства: <http://lcweb.loc.gov/z3950/agency>.

1.2 Область действия протокола *Z39.50*

Протокол *Z39.50* определяет порядок взаимодействия клиента и сервера, процедуры поиска и извлечения информации из баз данных, и форматы представления этой информации.

Протокол *Z39.50* не определяет форматы хранения данных в конкретных БД, способы их индексации и процедуры функционирования различных СУБД. Протокол *Z39.50* также не определяет интерфейсы взаимодействия пользователя и клиента.

В идеологии *Z39.50* в рамках одной схемы все базы данных совершенно одинаковы, несмотря на их физические различия по используемой СУБД, полям и синтаксису запросов. В *Z39.50* для клиента не существует возможности определить, под управлением какой СУБД хранятся извлекаемые им данные. Это, казалось бы, ограничение, имеет большой смысл, суть которого будет выяснена в следующих разделах. Однако можно отметить, что подобная информация клиенту не нужна, ибо он работает всегда с одной и той же системой запросов и получает данные в одних и тех же форматах.

1.3 Модель взаимодействия

Протокол *Z39.50* не определяет транспортный уровень сетевого взаимодействия. Поэтому возможны реализации *Z39.50* поверх любых транспортных протоколов (*TCP/IP*, *SPX/IPX*, *OSI* и т.д.).

Наибольшее значение имеют реализации *Z39.50* поверх *TCP/IP*, так как именно этот транспортный протокол принят в Интернет.

По умолчанию для запросов *Z39.50* предусматривается *TCP* порт 210. Допускается использование других номеров портов, как правило, непривилегированных. Поэтому обращение к серверу *Z39.50* должно содержать кроме *IP*-адреса или *DNS*-имени сервера дополнительно его номер *TCP* порта.

Протокол *Z39.50* описывает сетевое взаимодействие субъектов в архитектуре *клиент-сервер*. Однако это взаимодействие несколько отличается от классической архитектуры *клиент-сервер*, в которой инициатором любого запроса может быть только клиент, а серверу всегда отводится пассивная роль ожидающего и отвечающего. Как будет видно ниже, в *Z39.50* это не всегда так. Может быть поэтому разработчики протокола изменили терминологию, заменив термины *клиент* и *сервер* на термины *origin* и *target* соответственно. За редким исключением понятия *клиент* – *origin* и *сервер* – *target* совпадают.

Для описания логики сетевого взаимодействия в *Z39.50* стандартом определены следующие компоненты:

- Origin*** – компонента, инициирующая сеанс связи *Z39.50*;
- Target*** – компонента, ожидающая открытия сеанса *Z39.50*;
- Database user*** – пользователь данных;
- Database provider*** – поставщик данных;
- Service provider*** – компонента, принимающая сетевые запросы и ответы;
- Service user*** – компонента, использующая сетевые запросы и ответы;
- Client*** – приложение, включающее *origin* и *database user*;
- Server*** – приложение, включающее *target* и *database provider*.

В стандарте описаны четыре сервисных примитива, в терминах которых формулируются правила сервисных процедур *Z39.50*:

- Request*** (запрос) – примитив, используемый *origin* для инициализации его сервис-провайдером соответствующей сервисной процедуры;
- Indication*** (индикация) – примитив, используемый *target* для передачи от сервис-провайдера к сервис-потребителю;
- Response*** (ответ) – примитив, используемый *target* для инициализации его сервис-провайдером ответа;
- Confirmation*** (уведомление) – примитив, используемый *origin* для передачи от сервис-провайдера к сервис-потребителю.

Последовательность сервисной процедуры может быть проиллюстрирована на примере запроса *Search*, инициатором которого является *origin*:

1. *Origin* сервис-потребитель формирует для своего сервис-провайдера запрос *SearchRequest*;
2. От *origin* к *target* по сети пересылается *APDU SearchRequest*;
3. *Target* сервис-провайдер указывает своему сервис-потребителю о поступлении *SearchRequest*;
4. *Target* сервис-потребитель формирует *SearchResponse* для своего сервис-провайдера;
5. От *target* к *origin* пересылается *APDU SearchResponse*;
6. *Origin* сервис-провайдер уведомляет своего сервис-потребителя об *APDU SearchResponse*.

Обычно так выглядят все сервисные протокольные процедуры (*Init, Search, Present, Delete, Resource-report, Sort, Scan, Extended-services*). Однако есть и такие, в которых роли *target* и *origin* меняются местами. Инициатором процедур (*Access-control* и *Resource-control*) является *target*!

Как видно из описания протокольных процедур, по сети от *origin* к *target* и наоборот передаются специальные пакеты – *APDU* (*Application Protocol Data Unit*) для обмена информацией между субъектами. Каждой протокольной процедуре соответствуют свои *APDU*, структура которых жестко регламентирована.

В *Z39.50 APDU* принято описывать в синтаксисе *ASN.1* (*Abstract Syntax Notation*), который определен как международный стандарт *ISO-8824* [3,4].

ASN.1 позволяет определять передаваемое значение, не задавая конкретное его представление. Способ представления определяется заданием алгоритмов, называемых правилами кодирования. Правила кодирования определяют конкретный вид байтов на сеансовом уровне, с помощью которых передаются значения прикладного уровня, и позволяют получателю распознать переданную информацию как конкретное значение конкретного типа. Используемые в *Z39.50* правила кодирования *BER* будут рассмотрены ниже.

Основой *ASN.1* является теговая структура. Под тегом понимается некая структура данных определенного типа, поименованная совокупность значений. Имя тега однозначно идентифицирует тип содержащихся в нем данных, т.к. для правильной интерпретации конкретного представления значения необходимо знать тип передаваемых данных. Тег может быть определен в *ASN.1*, либо определяется пользователем.

ASN.1 является как средством определения сложных типов данных, так и средством задания конкретных значений этих типов без указания конкретного способа представления (в виде последовательности октетов) значений данного типа при их передаче. *ASN.1* позволяет описать любые многоуровневые структуры, включающие различные типы данных, а, следовательно, является универсальным средством для передачи пользовательской информации.

В настоящее время *ASN.1* используется в следующих приложениях:

- является основой для кодирования данных *RMON.1* и *RMON.2*, используемых в системах дистанционного управления сетевыми устройствами;
- в записях служб справочников *X.500* и *LDAP*;
- для кодирования протоколов *X.400* и *SNMP*;
- является основой для описания и кодирования *Z39.50*;
- и др.

Для передачи структур *ASN.1* по сети в виде последовательности байт с однозначным распознаванием на другой стороне необходимы дополнительные механизмы. Ситуация усугубляется еще и тем фактом, что компьютеры с разной архитектурой по-разному интерпретируют, например, целые числа, состоящие из нескольких байт (четыре байта для 32-х разрядных систем). Аппаратное представление целого числа различно для компьютеров с процессором Intel и компьютером с процессором SPARC. Это отличие, безусловно, нужно скрыть. Аналогичная ситуация и с числами с плавающей точкой.

Аппаратная независимость протокола *Z39.50* достигается установлением единых правил передачи октетов. Суть их такова:

1. Любое число из внутреннего представления HOST конвертируется во внешнее представление NET. Представление HOST соответствует внутреннему представлению числа, а представление NET – внешнему. Представление

NET совпадает с HOST-представлением числа для архитектуры SPARC, так как было придумано компанией *Sun Microsystems*. Конвертирование осуществляется вызовом стандартных функций типа *htonl*.

2. При передаче октета по сети старший бит приходит первым.

Для передачи структур *ASN.1* в последовательном потоке октетов используется *BER*-кодирование [5] (*BER – Basic Encoding Rules, ISO-8825-1*).

Итак, *Z39.50* содержит механизмы аппаратно-независимой передачи по сети сложных структур *ASN.1*. Однако для однозначной интерпретации содержания этих структур требуется дополнительная информация. Здесь вступает в силу стандартизация объектов.

Для однозначной идентификации всех объектов в *Z39.50* принята система присвоения каждому из них числового идентификатора. Справедливости ради следует отметить, что эта практика присуща не только *Z39.50*, а всем протоколам, описанных в терминах *ASN.1* (см. выше). В этом смысле сам *Z39.50* является объектом более общего класса. Иерархия классов следующая

iso (1) – member_body (2) – US (840) – ANSI-Z39.50 (10003)

что дает для *Z39.50* идентификатор *OID* (*object identifier*) 1.2.840.10003. В дальнейшем этот *OID* будет обозначаться как Z. Все объекты, подлежащие стандартизации, в *Z39.50* получают в своем *OID* префикс Z.

Классы объектов, подлежащих стандартизации в *Z39.50*, приведены в Табл. 1.1.

OID	Класс объектов
Z.1	application context definitions
Z.2	abstract syntax definition for APDUs
Z.3	attribute set definitions
Z.4	diagnostic definitions
Z.5	record syntax definitions
Z.6	transfer syntax definitions
Z.7	resource report format definitions
Z.8	access control format definitions
Z.9	Extended services definitions
Z.10	user information format definitions
Z.11	element specification format definitions
Z.12	Variant set definitions
Z.14	tag set definitions
Z.15	negotiation definitions
Z.16	query definitions

Таблица 1.1: Классы объектов *Z39.50*

1.4 Сеанс связи

Протокол *Z39.50* является протоколом, который ориентирован на постоянное соединение *origin-target*. Сеанс необходим для поддержки *target* специальных сеансовых переменных, которые создаются динамически при открытии сеанса и уничтожаются при его закрытии.

В сеансовых переменных *target* может хранить всяческую информацию, касающуюся текущего сеанса: историю запросов, настройки, сведения о пользователе и т.п. Именно в сеансовых переменных *target* сохраняет именованные наборы данных, доступные для использования при повторных запросах.

Общение *origin* с *target* вне сеанса возможно только при помощи *APDU initRequest*, предназначенного для установления сеанса.

Для закрытия сеанса по инициативе *origin* он должен послать *APDU close* и получить в ответ от *target* такой же *APDU*.

Следует особо отметить, что инициатором *APDU close* может быть не только *origin*, но и *target*. Эта особенность отличает *APDU close* от других. Причина этого проста: может возникнуть ситуация, когда сервер вынужден закрыть сеанс или все сеансы для корректной работы, например, при его выключении администратором.

1.5 Модель поиска

В *Z39.50* поисковые запросы всегда формулируются не к реальной базе данных, а к абстрактной. Эта абстрактная база данных не имеет никакой структуры и характеризуется только поисковыми атрибутами. Поэтому и принято ее называть "Набор атрибутов" (*attributeSet*).

Получив от *origin* запрос в виде термов и поисковых атрибутов, *target*, а точнее, *database provider*, конечно же, свяжет его с реальной базой данных и конвертирует его в реальный синтаксис, но это процедура для конечного пользователя останется незаметной. При таком подходе к процедуре поиска все базы данных становятся для пользователя одинаковыми, если поддерживают один и тот же набор поисковых атрибутов.

Наборы поисковых атрибутов составляют класс объектов *Z39.50 (Z.3)*, подлежащих стандартизации. В классе *Z.3* в настоящий момент стандартизованы наборы, представленные в Табл. 1.2.

OID	Набор	Комментарий
Z.3.1	bib-1	Библиографическая информация
Z.3.2	exp-1	<i>Explain</i>
Z.3.3	ext-1	Расширенный сервис
Z.3.4	ccl-1	<i>Type-2 (ISO-8777)</i> и <i>Type-100 (Z39.58)</i> запросы
Z.3.5	gils	Для поиска <i>GILS</i>
Z.3.6	stas	Научно-техническая информация
Z.3.7	collections-1	Навигация по электронным коллекциям
Z.3.8	cimi-1	Информация по музейным коллекциям
Z.3.9	geo-1	Пространственные метаданные
Z.3.10	ZBIG	Биологическая информация
Z.3.11	util	Утилиты
Z.3.12	xd-1	Междоменный набор (<i>Dublin Core</i>)
Z.3.13	<i>Zthes</i>	Навигация по тезаурусам

Таблица 1.2: Стандартные наборы поисковых атрибутов

Для поиска библиографической информации используется набор атрибутов *bib-1*. Он приведен в Приложении А.1. Некоторые наборы атрибутов (*gils*, *geo-1*)

включают в себя набор *bib-1*, поэтому этот набор является основным.

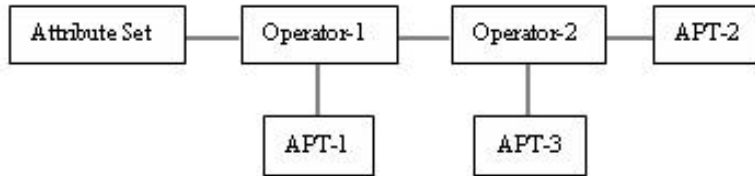


Рис. 1.1: Структура запроса *RPN*

В *Z39.50* для всех серверов обязательна поддержка запросов *type-1* (*RPN* – *Reverse Polish Notation* – обратная польская нотация). Поддержка других типов запросов серверами *Z39.50* является факультативной.

Запрос *RPN* можно представить в виде дерева, в узлах которого находятся связывающие операторы (*AND*, *OR*, *AND-NOT*). Листьями этого дерева являются блоки "атрибуты+терм" (*APT*).

В *Z39.50* запрос *RPN* не является строкой символов, а является структурой, которую можно изобразить в виде строки, но только для наглядности. Запрос *RPN* определен в виде структуры *ASN.1*, схематично представленной на Рис. 1.1.

1.6 Модель извлечения данных

Идеология максимального абстрагирования от структур реальных баз данных приводит к весьма изощренной схеме извлечения данных, описанной в стандарте *Z39.50*: записи из результирующих наборов отображаются в записи абстрактной базы данных через схему, определяющую абстрактную структуру записи в виде дерева элементов, специфицируемых метками (*tag*) из стандартных наборов (*tagSet*); затребованные элементы выбираются в нужной альтернативной форме (*variant*) из абстрактной записи и отображаются в экспортируемую структуру, определяемую форматом внешнего представления (*recordSyntax*). Все объекты описанной процедуры (*schema*, *tagSet*, *elementSpec*, *variantSet*, *recordSyntax*) определены в соответствующих классах с присвоением *OID*. Схематично модель извлечения записей представлена на Рис. 1.2.

1.7 Дополнительные функции

Дополнительные сервисы *Z39.50* не являются обязательными. Настройка на поддержку этих сервисов в сеансе связи происходит при инициализации сеанса через параметр *APDU options*.

Для представления об актуальности различных сервисов мы приводим информацию, снятую с *WEB*-сервера компании *IndexData* [6], которая показывает статистику поддержки различных сервисов в списке тестируемых серверов *Z39.50*. Распределение по поддерживаемым сервисам представлено в Табл. 1.3 (1 апреля 2001 года, 26 марта 2002 года, 2 марта 2003 года и 31 марта 2004 года).

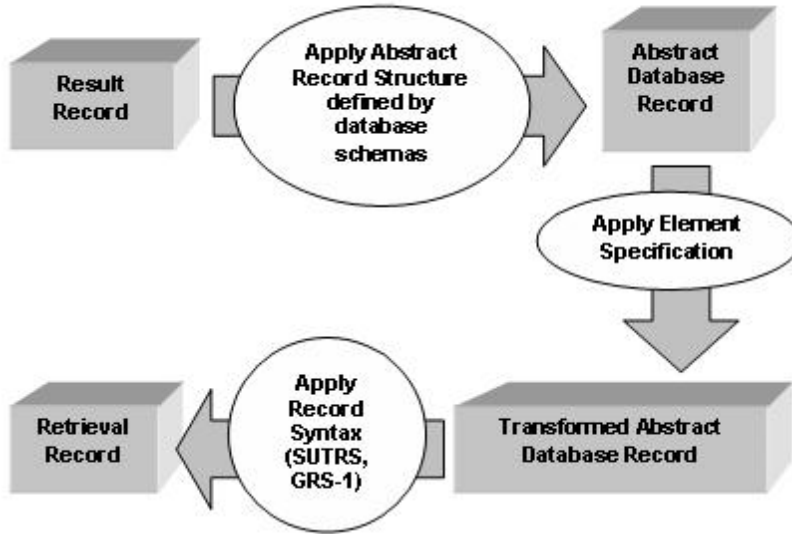


Рис. 1.2: Модель извлечения записей

Сервис	01.04.2001	26.03.2002	2.03.2003	31.03.2004
search	309 (100%)	302 (100%)	430 (100%)	623 (100%)
present	309 (100%)	302 (100%)	430 (100%)	623 (100%)
scan	226 (73%)	240 (79%)	319 (74%)	484 (78%)
sort	154 (50%)	135 (45%)	176 (41%)	280 (45%)
delSet	151 (49%)	166 (55%)	222 (52%)	359 (58%)
namedResSets	141 (46%)	175 (58%)	239 (55%)	388 (62%)
extendServices	80 (26%)	92 (30%)	130 (30%)	237 (38%)
concurrentOper	79 (26%)	108 (36%)	138 (32%)	255 (41%)
resourceCtrl	73 (24%)	52 (17%)	56 (13%)	56 (9%)
accessCtrl	54 (17%)	31 (10%)	37 (9%)	52 (8%)
triggerResCtrl	53 (17%)	29 (10%)	37 (9%)	37 (6%)
Level-1Segment	50 (16%)	27 (9%)	26 (6%)	38 (6%)
Level-2Segment	50 (16%)	28 (9%)	26 (6%)	25 (4%)
resourceReport	49 (16%)	32 (11%)	36 (8%)	48 (8%)

Таблица 1.3: Распространенность сервисов Z39.50

1.8 Сервис *Explain*

Протокол *Z39.50* позволяет достаточно эффективно работать с базами данных независимо от их физической структуры и месторасположения. Но, как известно, за все приходится платить. Абстрагировавшись от структуры реальных данных, мы получили взамен необходимость знать другую информацию. Так, для того чтобы произвести поиск, необходимо знать имена баз данных и поддерживаемые поисковые атрибуты, а для извлечения данных – форматы внешнего представления и наборы элементов, которые поддерживает данный сервер.

Можно поступить просто, тупо перебирая все возможные комбинации, получая в ответ диагностические сообщения об ошибках. Через некоторое время портрет выбранного сервера будет сформирован. Правда, имя базы данных таким образом не подберешь, если не считать ту, которую обычно считают базой данных по умолчанию с именем *xxdefault*. Тем не менее, некоторые системы, собирающие информацию о *Z39.50* серверах, именно так и поступают, например, робот от *IndexData*.

В протокол *Z39.50* заложен более изящный механизм получения информации о серверах и всех их параметрах. Протокол предписывает каждому серверу *Z39.50* держать специальную базу данных с предопределенным именем *IR-Explain-1*, в которой должна храниться информация, соответствующая настройкам сервера и всех его баз данных. К этой базе данных клиент может обращаться с запросами *RPN* по набору атрибутов *Exp-1* и получать ответы в формате *Explain*.

Если бы все серверы *Z39.50* поддерживали *Explain*, можно было бы строить самосогласованные информационные системы и клиентские рабочие места с самонастраивающейся конфигурацией. В процессе сеанса *Z39.50* клиенты бы автоматически выясняли параметры серверов и соответственно подстраивали бы свои экранные меню. У пользователей не было бы проблем с выяснением списка баз данных, поддерживаемых поисковых атрибутов и форматов внешнего представления записей. Соединившись с сервером, клиент уже знал бы все его возможности.

К сожалению, реальная жизнь далека от совершенства. Возможности, заложенные в протокол *Z39.50*, производителями программного обеспечения и просто менеджерами информационных систем реализуются сегодня далеко не полностью.

Глава 2

Z39.50 как основа распределенных информационных систем

В настоящее время технологиям и практике построения распределенных систем уделяется пристальное внимание. При этом большая часть работ посвящена технологиям распределенных вычислений, другая часть – технологиям доступа к распределенным информационным ресурсам (коллекциям, базам данных и т.п.). Спектр обсуждаемых технологий достаточно широк, однако уровень взаимодействия с базами данных, как правило, сводится к поддержке *ODBC/JDBC*, которые не могут обеспечить универсальных способов работы с базами данных – необходимого абстрагирования от конкретных систем и платформ при сохранении высокой степени функциональности и предоставления пользователю единого интерфейса доступа к различным СУБД при интеграции информационных ресурсов [7].

Ниже мы попытались собрать под одной обложкой различную информацию, касающуюся построения РИС на основе протокола *Z39.50* [1].

2.1 Что такое распределенная информационная система

Словосочетание ”распределенная информационная система” для однозначной трактовки изложенного в настоящей работе материала требует некоторого пояснения. Дело в том, что в это понятие различные авторы вкладывают различный смысл.

Информационная система здесь и ниже будет пониматься как совокупность упорядоченных информационных ресурсов, программного обеспечения, обеспечивающего доступ к этим ресурсам и набора пользовательских интерфейсов, управляющих доступом к этим ресурсам.

Распределенная информационная система это информационная система, ресурсы которой могут находиться на физически различных связанных коммуникационными линиями (сетью) серверах, но которая сохраняет логическую целост-

ность информации и обеспечивает единообразный доступ к ней (единые пользовательские интерфейсы), через любую, возможно единственную, точку входа.

Согласно такому определению, в распределенных информационных системах (РИС) не требуется исполнение распределенных приложений, а, следовательно, класс актуальных для них задач, решаемых, вообще говоря, при помощи широко обсуждаемых сегодня технологий распределенных систем [7] (.NET, J2EE, CORBA и др.) может быть существенно сужен.

По характеру работы с информационными ресурсами задачи, которые решаются при работе с РИС, можно разделить на пользовательские и административные. К пользовательским задачам относятся задачи поиска, извлечения и просмотра информации. Все остальные задачи, такие как пополнение информации, модификация, удаление информации, сбор статистики работы системы, управление ограничением доступа, мониторинг внутренней целостности и т.д., являются административными. В зависимости от структуры и организации РИС список административных задач может быть существенно расширен. Однако с точки зрения пользователя различные РИС отличаются только внешними интерфейсами пользовательских задач. Поэтому организация этих интерфейсов и их универсальность являются наиболее важным критерием оценки эффективности РИС.

2.2 Требования к РИС

Исходя из задач, которые должны решаться распределенной информационной системой, можно сформулировать минимальный список требований, которым она должна удовлетворять [8].

Работа с распределенными данными – информационная система должна допускать возможность работы с данными, расположенными на разных физических серверах, различных аппаратно-программных платформах и хранящихся в различных внутренних форматах.

Логическая группировка данных – система должна позволять обрабатывать все запросы на логических группах баз данных, полностью скрывая тем самым физическое расположение последних.

Абстрактная модель данных – информационная система строится на основе абстрактной схемы данных, на которую должны быть отображены конкретные базы данных. Это позволяет объединять данные из разнородных систем в одной логической группе, обеспечивая тем самым требование логической группировки.

Абстрактная система запросов – система должна оперировать не конкретным синтаксисом запросов, а его логической сутью на основе абстрактных атрибутов. Привязка поиска к синтаксису запросов конкретной подсистемы или СУБД ограничивает возможности сквозного поиска в других подсистемах.

Метаинформация – система должна предоставлять полную и достоверную информацию о себе и о всех своих ресурсах.

Разграничение доступа – система должна быть способна предоставлять различные уровни привилегий для пользователей по доступу к информации.

Учет и контроль – система должна уметь собирать статистику по запросам пользователей и вести их бюджеты.

Открытость – система должна легко расширяться и быть основана на открытых стандартах и протоколах.

Связь с другими системами – возможность интегрировать свои ресурсы с ресурсами других информационных систем. Это требование может быть удовлетворено лишь при выполнении требования открытости.

Легкость в общении – для пользователей система должна предоставлять различные, в том числе и простые, интерфейсы доступа к информации. Необходимым элементом системы сегодня является наличие *WEB*-шлюза, т.к. доступ через *WEB* сегодня является наиболее демократичным.

Следует заметить, что перечисленные выше требования не исчерпывают список требований к РИС в части решения административных задач. Приведенный список ограничен требованиями к внешним интерфейсам РИС, которые обеспечивают взаимодействие данной РИС с пользователями и другими системами.

2.3 Технологии построения распределенных систем

Для разработки распределенных информационных систем сегодня используются различные технологии [7]. Большая часть этих технологий по своей сути ориентирована не на работу с базами данных, а на сетевое взаимодействие программ и распределенные вычисления в гетерогенных средах. Даже беглый обзор используемых или обсуждаемых технологий может составить предмет отдельного увесистого тома (см. [9, 10]). Поэтому ниже упомянуты лишь некоторые технологии.

RPC – технология удаленного вызова процедур широко применяется для разработки сетевых приложений, в том числе и приложений для доступа к базам данных. Технология обеспечивает аппаратно независимый обмен данными между различными платформами, но в общем случае требует одновременной разработки серверного и клиентского приложения. Это делает ее малопригодной для построения распределенных информационных систем с разнородными данными без дополнительной стандартизации. Тем не менее, именно эта технология лежит в основе *DCE* [11], *DCOM/COM+* и многих специальных закрытых фирменных систем.

DCOM/COM+ – технологии фирмы *Microsoft* для работы с распределенными программными объектами общего назначения. Получила широкое распространение для платформы *Windows NT/2000*. Для работы с базами данных требует построения специальных объектов, реализующих специфичные для каждой СУБД методы. Фирма *Microsoft* предлагает набор объектов *DCOM* под общим названием *OLE DB* [12], позволяющий скрыть различия конкретных СУБД за общими программными интерфейсами, в том числе высокоуровневыми – *MS ADO*. Недостатки технологии *OLE DB* – уклон в сторону реляционной модели данных и, несомненно, недоступность на платформах, отличающихся от *Intel-Windows*.

RMI – технология распределенных вычислений в среде *Java* – набор интерфейсов для удаленных вызовов *Java*-приложений [10, 13, 14]. Технология ориентирована на взаимодействие программ, но не на взаимодействие баз данных. Для баз данных в *Java* существуют интерфейсы *JDBC* – аналог интерфейсов *ODBC* для среды *Windows*. Недостаток этих интерфейсов –

уклон в сторону реляционных систем и медлительность, которая является следствием большого объема промежуточных преобразований на линии "клиент-сервер". Для *Java*-приложений эта медлительность усугубляется медлительностью виртуальных *Java*-машин.

Специальные протоколы доступа к базам данных используют ведущие производители коммерческих СУБД для своих продуктов. Подобные технологии эффективны при условии построения информационной системы на базе одной СУБД, например, *ORACLE*. В противном случае требуется программное обеспечение промежуточного слоя, конвертирующее закрытые протоколы обмена данными в открытые, например, в *Z39.50*.

Z39.50 – протокол сетевого доступа к базам данных. Разработан специально для унификации интерфейсов доступа к различным базам данных на основе *ASN.1* в 1988 году [1]. Включает описание не только протокольных процедур взаимодействия "клиент-сервер", но и идеологии организации этого взаимодействия.

HTTP – протокол обмена данными в среде *WEB* – ориентирован на обмен неструктурированными данными, что делает его малоприменимым для доступа к базам данных. Однако, ввиду большой популярности *WWW*, **HTTP** может успешно применяться на стороне клиента для просмотра данных через *WEB*-серверы, выступающие в этом случае как серверы промежуточного слоя. В последнее время протокол **HTTP** часто используют как транспортный для передачи структур *XML*.

XML/SOAP – протокол работы с так называемыми *WEB*-сервисами. Выполняется поверх протокола **HTTP**. В последнее время является весьма популярным [15, 16] ввиду относительной простоты и возможности расширения в рамках принятых правил *XML*. Активно поддерживается ведущими разработчиками программного обеспечения. Однако ввиду новизны полностью отсутствует опыт эксплуатации систем на основе этой технологии.

XML/SOAP/SRW – технология, аккумулирующая лучшие свойства как *Z39.50*, так и *XML/SOAP*. Фактически представляет собой урезанный вариант структур *Z39.50*, представленных в виде структур *XML*. Позиционируется сегодня как следующая генерация технологий *Z39.50*. Находится в стадии активного обсуждения [17, 18].

На сегодняшний день обозначились три основных группы технологий построения распределенных систем

.NET – технология *Microsoft* для платформы *Windows 2000*. Технология включает поддержку специальной среды (*CLR*) для исполнения приложений, написанных на разных языках, среду *ASP.NET* для динамических *Web*-приложений. Для доступа к базам данных **.NET** использует технологии *MS ADO* и *OLE DB*.

J2EE (*Java 2 Enterprise Edition*, 1999) – технология *Sun Microsystems*. **J2EE** [10] включает описание компонентной модели *Enterprise JavaBeans*, описания правил для *WEB*-приложений *Java Servlets*, *Java Server Pages (JSP)* и правила доступа к базам данных *JDBC*. Технология **J2EE** активно используется при построении серверов приложений.

CORBA – проект *OMG (Object Management Group)* по стандартизации набора спецификаций для интеграции приложений, созданных в различной среде и выполняющихся в различных гетерогенных программно-аппаратных

средах. На сегодняшний день (*CORBA 3*, [10]) проект содержит наиболее полный набор спецификаций для распределенных систем.

В каждой из перечисленных групп предусматривается взаимодействие с базами данных. К сожалению, спецификации этого взаимодействия ограничены спецификациями *MS ADO*, *ODBC* или *JDBC*, стандартизация взаимодействия сводится к стандартизации языка *SQL* и метаданных для реляционных систем.

Как видно даже из этого беглого обзора, применяя перечисленные технологии отдельно или в комплексе, несомненно, можно организовать сетевую работу с базами данных. Более того, для каждой конкретной СУБД можно подобрать технологию, обеспечивающую высокий уровень эффективности и функциональности. Однако ни одна из упомянутых технологий, кроме *Z39.50*, не обеспечивает универсальных способов работы с базами данных – необходимого уровня абстрагирования от конкретных систем и платформ при сохранении высокой степени функциональности и предоставления пользователю единого интерфейса доступа к различным СУБД. В пользу технологии *Z39.50* для доступа к базам данных говорит тот факт, что обсуждается возможность включения *Z39.50* как в *.NET* и в *J2EE*, так и в *CORBA* [16, 19–24].

Для построения распределенных информационных систем, не требующих выполнения распределенных приложений, привлечение упомянутых выше технологий типа *CORBA* не является оправданным, тем более, что доступ к базам данных все равно лучше всего реализовывать через *Z39.50*. Ниже будет показано, что РИС можно построить, используя в основном только технологии *Z39.50*.

2.4 Протокол Z39.50 как технология построения РИС

Протокол *Z39.50* является, по-видимому, оптимальным протоколом для работы с базами данных в больших гетерогенных распределенных системах. Разрабатываемый изначально для доступа к библиографическим данным, этот протокол на сегодняшний день используется сотнями библиотек мира и крупнейшими информационными центрами как для предоставления доступа к своим ресурсам, так и для технологического обмена данными. Область применения протокола *Z39.50* в мировом информационном сообществе вышла за рамки своего изначального назначения. Сегодня протокол используется для доступа не только к библиографическим ресурсам, но и для доступа к различным научно-техническим и другим данным, например, к данным ГИС, к справочной информации и др. Также расширяются возможности самого протокола.

2.5 Недостатки Z39.50 как технологии построения РИС

Содержание описания стандарта *Z39.50* [1] позволяет утверждать, что протокол *Z39.50* изначально не предназначался как базовый протокол распределенных информационных систем. Несмотря на подробное описание регламента взаимодействия с базами данных в модели "клиент-сервер", в стандарте *Z39.50* не описаны некоторые детали, необходимые для построения распределенных информационных систем, к которым можно отнести следующие:

- Отсутствие описания модели 3-х уровневое взаимодействия с одной стороны не приводит к недоразумениям, т.к. любое взаимодействие типа "клиент-Х-...-Х-сервер" в конце концов может быть представлено как последовательность 2-уровневых связей типа "клиент-сервер", но, с другой стороны, не регламентирует функционирование серверов промежуточного слоя и порождает тем самым возможность реализации различных режимов их работы.
- Отсутствие описания механизмов переадресации запросов является следствием предыдущего недостатка. Однако без этих механизмов трудно представить РИС, в которой физическое расположение информации отличается от логического, обеспечение которого требует выполнения множества внутренних запросов для формирования правильной логической структуры.
- Отсутствие описания механизмов контроля маршрутизации запросов также является следствием первого недостатка. Этот контроль необходим в случаях возникновения ситуации множественной переадресации и петель переадресации.
- Отсутствие описания распределенного учета ресурсов приводит к невозможности построения в рамках стандартных средств распределенной базы данных *IR-Explain-1*. Дело в том, что в соответствии с идеологией *Z39.50* каждый сервер *Z39.50* в РИС может поддерживать собственную базу данных *IR-Explain-1*, в которой содержится исчерпывающая информация об его собственных ресурсах. Однако эти локальные базы данных, имеющие к тому же все одно и то же имя, ввиду своей жесткой предопределенной структуры не могут быть интегрированы в одну логическую базу данных, содержащую описание всего ресурса РИС в целом. Это ограничение, несомненно, отрицательно сказывается на возможности реализации службы обеспечения учета интегрированных ресурсов РИС.
- Отсутствие описания механизма выполнения распределенных транзакций затрудняет выполнение административных задач управления распределенными данными. Однако это совершенно не сказывается на пользовательских задачах, которые, как было сказано выше, являются основными задачами РИС.

Наличие перечисленных недостатков *Z39.50* при использовании его в качестве базового протокола РИС является досадным, но преодолимым препятствием. Дело в том, что в рамках стандарта предусмотрены механизмы расширения его возможностей и подстройки их под требования конкретных информационных систем.

2.6 Расширения базового *Z39.50* для построения РИС

Проектировщики *Z39.50* заложили в базовый стандарт возможность функционального его расширения, предоставив разработчикам программного обеспечения самостоятельно доопределять или определять некоторые структуры.

Существует несколько возможностей расширения возможностей базового набора средств *Z39.50*.

2.6.1 Локальные идентификаторы объектов

Применение локальных идентификаторов объектов (*OID*) является наиболее мощным и универсальным средством расширения. Суть этого метода сводится к "локальной стандартизации" любых дополнительных объектов и присвоению им уникального числового кода (*OID*).

Дело в том, что Z39.50, как и многие другие протоколы (*LDAP*, *SNMP* и др.) определен в терминах *ASN.1* [3] и для однозначной интерпретации объектов использует рубрику *ISO* (см. раздел 1.3). Z39.50 имеет в этом рубрикаторе свою ветвь, определяемую кодом

iso (1) – member_body (2) – US (840) – ANSI-Z39.50 (10003)

что дает для Z39.50 *OID*:1.2.840.10003. В дальнейшем этот *OID* будет обозначаться как (Z). Все объекты, подлежащие стандартизации, в Z39.50 получают префикс Z в своем *OID*. Стандартом определены классы объектов, приведенные в Табл. 1.1.

Следует заметить, что последние два класса объектов были определены после утверждения стандарта в 1995 году и не вошли в официальное издание стандарта.

Внутри каждого класса объектам присваивается идентификатор, в префиксе которого содержится *OID* класса. В частности, в классе Z.5 (см. Табл. 1.1) *RecordSyntaxDefinitions* содержится объект *SUTRS*, *OID* которого Z.5.101. Полный список актуальных *OID* всегда можно найти в глобальном реестре объектов:

<http://lcweb.loc.gov/z3950/agency/defns/oids.html>

Все вышесказанное касалось только объектов, определенных глобально. Благодаря этим определениям любое программное обеспечение Z39.50 любого производителя может однозначно интерпретировать и обработать тот или иной объект. Кроме этого стандарт Z39.50 допускает существование локально определенных объектов, область действия которых ограничена программным обеспечением того или иного производителя. Локально определенным объектам также присваивается *OID*, но формируется он добавлением после префикса класса кода 1000 и идентификатора разработчика. Глобальный реестр разработчиков программного обеспечения Z39.50 находится по адресу:

<http://lcweb.loc.gov/z3950/agency/register/entries.html>

Следуя этому правилу, в любом из указанных выше классов можно определить новый объект и присвоить ему локальный *OID*. Наиболее экзотическими являются новые объекты в классах

- Z.1 application context definitions
- Z.2 abstract syntax definition for APDUs
- Z.6 transfer syntax definitions

Новые объекты в них могут изменить Z39.50 до неузнаваемости (буквальной) и сделать его непригодным для применения вне программного обеспечения конкретного разработчика. Тем не менее, при этом буква закона (стандарта) Z39.50 нарушена не будет.

Однако возникает вопрос, как информация о локальных объектах может быть доведена до всех их использующих. Существует два способа:

1. *OID* локального объекта вместе с описанием объекта публикуется в глобальном реестре объектов. Это самый правильный способ, поскольку не только позволяет всем разработчикам получить необходимую информацию через *WWW*, но и при некоторых обстоятельствах, в конце концов, приводит к стандартизации объекта и выделении для него глобального *OID*.
2. Описание локального объекта помещается в базу данных *IR-Explain-1* в соответствующей категории на серверах, которые эти объекты поддерживают. Этот способ дополняет первый. К сожалению, далеко не все серверы *Z39.50* поддерживают сервис *Explain*.

2.6.2 Использование специальных полей в *APDU*

Согласно *Z39.50*, при взаимодействии клиент и сервер обмениваются прикладными пакетами (*APDU*), структура которых регламентирована в классе *Z.2*, содержащем 23 глобально стандартизированных объекта. В структуре всех типов *APDU* стандартом (v.3) предусмотрено необязательное поле *OtherInfo*, представляемое в терминах *ASN.1* как

```
OtherInformation ::= [201] IMPLICIT SEQUENCE OF SEQUENCE{
  category          [1] IMPLICIT InfoCategory OPTIONAL,
  information       CHOICE{
    characterInfo   [2] IMPLICIT InternationalString,
    binaryInfo      [3] IMPLICIT OCTET STRING,
    externallyDefinedInfo [4] IMPLICIT EXTERNAL,
    oid             [5] IMPLICIT OBJECT IDENTIFIER}}

InfoCategory ::= SEQUENCE{
  categoryId [1] IMPLICIT OBJECT IDENTIFIER OPTIONAL,
  categoryValue [2] IMPLICIT INTEGER}
```

Это поле содержит последовательность структур, идентифицируемых, как правило, по *OID*, в том числе и локальным. Информационное наполнение структур определяется структурой соответствующих объектов и в случае локально определенных объектов может содержать контент, требуемый для функционирования РИС.

2.6.3 Использование внешних определений

Для передачи структурированной информации, в том числе через поле *OtherInfo*, стандартом *Z39.50*, вернее *ASN.1*, предусмотрено наличие внешнеопределенных типов данных *EXTERNAL*. Объекты типа, *EXTERNAL* как правило, идентифицируются по *OID*, в том числе и локально определенным, и формально могут принадлежать любому классу. Использование данных типа *EXTERNAL* в структурах *OtherInfo* позволяет реализовать механизм передачи нестандартных структурированных данных любой сложности в стандартных *APDU*, не прибегая к переопределению последних.

2.7 *CIP* – попытка модернизации *Z39.50* для РИС

Сознавая привлекательность технологий *Z39.50* для построения РИС, с одной стороны, и отсутствие в *Z39.50* необходимых компонент для функционирования РИС, с другой стороны, *CEOS* (*Committee on Earth Observation Satellites*)

разработал специальный протокол для построения своей распределенной информационной системы – CIP (*Catalogue Interoperability Protocol*) [25,26] как модернизированный протокол Z39.50. Система предназначена для доступа к данным по дистанционному зондированию Земли, а членами CEOS являются

ESA <http://www.esrin.esa.it>
 NASA http://www.gsfc.nasa.gov/NASA_homepage.html
 DLR <http://www.dlr.de>
 NASDA <http://www.nasda.go.jp>
 CCRS <http://www.ccrs.nrcan.gc.ca>
 BNSC <http://www.open.gov.uk/bnsc/bnschome.htm>
 CEO <http://www.ceo.org>

Описание протокола и профиля CIP составляет довольно увесистый том и здесь, конечно, рассматриваться не будет. Следует лишь обратить внимание на специфические особенности CIP:

- CIP ориентирован на обработку распределенной информации, хранящейся на различных серверах и организованной в иерархические коллекции (базы данных).
- Для обеспечения обработки распределенных данных в CIP переопределены многие объекты в классах Z39.50, в том числе и объекты в классе APDU. Это существенно отличает расширения CIP от обычных расширений Z39.50.

Name	OID
CIP Releasee В APDU	Z.2.1000.99.1
CIP attribute set	Z.3.1000.99.1
CIP diagnostic set	Z.4.1000.99.1
CIP order extended service	Z.9.1000.99.1
CIP collection schema	Z.13.1000.99.1
CIP product schema	Z.13.1000.99.2
CIP user schema	Z.13.1000.99.4
CIP tag set	Z.14.1000.99.1

Таблица 2.1: Объекты Z39.50, переопределенные в CIP

- CIP включает дополнительные спецификации для поиска и извлечения информации, для поддержки *Explain*, расширенного сервиса, исполнения заказа.
- В CIP определены спецификации базы данных пользователей (*IR-User-1*), которая используется в том числе и для аутентификации.

Несмотря на привлекательность CIP, он не может быть использован для построения РИС универсального содержания, т.к. направлен на специфический вид информации. Однако многие компоненты, проработанные в CIP, могут быть применены для построения РИС на основе протокола Z39.50.

2.8 Выводы

Современный уровень развития информатизации требует наличия распределенных информационных систем, построенных на основе открытых международных стандартов с высокой степенью интероперабельности. Эти системы должны удовлетворять вполне определенным требованиям, которые продиктованы опытом внедрения и эксплуатации информационных систем.

Единственной технологией, способной сегодня обеспечить промышленное функционирование РИС в части доступа к распределенным данным, является технология на основе протокола Z39.50. Все остальные технологии или не удовлетворяют требованиям, указанным выше, или не имеют статуса технологий, прошедших проверку промышленной эксплуатацией.

Несмотря на то, что стандарт Z39.50 при своем создании не предназначался для построения РИС и не описывает режимы их функционирования, его базовых средств достаточно для построения РИС.

Для построения РИС стандарт Z39.50 должен быть расширен. Это расширение можно обеспечить собственными механизмами Z39.50.

Глава 3

Архитектура распределенных систем на основе *Z39.50*

Несмотря на то, что, как отмечалось в первой главе, протокол *Z39.50* изначально не предназначался для построения распределенных информационных систем (РИС), заложенные в него возможности позволяют создавать такие системы. В настоящей главе рассматриваются различные аспекты построения таких систем. При этом под РИС понимаются системы, попадающие под определения, сформулированные в 2.1.

3.1 Основные типы информационных систем

Исходя из деталей реализации различных РИС, можно выделить несколько их типов. Основным классификационным критерием здесь может служить используемый ими способ накопления и обработки информации.

Моноблочные ИС построены в виде единого комплекса программно-аппаратных средств. Все составляющие таких ИС, как правило, хорошо взаимодействуют между собой, используя недокументированные или плохо документированные интерфейсы и протоколы. Моноблочные ИС чаще всего ориентированы на работу с конкретными СУБД, трудно поддаются расширению и обладают плохой интероперабельностью. Тем не менее, в конкретной тематической области они позволяют создать достаточно эффективные системы, обеспечивающие быстрый поиск и извлечение информации, полное администрирование и высокую степень защиты данных на различных уровнях. Моноблочные ИС трудно считать распределенными, т.к. они, как правило, функционируют или на конкретном компьютере или на заведомо предопределенном их списке (кластере). "Распределенность" моноблочных ИС может достигаться применением распределенных СУБД или разнесением функциональных блоков ИС по разным компьютерам. Впрочем, последнее переводит моноблочную ИС в другую категорию.

Модульные ИС построены в виде отдельных функционально самостоятельных программных модулей. В идеальном случае каждый такой модуль может функционировать на отдельном компьютере. При этом ИС становится распреде-

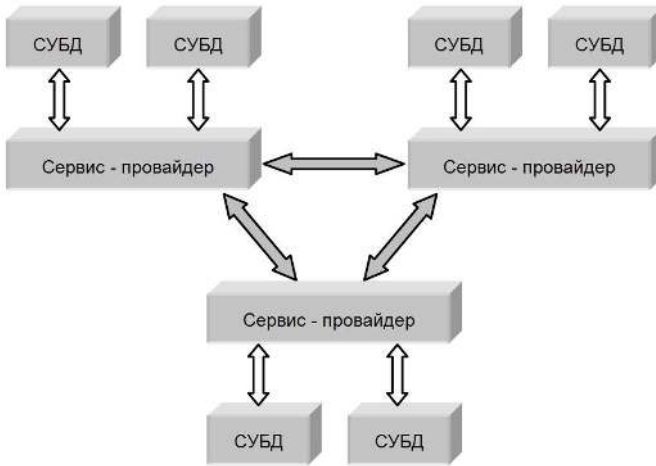


Рис. 3.1: Типичная схема распределенной модульной информационной системы

ленной. Модульные ИС требуют от разработчиков хорошо продуманных интерфейсов и протоколов взаимодействия отдельных модулей. Здесь можно заметить, что если все модули ИС создаются одним разработчиком (или одной фирмой), т.е. модульная ИС является программно моногенной, детальная документированность межмодульного взаимодействия ИС может отсутствовать. Такие ИС можно называть закрытыми, т.к. расширение их функциональных свойств за счет добавления новых модулей затруднено. В противоположность такой ситуации можно указать на программно гетерогенные ИС с полностью документированным межмодульным взаимодействием. Идеальный случай – унификация и стандартизация всех интерфейсов и протоколов межмодульного взаимодействия. Только в этом случае модульную ИС можно считать открытой, т.к. она открыта для добавления новых модулей и расширения тем самым своих функциональных возможностей.

Наиболее общий вид распределенной модульной информационной системы изображен на Рис. 3.1. Стрелками показаны сетевые взаимодействия модулей. При этом можно явно выделить семь логических уровней, которые представлены как на Рис. 3.2

1. **Уровень клиента** – совокупность программно-аппаратных средств клиента РИС, включающая средства поиска и визуализации информации.
2. **Транспортный уровень "клиент – сервис-провайдер"** – сеть, т.е. среда, обеспечивающая связь клиента с сервисными службами РИС. На сегодняшний день эта среда чаще всего представляет собой среду Интернет.
3. **Уровень сервисных служб РИС** – совокупность сетевых точек доступа в РИС. На этом уровне производится предварительная обработка запросов клиента для приведения их в соответствие с конкретными СУБД. В качестве субъектов этого уровня могут выступать различные серверы приложений, WEB-серверы с соответствующей настройкой, серверы Z39.50.

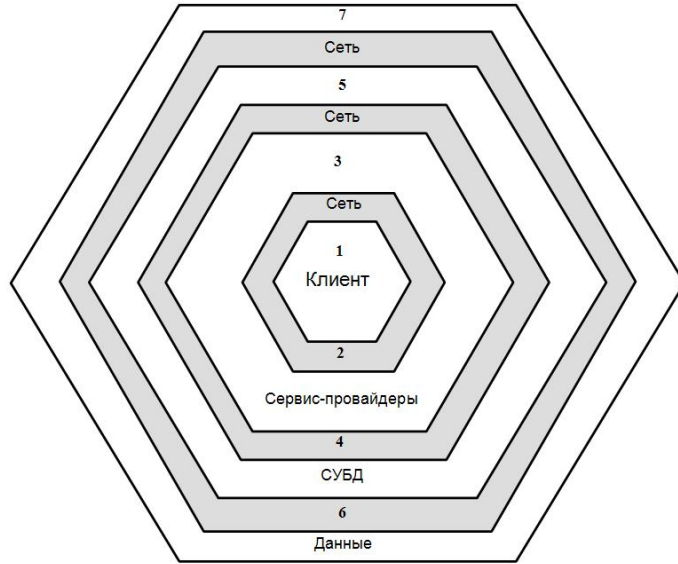


Рис. 3.2: Схема многослойной РИС

4. **Транспортный уровень "сервис-провайдер – СУБД"** – характерен для СУБД, имеющих сетевые интерфейсы, например, *SQL*-серверы (*ORACLE*, *MS SQL Server* и др.). Как правило, в качестве среды на этом уровне используются защищенные сети и ЛВС. Однако очень часто используется и Интернет.
5. **Уровень СУБД** – совокупность серверов с программным обеспечением СУБД
6. **Транспортный уровень "СУБД – данные"** – как правило, ЛВС или защищенные сети.
7. **Уровень данных** – совокупность файловых структур, в которых хранятся собственно данные. Организация зависит от конкретных СУБД.

В соответствии с общими правилами, сетевое взаимодействие между уровнями представленной модели описывается интерфейсами, а сетевое взаимодействие модулей внутри одного уровня – протоколами. В дальнейшем мы будем использовать эту терминологию. Здесь следует заметить, что на Рис. 3.2 не показаны сетевые взаимодействия модулей одного уровня. Однако следует понимать, что каждый представленный на рисунке уровень состоит из совокупности взаимодействующих по сети модулей. При этом правила этого взаимодействия регламентируются соответствующими протоколами..

3.2 Основные блоки РИС на основе Z39.50

Стандарт *Z39.50* [1] определяет следующие компоненты субъектов сетевого взаимодействия в модели *клиент-сервер*:

<i>Origin</i>	– компонента, инициирующая сеанс связи Z39.50;
<i>Target</i>	– компонента, ожидающая команды от-крытия сеанса Z39.50;
<i>Database user</i>	– пользователь данных;
<i>Database provider</i>	– поставщик данных;
<i>Service provider</i>	– компонента, принимающая сетевые за-просы и ответы;
<i>Service user</i>	– компонента, использующая сетевые за-просы и ответы;
<i>Client</i>	– приложение, состоящее из <i>origin</i> и <i>database user</i> ;
<i>Server</i>	– приложение, как правило состоящее из <i>target</i> и <i>database provider</i> ;

При этом фиксируется три четко выделенных уровня сетевого взаимодей-ствия. На Рис. 5 представлены эти уровни для клиента и сервера. Следует за-метить, что два верхних уровня относятся к последнему (Application) уровню семиуровневой модели *OSI*, а Z-транспортный уровень включает в себя осталь-ные шесть. Более тонкая структура уровней в Z39.50 не определена.

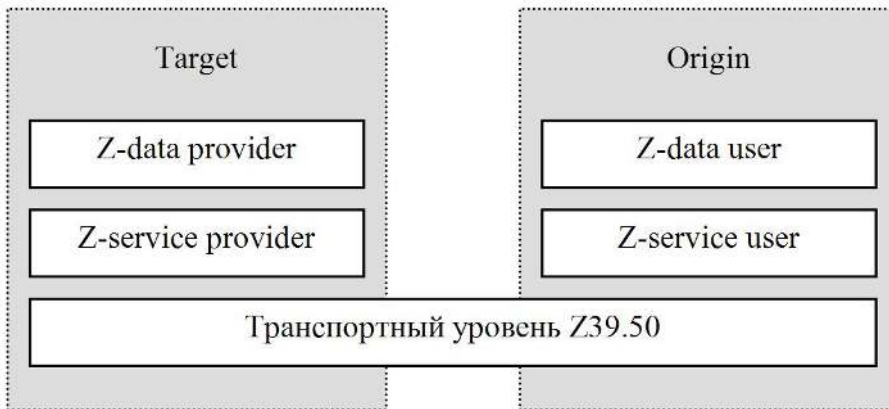
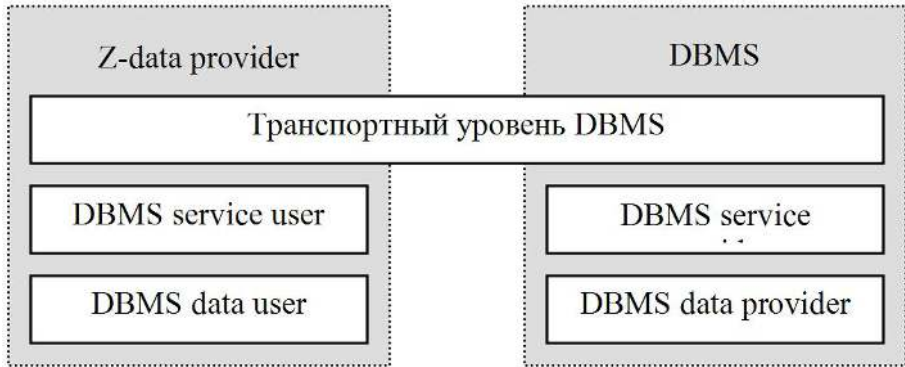


Рис. 3.3: Уровни сетевого взаимодействия в Z39.50

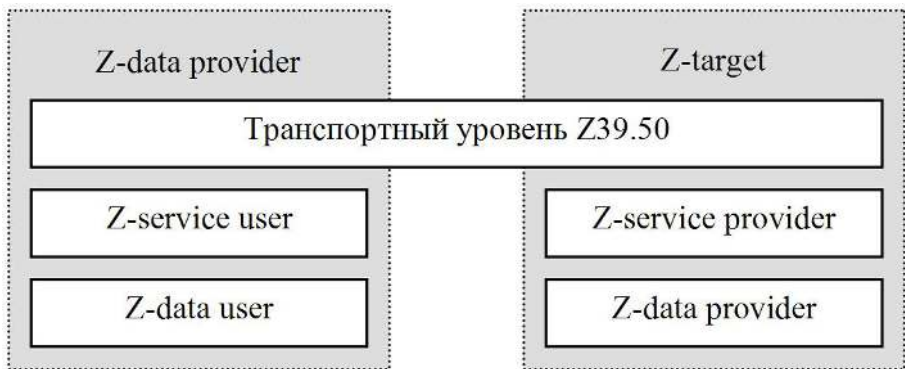
Наибольший интерес представляет верхний уровень сервера, определяющий взаимодействие с базами данных (*Z-data provider*). Здесь возможно несколько вариантов.

1. СУБД, управляющая данными, исполняется на том же самом сервере, на котором выполняется *target* (см. Рис. 3.3). В этом случае цепочка доступа к данным выглядит следующим образом:

- Запрос конвертируется из стандартного синтаксиса Z39.50 в синтаксис конкретной СУБД и передается ей.
- СУБД исполняет запрос и возвращает данные в своем внутреннем формате.
- Данные из внутреннего формата СУБД преобразуются в требуемый внешний формат Z39.50 и возвращаются на уровень *Z-service provider*.

Рис. 3.4: Уровни сетевого взаимодействия *Z-data provider* с сетевой СУБД

2. СУБД, управляющая данными, исполняется на другом сервере (см. Рис. 3.4). При этом *Z-data provider* должен реализовать функции клиента СУБД, т.е. состоять из блоков "data user – service user – data transport", первые два из которых относятся к седьмому уровню модели *OSI*, а последний включает в себя оставшиеся шесть. В этом случае сервер *Z39.50* играет роль сервера промежуточного слоя.

Рис. 3.5: Уровни сетевого взаимодействия *Z-data provider* с другим сервером *Z39.50*

3. Исключительный случай представляется, когда в качестве сетевой СУБД выступает другой сервер *Z39.50* (см. Рис. 3.5). Принципиально он не отличается от предыдущего, однако в качестве транспортного уровня СУБД здесь выступает транспортный уровень *Z39.50*.

В распределенной информационной системе могут сосуществовать все три описанных выше варианта. Но основное внимание ниже будет уделено третьему варианту, когда межсерверный обмен происходит по протоколу Z39.50.

Таким образом, распределенную информационную систему на основе протокола Z39.50 можно представить как совокупность серверов Z39.50, распределяющих между собой обработку запросов клиентов. Схематично такая система изображена на Рис. 3.6, на котором под базами данных (БД) следует понимать клиентские части сетевых СУБД.

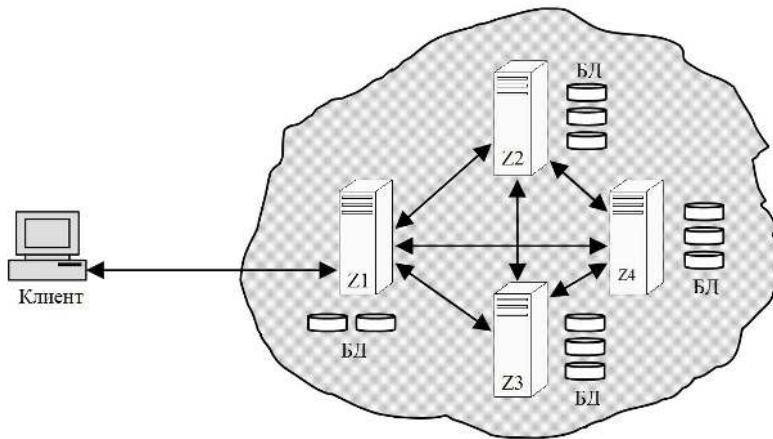


Рис. 3.6: Модель РИС на основе протокола Z39.50

3.3 Способы и механизмы обеспечения "распределенности" в Z39.50

Протокол Z39.50 явно не описывает способов построения распределенных систем и механизмов обеспечения "распределенности". Однако многие описанные в нем протокольные процедуры могут быть или непосредственно применены к распределенным системам, или в рамках технологий Z39.50 модифицированы до приобретения необходимых свойств.

3.3.1 Объединение баз данных

В ситуации, когда для пользователя все базы данных одинаковы, если

- поддерживаются одинаковые поисковые атрибуты
- записи извлекаются в одном формате внешнего представления

создается возможность для логического объединения этих баз данных [27, 28]. Пользователю нет необходимости знать реальные имена баз данных, достаточно знать некие условные имена, скрывающие группы однотипных баз данных. Указывая такое имя в качестве имени базы данных в запросе на поиск, пользователь

осуществит его во всей группе, а всю "кухню" одновременного поиска обеспечит сервер.

Конечно, пользователь нечто подобное может проделать и сам, т.к. в запросе на поиск (*searchRequest*) серверу *Z39.50* передается не имя базы данных, а список имен. Согласно протоколу сервер должен выполнить поиск во всех указанных базах данных. Однако первый вариант кажется более привлекательным, т.к. в этом случае у пользователя отпадает необходимость запоминать длинный список имен баз данных и условия их объединения в одном запросе.

Что касается извлечения данных, то здесь проблемы объединения баз данных не существует. Это определяется тем, что данные извлекаются не из баз данных, а из промежуточных наборов, созданных сервером в момент осуществления поиска. Поэтому к моменту запроса на извлечение данных эти данные, извлеченные из разных баз в момент поиска, уже будут объединены в одном промежуточном наборе.

Наконец, процедура просмотра поисковых термов (*scan*) тоже может выполняться над группой баз данных. Все необходимое при этом, т.е. сортировку общего списка термов, сервер должен взять на себя.

Итак, протокол *Z39.50* предоставляет замечательные возможности логического объединения баз данных, которое для пользователя выглядит как физическое объединение. Это незаменимое свойство при работе в системах, в которых существует множество однотипных баз данных.

3.3.2 Переадресация запросов

Рассмотренное выше свойство протокола *Z39.50* объединять базы данных может быть существенно усилено при возможности перенаправления сервером запросов на другие серверы *Z39.50*.

Здесь имеется в виду следующее. Предположим, клиент посылает серверу А запрос на поиск в базе данных – *APDU searchRequest*. Сервер А, проверяя список указанных в запросе баз данных, обнаруживает, что базы данных DB1 у него нет, но есть на сервере В. Сервер А посылает *APDU searchRequest* серверу В с указанием базы данных DB1 и получает ответ – *APDU searchResponse*. Далее сервер А комбинирует полученный ответ с результатом поиска в своих базах данных и отправляет ответ клиенту – *APDU searchResponse*. Клиент, получив ответ, будет удовлетворен, так и не узнав, что выполнение его запроса происходило на двух серверах *Z39.50*.

Описанный сценарий легко расширяется на процедуры извлечения данных и просмотр термов. Такой подход позволяет объединять базы данных расположенные не только на одном сервере, но и на различных серверах. Единственное требование – доступность по протоколу *Z39.50*.

Следует заметить, что в описанном сценарии клиент и сам мог открыть сеанс связи со вторым сервером и выполнить поиск в его базах данных. Однако когда речь идет о десятках баз данных, в которых нужно осуществить поиск, и местоположение которых знает только администратор, вариант с открытием клиентом многочисленных сеансов связи с множеством серверов представляется не очень удобным.

Тем не менее, имеет право существовать промежуточный вариант, когда роль клиента *Z39.50* выполняет шлюз *Z39.50-HTTP*. Этот шлюз выполняется на *WEB*-сервере и легко управляется администратором. Пользователь при этом не является клиентом *Z39.50*, а является клиентом *WEB*. Диспетчеризацию запросов к различным *Z*-серверам осуществляет шлюз, предоставляя тем самым

пользователю работать в распределенной информационной системе без серверного механизма перенаправления запросов.

Таким образом, используя потенциальную возможность сервера перенаправлять запросы Z39.50 на другие серверы или потенциальную возможность шлюза Z39.50-HTTP открывать многочисленные сеансы связи с серверами Z39.50, можно построить распределенную информационную систему. Термин "потенциальная возможность" здесь употреблен в том смысле, что или сервер должен иметь такого *database provider*, который мог бы выполнять роль клиента Z39.50 (третий вариант реализации *data provider* раздела 3.2), или шлюз должен "уметь" работать в многосерверном окружении.

3.3.3 Контроль маршрутизации запросов

Возможность перенаправлять запросы на другие серверы Z39.50 может привести к возникновению ситуаций [7]:

- Множественная переадресация реализуется при неоптимальном маршруте запроса. Если сервер А направляет запрос на сервер В, а сервер В отправляет его на сервер С (маршрут А-В-С), то оптимальным был бы маршрут запроса А-С.
- Петля переадресации возникает при появлении замкнутых повторяющихся маршрутов типа А-В-А-В-А-В-А. Без дополнительных ограничений появление петли в маршруте прохождения запроса приводит к непроизводительным потерям серверных ресурсов и бесконечному времени выполнения запроса.

Проконтролировать возникновение этих ситуаций можно при наличии полной информации о маршруте запроса, отсутствующей в стандартных структурах *APDU*. Задача решается довольно просто с применением механизмов, описанных в разделе 2.6. Через поле *OtherInfo* может быть передан объект типа *EXTERNAL*, содержащий список всех серверов, через которые прошел запрос. Этот список должен пополняться при каждом акте переадресации. Появление одинаковых строк в списке будет указывать на возникновение петли. Ситуация при этом будет контролируема. Более того, список серверов может быть дополнен полем времени, анализ которого позволяет получить временные характеристики прохождения запроса.

3.3.4 Построение распределенной БД Explain

Как отмечалось в разделе 2.5 в базовом Z39.50 отсутствует возможность построения базы данных *IR-Explain-1* (см. также раздел 1.8) для множества серверов Z39.50, которые задействованы в распределенной информационной системе. Формально задача может быть решена, если

1. Создать логическую распределенную базу данных, которая объединяет локальные БД *IR-Explain-1*. Задача решается методами, описанными в разделе 3.3.1.
2. В записи *Explain* в каждой категории определить элемент, содержащий информацию об идентификации сервера Z39.50, который эту запись формирует. Следует заметить, что стандартно идентификационная информация присутствует только в категории *TargetInfo* (элемент */name*). Тем не менее, наличие в каждой категории элемента *commonInfo/otherInfo* позволяет, например, через тип *EXTERNAL* (см. раздел 2.6) зафиксировать имя сервера и сделать его точкой доступа при поиске.

3. Переопределить смысл и способ формирования записи в категории *CategoryList*. Дело в том, что при формальном объединении баз данных мы получим количество записей этой категории равное количеству серверов *Z39.50*. Реально же должна существовать одна запись *CategoryList*, содержащая список поддерживаемых категорий Explain в распределенной базе данных.
4. Предоставить возможность поиска в распределенной БД *Explain* по имени сервера. При выполнении предыдущих требований, это реализуется через поисковый атрибут *Exp-1 USE=4 (TargetName)*.

3.4 Оптимизация распределенных запросов

Как правило, в гетерогенных информационных системах на основе протокола *Z39.50* существует одна или несколько точек доступа, реализованных в виде шлюзов *Z39.50-HTTP* или в виде серверов *Z39.50* с функциями перенаправления запросов.

При многобазовом поиске информации в распределенных информационных системах происходит обращение к множеству баз данных, в том числе расположенных на различных серверах. При этом каждый сервер, входящий в информационную систему, может управлять одной или несколькими базами данных, хранящимися на различных физических носителях информации. Если учесть, что все серверы информационной системы имеют различное быстродействие и реальную загрузку, что они соединены каналами связи с различной пропускной способностью, то возникает вопрос о наиболее оптимальной конфигурации распределенной информационной системы с точки зрения скорости обработки сложного запроса, т.е. времени, прошедшего с момента прихода запроса в точку входа в РИС до отправки клиенту ответа (см., [29, 30]).

Следует заметить, что функционирование сервера *Z39.50* в составе распределенной информационной системы возможно в трех режимах. Эти три режима проиллюстрированы на Рис. 3.7, Рис. 3.8 и Рис. 3.9 в виде временных диаграмм взаимодействия клиента и серверов в распределенной системе.

Режим 1 соответствует диаграмме (см. Рис. 3.7) сеанса связи клиента и сервера 1, перенаправляющего запросы серверу 2, который взаимодействует с СУБД. Сеанс состоит из процедуры открытия сеанса (1), двух запросов на поиск (2, 7) и процедуры закрытия сеанса. При этом обработка запросов на поиск сервером 1 сводится к открытию им сеанса связи с сервером 2 (3, 8), отправки ему запроса на поиск (4, 9) и закрытия сеанса (8, 11). Доступ к СУБД имеет сервер 2 (5, 10) при обработке им запроса на поиск от сервера 1. Заметим, что в этом режиме любое обращение сервера 1 к серверу 2 сопровождается установлением сеанса связи, выполнением операции и закрытием сеанса. С одной стороны, это позволяет не заботиться о сохранении контекста вторичных сеансов, что экономит сеансовую память сервера 1, но, с другой стороны, приводит к увеличению сетевого трафика и времени ожидания клиентом окончания операций. Дело в том, что время, необходимое на открытие сеанса, может быть намного больше времени выполнения любой операции, например, поиска в базе данных. При этом непроизводительные потери клиентского времени растут пропорционально количеству операций, выполняемых клиентом в течение одного сеанса. Тем не менее, именно этот режим работы сервер свободен от ловушек, о которых речь пойдет ниже, и прост для администрирования распределенной системы в целом. Также следует заметить, что именно режим 1 принят основным в распределенных

системах CIP (*CIP – Catalogue Interoperability Protocol* – прикладной профиль Z39.50, созданный для космического мониторинга земной поверхности) [25].

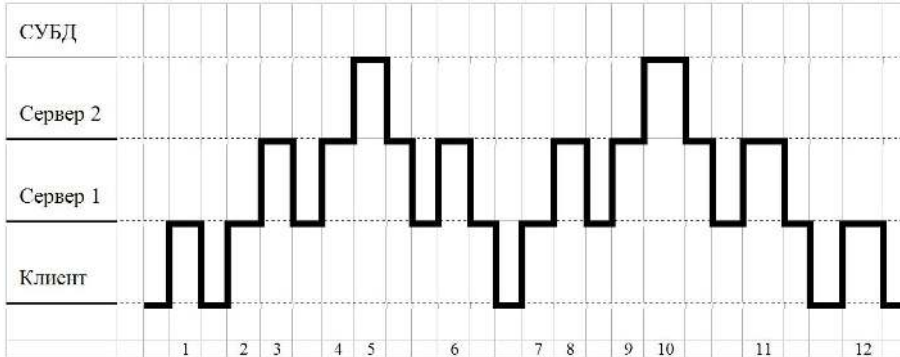


Рис. 3.7: Временная диаграмма режима 1 взаимодействия клиента и серверов в РИС

Режим 2 соответствует диаграмме, изображенной на Рис. 3.8. В этом режиме при открытии клиентом сеанса связи с сервером 1 (1) последний сразу открывает сеанс связи с сервером 2 (2) и сохраняет контекст этого сеанса в своих сеансовых переменных. Поэтому последующие операции клиента, например, операции поиска (3, 6), для сервера 1 сводятся лишь к передаче соответствующих запросов серверу 2 (4, 7) в контексте открытого ранее сеанса. При закрытии сеанса связи клиентом (9) закрывается сеанс связи сервера 1 с сервером 2 (10).

Как видно из логики работы сервера, режим 2 свободен от непроизводительных потерь клиентского времени в режиме 1. Однако следует иметь в виду, что в распределенной системе сервер 1 может общаться не только с сервером 2, но и с другими серверами, например, серверами 2, 3, 4. При этом для каждого из них в контексте сеанса "клиент-сервер1" должен быть открыт свой сеанс связи, например, "сервер1-сервер2", "сервер1-сервер3", "сервер1-сервер4" и т.д. Процесс открытия этих вторичных сеансов требует времени, а сохранение их контекста – ресурсов сервера 1. И если последнее не очень существенно при достаточной оперативной памяти, то время ожидания клиентом открытия сеанса связи с сервером 1 в режиме 2 будет пропорционально количеству серверов в распределенной системе и может оказаться очень большим.

Впрочем, проблема большого времени ожидания открытия сеанса решается на уровне архитектуры провайдера. Дело в том, что поскольку операции открытия сеансов связи с разными серверами независимы, они могут происходить в отдельных потоках и тем самым выполняться практически одновременно.

Что же касается запросов с однократной переадресацией, то следует ожидать, что в режиме 2, обсуждавшемся выше, работа распределенной информационной системы будет эффективней.

Однако, при работе нескольких серверов в распределенной системе в режиме 2 возможна ситуация возникновения инициализационных петель. Такая петля появляется уже в системе из двух серверов, работающих в режиме 2. Суть ее состоит в том, что если сервер 1 настроен на открытие вторичного сеанса с сервером 2, а сервер 2 – на открытие вторичного сеанса с сервером 1, т.е. создана

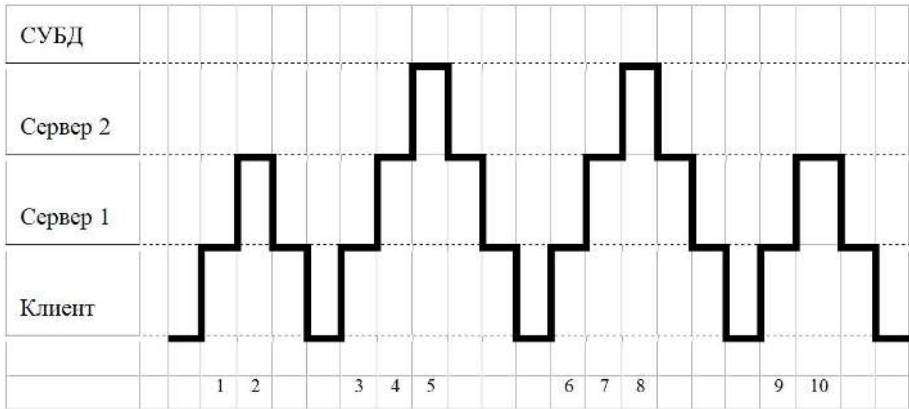


Рис. 3.8: Временная диаграмма режима 2 взаимодействия клиента и серверов в РИС

симметричная конфигурация серверов, то клиент никогда не сможет открыть сеанс связи ни с сервером 1, ни с сервером 2. Действительно, в этом случае событийная цепочка будет выглядеть следующим образом:

- клиент запрашивает сервер 1 для открытия сеанса "клиент-сервер1"
- открывая этот, сеанс сервер 1 запрашивает сервер 2 для открытия вторичного сеанса "сервер1-сервер2"
- открывая этот, сеанс сервер 2 запрашивает сервер 1 для открытия вторичного сеанса "сервер2-сервер1"
- открывая этот, сеанс сервер 1 запрашивает сервер 2 для открытия вторичного сеанса "сервер1-сервер2"
- и т.д. до бесконечности

Другой момент, который ухудшает временные характеристики распределенной системы и присущ как режиму 2, так и режиму 1, состоит в возможности появления вторичной переадресации. Эту ситуацию можно проиллюстрировать на примере. Пусть сервер 1 переадресует запрос серверу 2, а сервер 2 его не обрабатывает, а переадресует запрос серверу 3. В этом случае конфигурация, когда сервер 1 сразу бы обращался к серверу 3, была бы более экономичной. Переадресации уровня выше первого в распределенной системе не должно быть вообще, т.к. ее наличие ничего кроме проблем не дает. Более того, вторичная переадресация может привести к появлению петли переадресации, в том числе и в режиме 1.

Для защиты от избыточной переадресации, инициализационных петель и петель переадресации серверы РИС должны обеспечивать контроль маршрутов запросов, упоминавшийся выше.

По способу инициализации вторичных сеансов можно указать еще один режим функционирования РИС.

Режим 3 соответствует диаграмме, изображенной на Рис. 3.9. В этом режиме сервер 1 открывает сеанс связи с сервером 2 лишь при необходимости, но всегда сохраняет контекст открытого вторичного сеанса до закрытия первичного. После

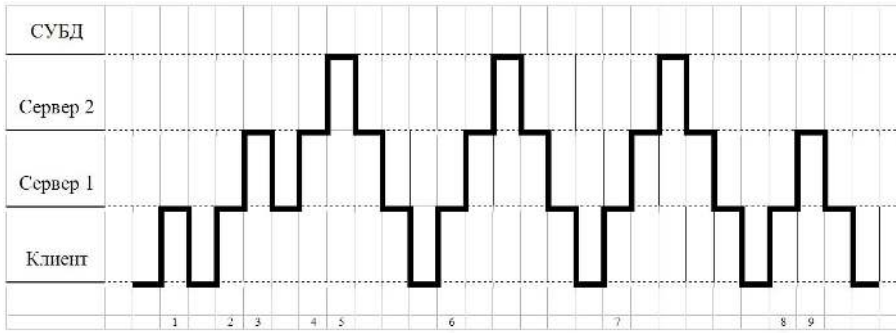


Рис. 3.9: Временная диаграмма режима 3 взаимодействия клиента и серверов в РИС

открытия вторичного сеанса связи "сервер 1 – сервер 2" все взаимодействие этих серверов происходит в контексте открытого сеанса.

Режим 3 имеет преимущества перед режимами 1 и 2:

- Минимизируются непроизводительные потери времени на открытие вторичных сеансов по сравнению с режимом 1.
- Отсутствует необходимость открытия сеансов связи с серверами, с которыми связь может не понадобиться, как в режиме 2.
- Отсутствует возможность появления петель инициализации, характерных для режима 2.

С точки зрения минимизации потерь времени режим 3 является наиболее оптимальным.

3.5 Администрирование распределенных систем

Как отмечалось в разделе 2.1, список административных задач, которые может или должна решать распределенная информационная система, может варьироваться в зависимости от ее структуры. Ниже перечисляются некоторые из этих задач, приводится описание технологии их решения для РИС на основе Z39.50.

3.5.1 Пополнение, модификация и удаление информации

Для систем на основе Z39.50 модификация информации в РИС сводится к модификации записей в базах данных. Для Z39.50 эта модификация является протокольной процедурой, реализуемой через функции расширенного сервиса. При этом транзакция изменения записей в БД иницируется при получении сервером *APDU extendedServiceRequest* и должна быть завершена (принята или отвергнута) компонентом *Z-data provider* до отсылки клиенту *APDU*

extendedServiceResponse. Все промежуточные действия протоколом не регламентируются.

В распределенных системах имеет право существовать возможность модификации записей в распределенных базах данных. При этом *Z-data provider* каждого сервера *Z39.50* в РИС должен обладать свойством контроля распределенных транзакций. Авторам неизвестна ни одна реализация этой возможности в РИС. Все известные серверы *Z39.50* если и используют функции расширенного сервиса для модификации записей в базах данных, то применяют их только для локальных баз данных.

Гораздо большей популярностью пользуется вариант модификации записей в базах данных средствами соответствующих СУБД локально вне пространства *Z39.50*. Немаловажную роль при этом играет разграничение ответственности за информацию между поставщиками информации в РИС. Каждый поставщик ведет свои базы данных своим инструментарием, открывая доступ к своим базам для остальных только на чтение. Именно в таком режиме сегодня пополняются все распределенные информационные системы.

3.5.2 Учет ресурсов РИС

Стандарт *Z39.50* для учета ресурсов любого сервера предписывает ведение и поддержание в актуальном состоянии специальной базы данных с предопределенным именем *IR-Explain-1* (см. раздел 1.8) и предоставление пользователям доступа к ней. Выше (см. раздел 3.3.4) был описан способ построения распределенной БД *IR-Explain-1* в рамках стандартных структур *Z39.50*, что позволяет организовать сервис информирования пользователей обо всех ресурсах РИС с описанием структуры этих ресурсов и способов доступа к ним. Несмотря на кажущуюся простоту организации распределенной БД *IR-Explain-1*, авторам на сегодняшний день не известен ни один факт ее реализации в известных РИС.

Более того, в сообществе разработчиков программного обеспечения для *Z39.50* в России (и не только в ней) сам сервис *Explain* не пользуется большой популярностью. В частности, на сервере компании *IndexData* [6] представлена статистика поддержки *Explain* в списке тестируемых серверов *Z39.50*. Серверов с поддержкой *Explain* в мире зарегистрировано всего 14 из 304 на 01.04.2001, 18 из 304 на 26.03.2002, 21 из 430 на 2.03.2003 и 33 из 613 на 29.01.2004. Ситуация в России несколько лучше, на 7.12.2003 *Explain* поддерживало 15 из 64 зарегистрированных [31] серверов *Z39.50*. Распределение по поддерживаемым категориям *Explain* (данные *IndexData*) представлено в таблице 3.1.

Также следует заметить, что различия между декларированием поддержки сервиса *Explain* и действительной поддержкой этого сервиса вносят некоторую путаницу и искажают даже эту не очень оптимистичную картину. В качестве примера можно привести сервер *RUSLAN*, который, декларируя поддержку целого списка категорий *Explain*, реально поддерживает только *CategoryList* (см. z3950s://ruslan.ru:210).

Для доступа к информации о ресурсах РИС можно использовать и отличные от *Z39.50* технологии. В частности, существуют решения [32–35], основанные на протоколе *LDAP*. Рассмотрение этих технологий выходит за рамки настоящей статьи.

3.5.3 Сбор статистической информации о работе РИС

Сбор и анализ информации о работе распределенной информационной системы в целом и отдельных ее компонент является обязательной административной

Explain Category	01.04.01	26.03.02	2.03.03	29.01.04
CategoryList	14 (100%)	18 (100%)	19 (90%)	31 (94%)
TargetInfo	14 (100%)	18 (100%)	21 (100%)	33 (100%)
DatabaseInfo	14 (100%)	18 (100%)	21 (100%)	33 (100%)
AttributeSetInfo	13 (93%)	15 (83%)	16 (76%)	28 (85%)
SchemaInfo	8 (57%)	10 (56%)	8 (38%)	13 (39%)
TagSetInfo	8 (57%)	10 (56%)	9 (43%)	14 (42%)
AttributeDetails	8 (57%)	10 (56%)	14 (67%)	24 (73%)
RecordSyntaxInfo	4 (29%)	7 (39%)	7 (33%)	13 (39%)
TermListInfo	2 (14%)	2 (11%)	4 (19%)	4 (12%)
TermListDetails	2 (14%)	2 (11%)	4 (19%)	4 (12%)
ExtendedServicesInfo	1 (7%)	1 (6%)	0	0
SortDetails	0	0	1 (5%)	1 (3%)

Таблица 3.1: Поддерживаемость категорий *Explain* зарегистрированными серверами

задачей [36, 37].

В Z39.50 сбор статистической информации о работе конкретного сервера усложняется обилием типов *APDU* (23 стандартных типа) и достаточно сложной структурой этих *APDU*. Это приводит к проблематичности обычного метода записей в log-файлы, когда одна запись (одна строка log-файла) отражает суть одного запроса или ответа. Для однозначной интерпретации результатов протоколирования работы сервера Z39.50 необходимо протоколирование в структуры, аналогичные структурам обрабатываемых *APDU*. Решение напрашивается само собой – каждый сервер Z39.50 должен вести специальную базу данных статистики, структура записи которой соответствует структуре всех *APDU*, и пополнять ее при каждом акте обработки любого *APDU*.

К сожалению, в Z39.50 не существует глобального определения схемы данных для статистической информации, т.е. не определены такие стандартизуемые объекты как набор поисковых атрибутов, структура записи, набор меток, наборы элементов и форматы внешнего представления. Однако все эти объекты можно определить локально и обязать все серверы Z39.50 вести базу данных учета статистики.

Для РИС в целом учет статистики может быть сведен к определению распределенной логической базы данных статистики, состоящей из локальных баз.

Из известных авторам серверов Z39.50 различных разработчиков только сервер *ZooPARK* [38] поддерживает описанную технологию, но ни одна из известных РИС сегодня не собирает полную статистическую информацию, однако работы в этом направлении активно ведутся.

3.5.4 Аутентификация пользователей РИС

В Z39.50 определены механизмы регистрации пользователей для авторизованной работы с информационными ресурсами через *APDU initRequest* и *accessControlRequest/accessControlResponse*, но не регламентированы механизмы аутентификации.

Если для работы одного сервера Z39.50 достаточно локальной аутентификации, то при работе в РИС необходима или поддержка синхронизованных по всей

РИС локальных баз данных пользователей, или централизованная аутентификация пользователей через единую для всей РИС базу данных пользователей. При этом следует иметь в виду, что в *Z39.50* нет глобального определения схемы данных профиля пользователя. Поэтому требуются или локальные определения или использование других схем, например, схем, определенных в *CIP* [25].

Для синхронизации локальных баз данных могут быть задействованы механизмы *Z39.50* через функции расширенного сервиса.

Гораздо сложнее дело обстоит с централизованной аутентификацией. Дело в том, что для нее неэффективно использовать распределенную базу данных пользователей как логическое объединение локальных баз. Основная причина – дублирование имен. Если эту причину обойти, например, снабдив пользователей каждого сервера *Z39.50* уникальным префиксом в имени, механизм логического объединения локальных баз может быть задействован и тем самым реализована аутентификация пользователей в распределенной системе в рамках *Z39.50*. Другой, менее экзотический способ аутентификации пользователей в РИС требует применения технологий *LDAP* (см., например, [34, 35, 39]), что выходит за рамки *Z39.50*.

На сегодняшний день авторам неизвестна ни одна РИС на основе *Z39.50* в России, в которой бы использовалась аутентификация пользователей, отличная от локальной. Однако перспективы привлечения *LDAP* широко обсуждаются.

3.5.5 Ограничение доступа к ресурсам РИС

Право ограничения доступа к тому или иному ресурсу РИС принадлежит поставщику этого ресурса и выражается в создании списков доступа (*ACL – Access Control List*) к ресурсу со стороны, как пользователей, так и компьютеров клиентов и оформление их в виде локальной базы данных *ACL*. В распределенной информационной системе в рамках *Z39.50* естественным решением является создание распределенной базы данных *ACL* как логического объединения локальных баз *ACL*. Обязательным требованием при этом является необходимость формулирования локальных *ACL* в терминах имен распределенной базы данных пользователей, обсуждавшейся выше.

Альтернативой технологии *Z39.50* по контролю доступа к информационным ресурсам РИС может выступать технология распределенных каталогов *LDAP*, обсуждение которой выходит за рамки настоящей работы.

3.5.6 Мониторинг целостности РИС

Административный сервис мониторинга целостности РИС может быть реализован в рамках *Z39.50* через расширенный сервис (исполнение запросов по расписанию). Его реализация в РИС не требует дополнительных соглашений.

3.6 Выводы

Для построения РИС на основе протокола *Z39.50* требуется наличие программного обеспечения, реализующего специальные функции сетевого взаимодействия серверов *Z39.50*, при этом серверы *Z39.50* должны совмещать роли *target* и *origin* и уметь функционировать в качестве серверов промежуточного слоя трехзвенной модели.

Каждая РИС требует индивидуального подхода для построения наиболее оптимальной конфигурации. При этом должны быть учтены такие факторы как

количество и объемы БД в РИС, количество интегрированных в РИС серверов, пропускная способность каналов связи, тип доступной в РИС информации, интенсивность обращения к РИС и др.

Практически все административные задачи в РИС могут быть решены в рамках технологий Z39.50, но для их решения требуется создание большого количества служебных распределенных баз данных как логического объединения локальных.

Глава 4

Состав распределенных систем на основе Z39.50

Несмотря на разнообразие содержания большинство РИС подобны по типам обрабатываемых данных. В этой главе рассматриваются основные типы информации, которые могут быть доступны в РИС.

4.1 Метаданные как основа РИС

Информация, которая может быть доступна в РИС, носит разнообразный характер. Следствием этого является необходимость использовать различное программное обеспечение и различные пользовательские интерфейсы для доступа к ней. Однако существует ряд приемов, позволяющих получить доступ к различной информации однотипным способом и через однотипные интерфейсы:

Систематизация – все информация должна быть сгруппирована по типу и содержанию.

Каталогизация – каждая единица хранения каждого типа должна быть описана единым для всего типа образом. Полученные описания в совокупности тоже образуют некий информационный ресурс, который принято называть вторичным.

При правильном подходе к каталогизации информационных ресурсов получаемые описания могут быть хорошо структурированы и содержать всю основную информацию о первичном объекте. Такие описания принято называть метаданными.

Наиболее характерное свойство метаданных – одинаковая структура вторичных данных, соответствующих первичным объектам одного типа. Именно это свойство удобно использовать при организации поиска в РИС, когда поиск информации производится не по массивам первичной информации, а по массивам их описаний – метаданным. Более того, можно утверждать, что информационные ресурсы одного типа могут быть объединены только на уровне метаданных, т.к. только метаданные в разных системах могут иметь одинаковую структуру и семантику. Однако чтобы этого добиться требуется стандартизация типов метаданных.

4.2 Особенности метаданных в Z39.50

Идеология Z39.50 предусматривает работу в основном со стандартизованными (глобально или локально) объектами. Наряду с другими, объектами стандартизации выступают описания схем данных и соответствующих им абстрактных структур записей. При этом описание схемы данных и структуры записи производится в иерархической схеме в терминах меток, определенных в стандартизованных наборах.

Обсуждавшиеся выше метаданные, вернее их схемы данных, очень хорошо соответствуют схемам данных Z39.50. При этом каждому типу метаданных должна соответствовать своя схема данных Z39.50. Положительным моментом оказывается тот факт, что схемы метаданных при таком подходе получаются стандартизованными. А именно это требуется для интеграции информационных ресурсов в РИС.

4.3 Типы метаданных

4.3.1 MARC

Стандарт *MARC* (*Machine Readable Catalogue*) был создан на основе стандарта обмена *ISO-2709* [40] в конце шестидесятых годов с целью оказания помощи службам каталогизации в обмене каталожными записями. Он использовался в библиотечных автоматизированных службах, как основа для манипулирования библиотечными записями, для их индексирования и отображения. Варианты стандарта *MARC*, такие как *USmarc* [41] и *UKmarc*, а позднее *UNIMARC* [42], как основа для создания национальных версий форматов, появились как ответ на существующие конфликты форматов в качестве реализации индивидуальных требований различных библиотек.

Следует отметить, что хотя изначально стандарт *USmarc* не предназначался для описания сетевых ресурсов, он стал основой для проекта *OCLC*. Этот проект развил реализацию стандарта *MARC*, объединенного с правилами *AACR2* (*Anglo American Cataloguing Rules*) [43] для описания Web-ресурсов различных типов и осуществления доступа к ним.

Форматы семейства *MARC*, будь то *USmarc* или *RUSmarc* [44], или любой другой, являются ввиду своей жесткой структурированности трудными для модификации. Например, создание нового элемента требует согласования с сообществом *USmarc* или *RUSmarc*. Библиографические описания следуют набором правил *AACR2* [43] и *ISBD* (*International Standard for Bibliographic Description*) [45] для формата *USmarc* и правилам каталогизации, отраженными в ГОСТ-7.1-84 [46] для формата *RUSmarc*. При этом имеющаяся *MARC* модель обслуживает библиотеки и пользователей несовершенно, поскольку в основном она направлена на печатные первоисточники. Основная причина – плоская структура формата *MARC*, требующая фрагментации иерархически подчиненных компонентов в отдельные дискретные записи. Такая структура записи отражает требования компьютеров начала семидесятых годов, когда системы были жестко привязаны к возможностям и объемам хранения данных, но в настоящее время уже является несовершенной. В литературе все чаще и чаще отмечается, что стандарт *MARC* совершенно непригоден для взаимодействия с новыми технологическими требованиями информационно-библиотечных систем, но, по-видимому, он еще много лет будет оставаться на плаву из-за миллиардов библиотечных записей, функционирующих в информационно-библиотечных системах. По этой причине,

одной из основных задач, которую приходится сегодня решать – это различные отображения между стандартом формата *MARC* и другими форматами.

Если логическая структура записей различна в различных *MARC*-форматах, то физическая структура записи у всех одна – *ISO-2709* [40].

Формат *ISO-2709* в *Z39.50* не определяется и в терминах *ASN.1* не представляется. Он определен в другом стандарте *ISO* (OID: 1.0.2709).

Запись в формате *ISO-2709* в *Z39.50* кодируется как *OCTET STRING*, т.е. как произвольная байтовая последовательность. Структура этой последовательности определяется стандартом *ISO-2709*.

Каждая запись *ISO-2709* должна содержать:

- Маркер записи, состоящий из 24-х символов.
- Справочник.
- Поля данных переменной длины, отделенные друг от друга разделителем поля.
- Разделитель записи.

Ниже кратко описан каждый из этих блоков.

ISO-2709 предписывает, что каждая запись начинается с 24-х символьного маркера записи. Он содержит данные, относящиеся к структуре записи, определения которых даются в стандарте *ISO-2709*, а также некоторые элементы данных, выделенные *ISO-2709* для особого применения.

Смещение	Длина	Поле
0	5	Длина записи
5	1	Статус записи
6	4	
10	1	Длина индикатора
11	1	Длина идентификатора
12	5	Начало данных
17	3	
20	1	Длина записи о длине поля данных
21	1	Длина записи о начальной позиции
22	1	
23	1	Резерв

Таблица 4.1: Структура маркера записи *ISO-2709*

За маркером записи следует справочник. Каждая статья справочника состоит из трех частей:

- Метки, длина которой определяется значением байта 11 в маркере.
- Числа, указывающего длину поля данных. Это число занимает количество байт, соответствующее байту 20 маркера.
- Числа, указывающего начальную позицию данных. Это число занимает количество байт, соответствующее байту 21 маркера.

Другие символы в статье справочника не допускаются. Первая часть каждой статьи справочника – метка поля. Вторая часть статьи справочника определяет

число символов в поле, на которое указывает метка, приведенная в первой части статьи. В это число включаются все символы – индикаторы, идентификаторы подполей, текстовые или кодированные данные и разделитель полей. Третья часть статьи справочника содержит позицию первого символа поля относительно позиции первого символа той части записи, которая содержит переменные поля. Первый символ первого переменного поля имеет символьную позицию 0. Положение символьной позиции 0 внутри целой записи задается позициями символов 12-16 маркера записи. Справочник заканчивается разделителем поля.

Поля данных переменной длины следуют за справочником и содержат данные. Метки не содержатся в полях данных, а приводятся только в справочнике. Поля данных состоят из двух индикаторов и следующим за ними любым количеством подполей. Каждое подполе начинается с идентификатора подполя, который состоит из разделителя подполя и кода подполя, идентифицирующего подполе. За идентификаторами подполя следуют кодированные или текстовые данные произвольной длины, не превышающей указанной в начале описания поля. Последним символом данных в записи является символ конца записи, следующий за символом конца поля.

Наконец, относительно всех *MARC* следует сделать некоторые оговорки.

Во-первых, все *MARC* присутствуют в *Z39.50* в виде отдельных *OID* в классе 1.2.840.10003.5 (*Z.5*) – record syntax definitions. Например,

<i>USmarc</i>	Z.5.10
<i>RUSmarc</i>	Z.5.28

Наличие различных *OID* в классе физического представления записей (record syntax definitions) для записей в одной физической структуре (*ISO-2709*) вызывает некоторое недоумение. Это недоумение усиливается тем фактом, что в классе "record syntax definitions" отсутствует *OID* структуры записи собственно *ISO-2709*.

Во-вторых, при наличии одинаковой физической структуры все *MARC* отличаются семантикой, которую принято определять для *Z39.50* в классе *Z.13* – database schema definitions. Однако в этом классе не существует ни одного объекта, соответствующего какому-либо *MARC*.

Наличие этих противоречий в *Z39.50* носит исторический характер. Нам представляется, что все уже привыкли называть, например, *RUSmarc* форматом данных, а не схемой данных, в то время как он является именно схемой данных, а форматом является *ISO-2709*. Тем не менее, отсутствие в *Z39.50* для *MARC* четкого разграничения понятий формата (record syntax) и схемы данных приводит к некоторым трудностям при автоматической конвертации данных между различными схемами.

4.3.2 *GILS*

Из самого названия *GILS* (*Global Information Locator Service*¹) (см, [47–50]) следует, что профиль был создан для описания информационных ресурсов. *GILS* регламентирует работу с данными, которые содержат информацию о других данных.

В профиль *GILS* входят [47]

¹В первоначальном варианте *GILS* – *Government Information Locator Service*.

- Набор меток *GILS* – *tagSet GILS* (*OID: Z.14.4*), дополняющий стандартные наборы меток *tagSet-M* (см. В.1) и *tagSet-G* (см. В.2) для определения элементов абстрактной структуры записи. Набор содержит более 100 меток (см. Приложение В.3).
- Схема *GILS* – *schema GILS* (*OID: Z.13.2*), определяющая иерархическую абстрактную структуру записи. Количество элементов схемы данных – около 120 (см. Приложение С.1).
- Набор поисковых атрибутов *GILS* (*OID: Z.3.5*) – содержит специальные значения *USE* (более 70) для поиска, например, информации с географической привязкой (см. Приложение А.3).
- Определение элементов короткого набора элементов (В).
- Обязательные форматы внешнего представления – *USmarc*, *GRS-1*, *SUTRS*.

В отличие от *Dublin Core* (см. ниже) профиль *GILS* создан для *Z39.50* и жестко определяет структуру описания информационного ресурса. Следствием этого является возможность построения на основе *GILS* распределенных информационных систем с хорошо развитым поисковым сервисом [51].

Практическое применение *GILS* для построения РИС будет обсуждаться в последующих главах (см. разделы 7.5, 8.2)

4.3.3 *Dublin Core*

В сообществе разработчиков *WEB*-приложений термины *метаданные* и *Dublin Core* являются сегодня синонимами. Наверное, это неудивительно, потому что набор *Dublin Core*, несмотря на декларирования его универсальности, разрабатывался и продолжает совершенствоваться для описания информации, представленной на *WEB*-страничках. Основная идея проекта *Dublin Core* – сделать каталогизацию простой и доступной для создателей *WEB*-ресурсов (см., например, [52]).

Действительно, набор *Dublin Core (DC)* содержит всего 15 элементов, с интуитивно понятной семантикой. Однако, используя эти 15 элементов, можно создать лишь очень грубое описание ресурса. Поэтому для реальной работы с *DC* требуются уточнения многих элементов, как в части их семантики, так и в части форматов представления информации, что приводит к необходимости использования *DC*-квалификаторов, например, *DC.Date.Created* и указания различных атрибутов элементов, например, атрибута *lang* (указание языка описания) и атрибута *Scheme* (указание схемы кодирования). В результате простота *DC* становится не совсем очевидной, а строгость описания ресурса все равно не достигается.

В *Z39.50* не существует глобального профиля, который бы соответствовал *Dublin Core*. Однако на *DC* отображаются, конечно, с потерями информации, все схемы данных *Z39.50*. Более того, некоторые схемы данных *Z39.50*, например, *CIMI*, явно содержат элементы *DC*.

Тем не менее, *Dublin Core* ”присутствует” в *Z39.50* в виде поисковых атрибутов. Во-первых, существует глобально стандартизованный набор поисковых атрибутов *XD-1* (*Cross-Domain Attribute Set*, *OID: Z.3.12*, см. Приложение А.5) [53], а, во-вторых, утверждены поисковые атрибуты *DC* в наборе атрибутов *Bib-1* (*OID: Z.3.1*) [54].

Наличие поисковых атрибутов *Dublin Core* в наборах поисковых атрибутов *Z39.50* предоставляет потенциальную возможность включить ресурсы *Z39.50* в

Название атрибута	Значение USE Bib-1
DC-Title	1097
DC-Creator	1098
DC-Subject	1099
DC-Description	1100
DC-Publisher	1101
DC-Date	1102
DC-ResourceType	1103
DC-ResourceIdentifier	1104
DC-Language	1105
DC-OtherContributor	1106
DC-Format	1107
DC-Source	1108
DC-Relation	1109
DC-Coverage	1110
DC-RightsManagement	1111

Таблица 4.2: Поисквые USE-атрибуты *Bib-1* для *Dublin Core*

систему индексации *WEB*-ресурсов и сделать их доступными для поисковых *WEB*-машин. Обратная же задача, т.е. предоставление доступа по протоколу *Z39.50* к *WEB*-ресурсам, не является однозначной ввиду отсутствия в *Z39.50* схемы данных *DC*.

При наличии в РИС метаданных смешанного типа, т.е. ресурсов, каталогизированных на основе *DC* и метайнформации в базах данных *Z39.50*, необходимо обеспечение их взаимного преобразования.

4.3.4 *GEO*

Для однозначного описания такого информационного ресурса как книга достаточно полей, например, *RUSmarc*. Для описания информационных ресурсов Интернет требуется дополнительная информация, например, доступность по времени суток, правила ограничения доступа, географическое местоположение ресурса и т.п. Такая информация может быть представлена в схеме *GILS*, рассмотренной выше. Однако существует класс ресурсов, описание которых требует большей детализации, чем в состоянии предоставить схема *GILS*. Это всевозможные геоинформационные ресурсы, объекты различных ГИС. Описание этих ресурсов (географические карты, снимки земной поверхности, цифровые изображения и т.п.), требует указания масштаба, проекции, координат границ и многого-многого другого. Обычно эти ресурсы описываются по некоторой стандартной схеме метаданных, например, по схеме *FGDC* (*U.S. Federal Geographic Data Committee* [55]) (см., например, [56–58]).

Стандарт метаданных *FGDC*, так называемый *CSDGM* – *Content Standard for Digital Geospatial Metadata* [59], определяет структуру и содержание метаданных, но не определяет схему данных в стиле *Z39.50*. Для связи стандарта *CSDGM* и *Z39.50* разработан специальный профиль *GEO*.

Профиль *GEO* [60] включает в себя схему данных *GEO*, набор поисковых атрибутов *GEO-1*, обязательные наборы элементов, форматы внешнего представ-

ления (*SUTRS*, *GRS-1*, *XML*, *HTML*), допустимый синтаксис запросов (*RPN*), допустимую диагностику (*Bib-1*). Профиль позволяет предоставлять доступ к метаданным *FGDC* по протоколу *Z39.50*, обеспечивая тем самым возможность построения распределенных информационных систем на основе метаданных. Подобные системы сегодня функционируют в США. Практическое применение *GEO* в РИС рассматривается в разделе 8.6.

Таблица 4.3: Объекты *Z39.50* для *GEO*

Поисковые атрибуты:	Bib-1	Z.3.1
	GILS	Z.3.5
	GEO	Z.3.9
Диагностика:	Bib-1	Z.4.1
Форматы:	GRS-1	Z.5.105
	HTML	Z.5.109.3
	SUTRS	Z.5.101
	XML	Z.5.109.10
Форматы дополнительные:	SGML	Z.5.109.9
	USmarc ¹	Z.5.10
Схемы:	GEO Schema	Z.13.4

Схема *GEO* содержит около 450 элементов, набор поисковых атрибутов *GEO-1* – около 340 атрибутов *USE*. Здесь они рассматриваться детально не будут, т.к. принципиально схема *GEO* ничем не отличается от других схем данных.

4.3.5 *Digital Collections*

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.

Под цифровыми коллекциями принято понимать упорядоченный набор цифровых объектов, которые могут иметь, а могут и не иметь нецифровых аналогов. Примеры цифровых объектов – оцифрованные аудиозаписи, изображения картин, цифровые фотографии, программы, базы данных и т.п. Подобные объекты и коллекции этих объектов могут быть систематизированы и описаны, т.е. могут быть созданы метаданные для цифровых объектов. Естественно, что они будут существенно отличаться от обычных библиографических описаний. В частности, обязательным элементом описания цифрового объекта является ссылка на сам объект, например, по *URL*.

Для доступа к метаданным цифровых объектов и коллекций существует профиль *Digital Collections* [63].

Профиль *Digital Collections* определяет конформное подмножество определений *Z39.50-1995* для доступа к цифровым коллекциям, организованным посредством метаданных, структура которых описывается в этом профиле. Он также предоставляет семантику для навигации по цифровым коллекциям, обнаружения и получения объектов.

¹Взаимное соответствие полей USmarc и элементов GEO приведено в [61, 62]

Данный профиль предполагает, что сопутствующие профили (совместимые расширения или подмножества рассматриваемого профиля) будут разработаны независимо, расширяя или ограничивая использование этого профиля для определенных приложений или классов информации, например, музейных объектов, спутниковых фотографий, географических данных или химических соединений.

Таким образом, данный профиль непосредственно не решает все задачи доступа к цифровым коллекциям, однако предоставляет сопутствующим профилям каркас для этого:

- Цифровые объекты представляются атомарными, то есть их содержимое является "непрозрачным" и не обрабатывается профилем. Таким образом, профиль предоставляет поиск описательной информации, а не поиск цифровых объектов. Сопутствующие профили могут моделировать содержимое специфических объектов (например, музейных).
- Связанные описания (например, средство поиска, каталоги выставок и т.д.) рассматриваются как "непрозрачные" (их содержимое не обрабатывается профилем), хотя клиенты могут иметь в своем распоряжении дополнительные приложения, имеющие возможность их обработки и отображения. Сопутствующие профили могут моделировать содержимое связанных описаний.
- Профиль не моделирует сложные взаимодействия между объектами всех классов. Сопутствующие профили могут делать это для определенных классов объектов.

Хотя в рассматриваемом профиле особое внимание уделено логическому отделению описательной информации от содержимого, он не включает в себя руководства или спецификации, определяющих информацию как описание или содержимое. Профиль также не указывает принципы организации, т.е. пути сбора объектов в коллекции (см. [63–65]).

Не смотря на то, что данный профиль адресован на распределенные коллекции, он не нацелен на распределенные базы данных. Различные части коллекции могут управляться различными организациями таким образом, что различные базы данных, соответствующие одной коллекции, могут располагаться на различных серверах, однако индивидуальная база данных не может быть распределена по многим серверам.

Настоящим профилем используется только протокол Z39.50. Профиль, однако, не препятствует наличию мультипротокольных клиентов, шлюзов или цифровых коллекций, в которых одна часть доступна по протоколу Z39.50, другая же – по какому-либо другому протоколу.

В Табл.4.4 показано, с какими объектами Z39.50 оперирует данный профиль.

Для программного обеспечения, декларирующего совместимость с профилем *Digital Collections*, определяется три уровня совместимости, как для клиента, так и для сервера:

1. Базовая совместимость – поддержка версии 2 протокола Z39.50-1995, схемы и структуры описательной записи *Digital Collections* и синтаксиса записи *GRS-1*.
2. Базовая V3 совместимость – базовая совместимость и дополнительно поддержка версии 3 протокола Z39.50-1995.
3. Расширенная V3 совместимость – поддержка базовой V3 совместимости и дополнительно поддержка
 - параметра *CompSpec* сервиса *Present*;

Поисковые атрибуты:	Bib-1 Collection-1	Z.3.1 Z.3.7
Диагностика:	Bib-1	Z.4.1
Форматы:	GRS-1	Z.5.105
Схемы:	DigitalCollections Schema	Z.13.3
Наборы меток:	TagSet-M TagSet-G Collections TagSet	Z.14.1 Z.14.2 Z.14.5

Таблица 4.4: Объекты Z39.50 для *Digital Collections*

- спецификации элементов *eSpec*;
- набор вариантов *variant-1*;
- фрагментации;
- *Metadata* и *appliedVariant* для *GRS-1*.

Для структуры записи в этой схеме характерна одна деталь. Поскольку толкование цифрового объекта достаточно широкое, он может быть таким, что для его описания используется другая схема данных. В этом случае в записи присутствует *OID* другой схемы данных. Примером такого случая является музейная информация.

4.3.6 CIMI

Правила, которые регламентируют работу с музейной информацией в соответствии с Z39.50, сформулированы в документах консорциума *CIMI* (*Computer Interchange of Museum Information*) и составляют профиль *CIMI* [66]. 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 характерны для этого профиля.

В таблице показано с какими объектами Z39.50 оперирует профиль *CIMI*:

На основе этих объектов определена модель поиска и извлечения данных с музейной информацией.

Во-первых, все серверы *CIMI* обязаны поддерживать запросы *Type-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* [63].

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

Поисковые атрибуты:	Bib-1	Z.3.1
	CIMI-1	Z.3.8
Диагностика:	Bib-1	Z.4.1
Форматы:	GRS-1	Z.5.105
	SUTRS	Z.5.101
	Usmarc	Z.5.10
Схемы:	Digital Collections Schema	Z.13.3
	CIMI Schema	Z.13.5
Наборы меток:	TagSet-M	Z.14.1
	TagSet-G	Z.14.2
	Collections TagSet	Z.14.5
	CIMI TagSet	Z.14.6

Таблица 4.5: Объекты Z39.50 для CIMI

mrObject/rendition/resource. Некоторые аспекты применения схемы данных CIMI можно найти в работах [67–70].

В ситуации, когда музейное информационное сообщество России только начинает создавать информационные системы с предоставлением доступа к ним через Интернет, очень актуальным является 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.

4.4 Тезаурусы и рубрикаторы в РИС

В настоящее время, как в России, так и за рубежом существует значительное количество различных тезаурусов и классификационных схем, используемых в информационно-библиотечной практике. Отраслевой тезаурус по сельскому хозяйству, используемый в ЦНСХБ, русско-английская версия известного медицинского тезауруса *MeSH*, используемого в ГЦНМБ, Государственный рубрикатор НТИ (*ГРНТИ*), используемый так или иначе практически во всех информационных органах и библиотеках России, а также рубрикатор ВИНТИ, используемый этой организацией для классификации реферируемых источников – это лишь незначительный перечень из огромного количества тезаурусов и классификаторов, которые используются в настоящее время. Как правило, во всех организациях, в которых используются ресурсы этого класса, они оформлены в машиночитаемые базы данных, поддерживаемые собственным программным обеспечением и как следствие этому – затруднена интеграция и совместное использование таких баз данных в распределенных информационных системах.

Насколько известно авторам, ряд организаций, располагающих такими базами данных (далее БДТК – базы данных тезаурусов и классификаторов), в настоящее время ведут работы над тем, чтобы обеспечить унифицированный сетевой доступ к ним. Цель этих работ – обеспечение возможности не только просмотра тезаурусов и классификаторов в удобном интерфейсе, но и активное использование их при поиске соответствующей информации в базах данных, в том числе и библиографических.

4.4.1 Тезаурусы и рубрикаторы в архитектуре *клиент-сервер*

С развитием технологий построения больших распределенных информационных систем, включающих в себя множество различных баз данных, достаточно актуальным становится вопрос поиска информации в БД с использованием тезаурусов и классификационных схем. Более того, в распределенной информационной системе логично обеспечить доступ к базам данных тезаурусов и классификационных схем в той же самой технологии, в которой осуществляется доступ к БД, т.е. в технологии *клиент-сервер* (см., например, [71–73]).

4.4.2 Тезаурусы и рубрикаторы в технологии *Z39.50*

Исходя из общей идеологии *Z39.50*, доступ к любой базе данных, в том числе и к БДТК, должен осуществляться через единую стандартную схему данных, на которую должны быть корректно отображены все частные структуры БДТК. Такая схема данных, вернее профиль, существует – *Zthes* (*OID Z.13.8*) [74] и активно обсуждается (см. Приложение С.2).

Кроме описания абстрактной структуры записи профиль *Zthes* определяет типы термов и типы связей между термами. Терм может иметь тип:

ТТ – терм верхнего уровня, т.е. терм, не имеющий связанных термов более широкого класса (термов с типом связи ВТ).

НТ – не терм верхнего уровня, т.е. основной терм, имеющий связи типа ВТ

ND – не основной терм.

NL – фиктивный терм, т.е. терм, по которому не происходит индексация документов.

Типы связей могут быть:

ВТ – связь с родительским термом, т.е. с термом более широкого смысла.

НТ – связь с дочерним термом, т.е. с термом более узкого смысла. Связь ВТ-НТ является взаимно обратной.

USE – связь с термом, который используется вместо этого.

UF – взаимно обратная связь USE.

RT – связь, определяющая связанный терм.

Таким образом, профиль *Zthes* предполагает использование ограниченного набора типов отношений между термами из множества, рассматриваемого в [75]. Утверждается, что это сделано преднамеренно для большей совместимости.

В частности, одна из статей (“Мошки”) рубрикатор ВИНТИ выглядит следующим образом (представление *XML*):

```
<zthes>
  <local-control-number> 341.33.23.19.07.17.11
  </local-control-number>
  <termQualifier> 341.33.23.19.07.17.11 </termQualifier>
  <Title> Мошки </Title>
  <termType> NT </termType>
  <relation>
    <relationType> ВТ </relationType>
    <local-control-number> 341.33.23.19.07.17
    </local-control-number>
  </relation>
</zthes>
```

```

    <termQualifier> 341.33.23.19.07.17 </termQualifier>
    <Title> Двукрылые </Title>
</relation>
<relation>
  <relationType> NT </relationType>
  <local-control-number> 341.33.23.19.07.17.11.02
    </local-control-number>
  <termQualifier>341.33.23.19.07.17.11.02
    </termQualifier>
  <Title> Общие проблемы </Title>
</relation>
<relation>
  <relationType> NT </relationType>
  <local-control-number>341.33.23.19.07.17.11.07
    </local-control-number>
  <termQualifier>341.33.23.19.07.17.11.07</termQualifier>
  <Title> Морфология </Title>
</relation>
<relation>
  <relationType> NT </relationType>
  <local-control-number>341.33.23.19.07.17.11.09
    </local-control-number>
  <termQualifier>341.33.23.19.07.17.11.09</termQualifier>
  <Title>Систематика, фауна, географическое
    распространение </Title>
</relation>
<relation>
  <relationType> NT </relationType>
  <local-control-number>341.33.23.19.07.17.11.11
    </local-control-number>
  <termQualifier>341.33.23.19.07.17.11.11</termQualifier>
  <Title> Биология и экология </Title>
</relation>
<relation>
  <relationType> NT </relationType>
  <local-control-number>341.33.23.19.07.17.11.13
    </local-control-number>
  <termQualifier>341.33.23.19.07.17.11.13</termQualifier>
  <Title>Физиология и биохимия</Title>
</relation>
<relation>
  <relationType> NT </relationType>
  <local-control-number>341.33.23.19.07.17.11.15
    </local-control-number>
  <termQualifier> 341.33.23.19.07.17.11.15 </termQualifier>
  <Title>Медико-ветеринарное значение и меры борьбы</Title>
</relation>
</zthes>

```

Аналогичное отображение можно осуществить и для записей БД тезаурусов [76–80].

Согласно схеме *Zthes* поиск в БДТК должен осуществляться с использованием поисковых атрибутов *XD-1* [53], *Utility* [81] и *Zthes-1* [74]. Следующая таблица устанавливает связь поисковых атрибутов с элементами схемы данных:

AttrSet	Type	Value	Search For	Name
Utility	1	4	termID	local number
xd-1	1	1	termName	Title
zthes-1	1	1	termQualifier	
utility	1	11	all elements	anywhere
zthes-1	1	3	thesAdmin	
zthes-1	1	4	relatedTermID	
Utility	1	3	termLanguage	language
xd-1	1	4	termNote	description
zthes-1	1	2	termType	
Utility	1	1	termCreatedDate	Record date

Таблица 4.6: Связь поисковых атрибутов и элементов схемы данных *Zthes*

4.5 *Z+SQL* – доступ к реляционным СУБД

В качестве систем хранения и управления данными сегодня чаще всего применяются системы реляционного типа на основе различных *SQL*-серверов (*Oracle*, *Informix*, *MS SQL Server*, *IBM DB2* и т.п.). Используемая в них реляционная модель данных в основе содержит совсем другие принципы, чем иерархическая модель *Z39.50*. Это усложняет, а иногда делает просто невозможным, отображение данных реляционных таблиц в структуры записей *Z39.50*, что несомненно ограничивает область применения как унифицированного подхода к работе с базами данных *Z39.50*, так и использование *SQL*-серверов при построении гетерогенных распределенных информационных систем.

В свете вышеизложенного очень интересной выглядит идея введения в стандарт *Z39.50* возможностей работы с реляционными данными без нарушения общей идеологии (глобальные схемы данных, абстрактные поисковые атрибуты, унифицированный сетевой протокол и т.д.). В феврале 2000 года эта идея воплотилась в расширение стандарта *Z39.50* – *Z+SQL*, которое описывается документом [82]. *Z+SQL* включает в себя следующие дополнения к стандарту *Z39.50*:

- Новый механизм запросов: *Type-104* (*EXTERNAL*, *OID Z.16.1*).
- Новый формат внешнего представления (RecordSyntax): *SQL-RS* (*EXTERNAL*, *OID: Z.5.111*).
- Дополнительные диагностические сообщения.
- Дополнения формата внешнего представления *Explain*.

Профиль *Z+SQL* подразумевает использование запросов *Type-104* в двух вариантах: прямом и абстрактном:

```
--SQL Query Definition
SQLQuery ::= SEQUENCE {
  abstractDatabaseFlag [0] BOOLEAN OPTIONAL,
  queryExpression      [1] IMPLICIT InternationalString}

-- as defined in the SQL Standard [ISO/IEC9075]
-- with/without Z39.50 schema abstraction extension
-- as specified by the flag
```

Прямой вариант запроса (*abstractDatabaseFlag* = 0) – обычный запрос *SQL* в терминах реальных названий таблиц и полей. Например,

```
select title from collection where title like '%stamp%';
```

Допускается использование именованных результирующих наборов. Так, если в результате выполнения предыдущего запроса был создан набор Q1, запрос

```
select title from Q1 where title like '%stamp%';
```

будет выполнен только на множестве Q1.

Абстрактный вариант запроса (*abstractDatabaseFlag* = 1) отличается от прямого тем, что в нем указываются не реальные названия таблиц и полей, а *OID* и элементы схемы данных. Например, при отображении таблиц предыдущего примера на схему *GILS* предыдущий запрос можно представить в виде

```
select [(2,1)] from [1.2.840.10003.13.2, 1.2.840.10003.3.5]
       where [4] like '%stamp%';
```

где в квадратных скобках указаны:

1.2.840.10003.13.2	– <i>OID</i> схемы <i>GILS</i>
1.2.840.10003.3.5	– <i>OID</i> набора атрибутов <i>GILS</i>
(2,1)	– выбираемый элемент <i>title</i> схемы <i>GILS</i>
4	– поисковый атрибут <i>USE</i> = 4 – <i>title</i>

В абстрактном запросе можно ссылаться на наборы элементов, например

```
select *.brief from [1.2.840.10003.13.2, 1.2.840.10003.3.5]
       where [4] like '%stamp%';
```

и использовать именованные результирующие наборы. Таким образом, абстрактный вариант запроса совмещает синтаксис *SQL* с абстрактным представлением данных *Z39.50*.

Формат представления *SQL-RS* содержит два необязательных блока: блок описания данных (метаданные) и блок данных

```
SQL-Result ::= SEQUENCE {
  tableDescriptor [0] IMPLICIT SQLTableDescriptor OPTIONAL,
  listOfResultValues [1] IMPLICIT SEQUENCE OF SQLRowValue
                                OPTIONAL}
```

В блоке метаданных (*tableDescriptor*) определяется структура результирующей таблицы: название, имена полей, тип данных и т.п., в блоке данных (*listOfResultValues*) содержится собственно таблица данных. Допустимые в *SQL-RS* типы данных соответствуют типам данных *SQL-1999* [83–87], включая коллекции и структуры.

```
SQLValue ::= SEQUENCE {
  dataItem CHOICE {
    characterItem [1] OCTET STRING,
    numericItem [2] INTEGER,
    decimalItem [3] INTEGER,
    integerItem [4] INTEGER,
    smallIntItem [5] INTEGER,
    floatItem [6] REAL,
    realItem [7] REAL,
    doublePrecisionItem [8] REAL,
    dateTimeItem [9] InternationalString,
    intervalItem [10] InternationalString,
    varcharItem [12] OCTET STRING,
    booleanItem [13] BOOLEAN,
    bitItem [14] BIT STRING,
```


bitVarItem	[15] BIT STRING,
udtItem	[17] SEQUENCE OF SQLValue,
udtLocator	[18] OCTET STRING,
rowItem	[19] SEQUENCE OF SQLValue,
refItem	[20] OCTET STRING,
collectionItem	[21] SEQUENCE OF SQLValue,
collectionLocator	[22] OCTET STRING,
bLOBItem	[30] OCTET STRING,
bLOBLocator	[31] OCTET STRING,
cLOBItem	[40] OCTET STRING,
cLOBLocator	[41] OCTET STRING,
resultSetItem	[50] SEQUENCE OF SQLValue,
resultSetLocator	[51] OCTET STRING } OPTIONAL,
indicator	[50] IMPLICIT INTEGER { sqlnull (1), sqlempty (2), sqldefault (3) } OPTIONAL }

Последнее позволяет использовать вложенные таблицы для более адекватного отображения реляционной схемы данных *SQL* на иерархическую схему данных *Z39.50*.

Таким образом, *Z+SQL* позволяет на уровне запросов и формата внешнего представления данных объединить две различные идеологии доступа к базам данных – *SQL* и *Z39.50* [88–92]. Используя этот профиль можно

- организовать обмен данными между *SQL*-серверами различных производителей;
- обеспечить унифицированный доступ клиентов к реляционным базам данных без применения различных драйверов типа *ODBC*, шлюзов и другого программного обеспечения промежуточного слоя;
- интегрировать реляционные базы данных в рамках гетерогенных распределенных информационных систем.
- упростить интеграцию *Z39.50* и технологий *CORBA*, *.NET* и *J2EE*.

Следует заметить, что многие проблемы построения гетерогенных информационных систем были бы сняты, если бы производители коммерческих *SQL*-серверов договорились между собой об едином протоколе обмена данными. В этом смысле *Z+SQL* – хорошая платформа для таких начинаний, т.к. является международным стандартом. К сожалению, на сегодняшний день поддержка *Z+SQL* отсутствует как со стороны производителей *SQL*-серверов, так и со стороны производителей серверов *Z39.50* (см. [93, 94]).

4.6 Библиографические информационные системы

Человеческое сообщество развивалось таким образом, что основные информационные ресурсы столетиями собирались в специализированных центрах – библиотеках. На сегодняшний день именно в традиционных библиотеках сосредоточены основные информационные ресурсы общества. Эти ресурсы как правило не находятся в машиночитаемом виде, однако тенденция перевода этих ресурсов в машиночитаемый вид существует.

Если перевод библиотечных фондов в машиночитаемый вид является длительным и дорогостоящим процессом, то перевод в машиночитаемый вид биб-

лиотечных каталогов, т.е. вторичных информационных ресурсов или метаданных, носит массовый характер. При этом повсеместно в библиотеках создаются электронные каталоги как машиночитаемые аналоги традиционных библиотечных каталогов. Наличие электронных каталогов в традиционных библиотеках и организация доступа к ним через Интернет является предпосылкой для построения распределенных библиотечных информационных систем (РИБС).

4.6.1 Специфика РИБС

Библиотечные информационные системы обладают рядом специфических особенностей, некоторые из которых заключаются в следующем:

- Наличие достаточно больших массивов вторичной информации – электронных каталогов библиотечных фондов.
- Организация ведения электронных каталогов в соответствии со специфическими правилами каталогизации, которые, во-первых, являются национальными, т.е. разными в разных странах, и, во-вторых, ориентированы, как правило, не на электронные каталоги, а на каталоги традиционные, т.е. фактически регламентируют правила ведения библиографических карточек. Последнее существенно влияет на структуры записей в электронных каталогах и совсем не благоприятным образом.
- Широкое использование различных *MARK*-форматов из семейства *ISO-2709* (*USmarc*, *RUSmarc* и т.п.) как способа стандартизации библиографических записей. Более того, наблюдается явная фетишизация форматов *ISO-2709* в мировом библиотечном сообществе и насильственное внедрение его в области, которые отличаются от представления библиографической карточки. Ярким примером последнего является представление записей тезаурусов и рубрикаторов в этих форматах. И если для США с их большим объемом внедренных систем на основе *USmarc* это можно как-то объяснить, то внедрение представления тезаурусов и рубрикаторов в *RUSmarc* в России иначе как инертностью или недальновидностью людей, принимающих подобные решения, назвать трудно.
- Широкий спектр используемого программного обеспечения (СУБД) для управления электронными каталогами.

Последнее утверждение может быть несколько пояснено.

4.6.2 О программном обеспечении

В информационно-библиотечном сообществе России отмечается тенденция использования в качестве базового программного обеспечения для построения информационных систем сравнительно дешевых или свободно распространяемых СУБД целевого назначения, несмотря на то, что функции, выполняемые информационно-библиотечными системами, входящими в состав крупных информационных центров и библиотек, становятся значительно сложнее и затраты на их реализацию возрастают (см. [95–100]).

Такое положение дел обуславливается еще и тем, что при внедрении больших СУБД с целью создания информационно-библиотечных систем возникает ряд трудностей, которые могут быть сведены к следующим проблемам:

- имеющиеся в распоряжении у пользователей СУБД представляют собой достаточно сложные программные комплексы, для эксплуатации которых

требуется наличие высококвалифицированных специалистов и значительных машинных ресурсов;

- пользователи, имеющие опыт создания и эксплуатации систем, разработанных в соответствии с принципами "табличной" организации данных, часто бывают морально неподготовлены к восприятию идей и методов, связанных с организацией и ведением библиографических баз данных [98].

Таким образом, сложность существующих коммерческих СУБД, с одной стороны, предполагает наличие достаточно мощного машинного парка, а с другой стороны требует значительных усилий на их освоение и предъявляет повышенные требования к квалификации пользователей.

Разработанные в России и за рубежом СУБД для построения информационно-библиотечных систем отличаются как структурами хранимых данных, так и функциями, которые они выполняют.

При этом необходимо отметить, что степень их распространения в России нельзя считать неудовлетворительной. По оценкам специалистов из ГПНТБ России [98, 100] информационно-библиотечный рынок России достаточно насыщен. Однако, системы рассмотренные в [100] в основном предназначены для решения задач комплексной автоматизации библиотек.

Как показывает практика, наиболее успешно внедряются СУБД, не предъявляющие жестких требований к уровню квалификации пользователя, доступные по цене и в то же время пригодные для решения практических задач обработки данных в информационно-библиотечных системах [98].

При этом должно соблюдаться соответствие уровня сложности решаемых задач требованиям к обработке данных, предъявляемым в соответствующих библиотеках и информационно-библиотечных центрах. Установившиеся методы обработки информации, а также общий уровень внедрения автоматизации информационно-библиотечных процессов в организациях объективно определяют эти требования. Таким образом, хорошо применяются системы:

- не отличающиеся особой сложностью;
- характеризующиеся ориентацией на массовое применение в библиотеках и информационных центрах;
- доступные по цене или свободно распространяемые.

Очевидно, что при применении той или иной СУБД должно четко соблюдаться соответствие между набором средств используемой системы и спецификой решаемых задач. Уровень сложности применяемой системы должен соответствовать уровню сложности конкретного применения.

Очевидно, что при применении той или иной СУБД должно четко соблюдаться соответствие между набором средств используемой системы и спецификой решаемых задач. Уровень сложности применяемой системы должен соответствовать уровню сложности конкретного применения. Поэтому в информационно-библиотечной практике не получают широкого применения системы, обладающие широким диапазоном предоставляемых средств, но сложные в освоении к эксплуатации, например, *Oracle* [101, 102, 102–104], *Informix*, *MS SQL Server* [105], *IBM DB2* и ряд других, поскольку пользователи могут пойти на ограничение средств, предоставляемых ему системой, с целью получения практических результатов в более короткие сроки [98]. Кроме этого, сложность существующих коммерческих СУБД, с одной стороны, предполагает наличие достаточно мощного машинного парка, а с другой стороны, требует значительных усилий на их освоение и предъявляет повышенные требования к квалификации пользователей.

С этой точки зрения целесообразно применение такой СУБД, которая была бы ориентирована на массовое применение в информационно-библиотечной практике и сводила к минимуму различные ограничения. Применение такой системы позволяет сократить затраты, связанные с переходом к широкому использованию библиографических баз данных и непосредственной реализацией информационно-библиотечных систем.

Результаты анализа действующих систем показали, что СУБД, ориентированная на массовое применение в информационно-библиотечной практике, должна соответствовать следующим принципам:

- возможность реализации информационно-библиотечных задач в соответствии с потребностью соответствующих информационно-библиотечных учреждений;
- сочетание необходимого сервиса с простотой освоения и использования системы;
- сравнительно легкий переход от традиционного способа обработки информации к обработке с применением СУБД;
- дополнительные уровни сервиса за счет возможности введения в систему специальных программ и программных комплексов. Исходя из вышеперечисленного, можно сформулировать основные требования, предъявляемые к информационно-библиографическим системам:
- стоимость системы;
- простота архитектуры системы и организации данных;
- небольшой объем программного обеспечения;
- умеренные требования к техническому обеспечению;
- относительно высокая производительность обработки данных, не требующая значительных вычислительных ресурсов, поскольку техническая оснащенность библиотек и информационных центров всегда оставляла желать лучшего;
- формат внутреннего представления данных;
- возможность работы с международными коммуникативными форматами (*ISO-2709*) [40], исходя из специфики структуры библиографических данных;
- независимость данных на уровне записи;
- наличие развитых лингвистических и словарных средств, для обеспечения эффективного поиска в базах данных;
- отсутствие необходимости в реорганизации базы данных при обновлении описания существующей базы данных;
- включение дополнительных средств программирования для написания проблемных программ;
- возможность работы в сетевом режиме;
- оснащенность дополнительными сервисными программными средствами;
- эффективность эксплуатации системы;

Обзор рынка информационно-поисковых систем показал, что рынок информационно-библиотечных СУБД в России покрывают три системы: *ИРБИС*, *МАРК-SQL* и *Liber*, поскольку именно эти системы в настоящее время наиболее полно удовлетворяют перечисленным требованиям [98] и могут быть эффективно использованы в информационно-библиотечной практике.

4.6.3 Модели интеграции ресурсов

Если вернуться к построению интегрированных библиотечно-информационных ресурсов, можно выделить три модели интеграции:

Сводный каталог – представляет собой некую базу данных, в которую включаются записи из разных каталогов. Дублируемые записи при этом удаляются, а информация об источнике записи фиксируется в специальных полях. Создание сводного каталога достаточно трудоемко, т.к. связано с дополнительной, причем интеллектуальной, обработкой записей. Следствием этого является тот факт, что актуальность сводных каталогов оставляет желать лучшего.

Централизованное хранение каталогов различных поставщиков в различных базах данных и, возможно, объединение их в одну логическую базу данных. Эта модель является самой простой по реализации и повсеместно применяется на практике. Однако необходимость в синхронизации данных, которая осуществляется через их реплицирование, не дает полной уверенности в актуальности представленной информации.

Распределенные каталоги не требуют синхронизации данных и, следовательно, всегда актуальны. Именно эта интеграционная модель соответствует РИБС. Для обеспечения прозрачного доступа к распределенным каталогам принято использовать протокол *Z39.50*.

В заключение следует особо отметить, что протокол *Z39.50* первоначально создавался для потребностей именно библиотечного сообщества. Поэтому в нем явным образом присутствует специфика информационно-библиотечных систем.

4.7 Доступ к первичной информации

Как отмечалось выше, метаданные являются вторичным информационным ресурсом. И несмотря на то, что интеграция ресурсов в рамках распределенных систем возможна, как правило, только на уровне метаданных, основной интерес представляет собой несомненно ресурс первичный. Поэтому метаданные в любых схемах должны содержать ссылки на первичный ресурс. При наличии в записи метаданных такой ссылки извлечение первичной информации становится технической задачей организации пользовательских интерфейсов.

Таким образом, сценарий доступа к первичной информации в РИС выглядит следующим образом:

1. Пользователь производит поиск по распределенным базам метаданных.
2. Пользователь просматривает найденные записи и отбирает нужные.
3. В отобранных записях пользователь переходит по ссылке (чаще всего гипертекстовой) на просмотр (или заказ копии) первичного источника.

Именно по такому сценарию сегодня функционируют так называемые электронные библиотеки.

Глава 5

Сервер *ZooPARK* в РИС

5.1 История создания сервера *ZooPARK*

Сервер *ZooPARK* был создан в ОИГТМ СО РАН как многофункциональный сервер *Z39.50* в 1998 году.

Первая версия сервера был ориентирована на работу с базами данных *CDS/ISIS*, ее появление явилось следствием целой серии работ, проводимых в ОИГТМ СО РАН, по обеспечению сетевого доступа к базам данных *CDS/ISIS* [106–108]. Первая версия не обладала многими функциональными качествами последующих версий, но обеспечивала устойчивую работу по протоколу *Z39.50* v.3.

Начиная с версии 2 (конец 1999 года) сервер *ZooPARK* приобрел модульную структуру и унифицированный интерфейс провайдера данных. Последнее позволило локализовать специфику каждой из используемых СУБД в отдельных динамически загружаемых модулях с единым интерфейсом со стороны ядра сервера. Были созданы динамические провайдеры данных для *CDS/ISIS (Z-ISIS)*, *MS SQL Server (Z-MSSQL)*, провайдер удаленного сервера *Z39.50 (Z-REMOTE)*, провайдер данных *MS ADO (Z-MSADO)* и др. Сервер *ZooPARK* версии 2 был инсталлирован во многих библиотечных центрах России и стал самым массовым сервером *Z39.50* в России, в некоторых из центров эта версия функционирует и сегодня.

Версия 3 сервера *ZooPARK* (2002 год) стала обладать возможностью параллельной обработки поисковых запросов в разных базах данных, что существенно уменьшило время обработки запросов в распределенных системах.

Наконец, сервер *ZooPARK* версии 4 (июнь 2003 года) [38] стал включать не только сервер *Z39.50*, но и сервер *WWW*, работающий в том же самом порту, что и сервер *Z39.50*. Это позволило реализовать в теле *ZooPARK* встроенный шлюз *Z39.50-HTTP (Z-GW)*, обработку запросов *SRU* и модуль обработки запросов *SOAP/SRW* [17].

Из всех функционирующих в России серверов *Z39.50* сервер *ZooPARK* вместе со своим клоном *Z-IRBIS* являются самыми распространенными (50%).

5.2 Основные характеристики

Сервер *ZooPARK* предназначен для эксплуатации в качестве базового сервера *Z39.50 (ISO-23950)* распределенной информационной системы общего назначения на различных программно-аппаратных платформах. *ZooPARK* включает встроенный в себя *WEB*-сервер для доступа к встроенному шлюзу *Z39.50-HTTP* и к ресурсам *Z39.50* по протоколу *SOAP (SRW/SRU)*.

Установленные в узлах информационной системы однотипные серверы *ZooPARK* могут образовывать кластеры с перераспределением поисковых запросов. Серверы других производителей также могут быть включены в подобные кластеры, но с ограниченными возможностями.

Сервер поддерживает протокол *Z39.50* версии 3 над *TCP/IP*. В текущей версии сервера реализованы следующие функции [38]:

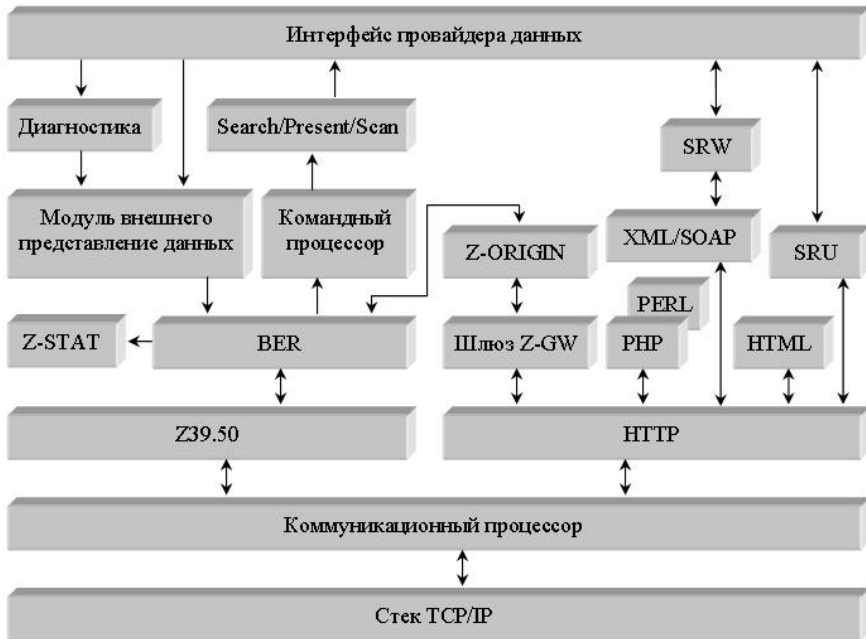
- Поиск, в том числе многопоточный в многобазовом окружении
- Представление данных в различных форматах (*SUTRS, GRS-1, USmarc, UNIMARC, RUSmarc, HTML, RTF*)
- Поддержка различных схем данных (*GILS, GEO, CIMI, ZTHES* и др.)
- Работа со словарем
- Создание и удаление именованных сеансовых наборов данных на сервере
- Расширенный сервис – добавление, модификация, удаление записей из баз данных
- Работа (*search, present, scan*) в многобазовом окружении
- Работа с логическими группами баз данных
- Контроль и ограничение доступа к базам данных по IP-адресам
- Аутентификация пользователей
- Поддержка различных кодовых таблиц (*DOS, WIN, ISO, MAC*) для кириллических текстов
- Поддержка *Explain*
- Поддержка *HTTP 1.0/1.1* в одном порту с *Z39.50*
- Поддержка *SOAP (SRW/SRU)*
- Поддержка *WEB*-сервером *PHP* и *PERL*.
- Поддержка встроенного шлюза *Z39.50-HTTP*

5.3 Архитектура сервера ZooPARK

Для *Win32* сервер *ZooPARK* реализует модель многопоточного приложения *Win32* (один клиент – один поток), для *UNIX* – модель как однопоточного (один клиент – один процесс), так и многопоточного приложения.

Программа сервера выполняется для *Windows NT/2000/XP/2003* как сервисная служба или как отдельное консольное приложение *Win32*, а для *UNIX* – как отдельное консольное приложение или как *daemon*, запускаемый через стартовые системные скрипты или через *inetd*. Конфигурирование сервера производится через конфигурационный файл сервера, каждой базы данных – через соответствующий настроечный файл.

Основные компоненты сервера представлены на Рис. 5.1, где явно выделены модули, реализующую обработку запросов *Z39.50* и *HTTP*. На Рис. 5.2 представлена схема доступа к данным через общий интерфейс провайдера данных.

Рис. 5.1: Основные блоки сервера *ZooPARK*

5.4 Модульность и расширяемость

Как уже отмечалось, сервер *ZooPARK* организован по модульному принципу и состоит из ядра и набора динамически загружаемых библиотек, часть из которых требует явной загрузки. Явная загрузка динамических библиотек позволяет не фиксировать список требуемых библиотек в момент запуска основного процесса сервера, т.е. сервер "не знает" какие динамические библиотеки ему понадобятся в процессе работы. Это касается в первую очередь библиотек динамических провайдеров данных и дополнительных модулей Web-сервера.

В качестве отдельных модулей сервера *ZooPARK* выступают следующие:

- Ядро, обеспечивающее функции *Z-service provider* и сервера *HTTP*.
- Динамически загружаемая библиотека основных функций обработки *APDU*.
- Динамически загружаемые библиотеки с функциями *Z-data provider*. Для каждой СУБД требуется своя динамическая библиотека. Загрузка этих модулей явная. Взаимодействие с ядром происходит через документированный интерфейс, что позволяет расширять список поддерживаемых СУБД созданием дополнительных модулей.
- Динамически загружаемая библиотека, реализующая функции клиента *Z39.50 (Z-origin)*. Используется шлюзом *Z39.50-HTTP* и провайдером данных удаленного сервера *Z39.50*.

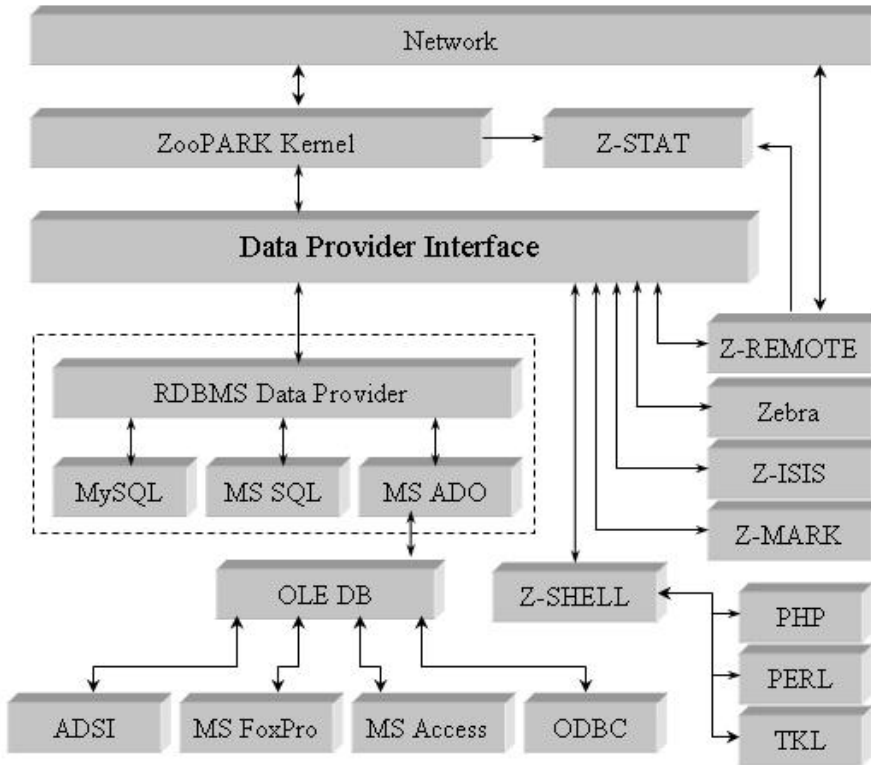


Рис. 5.2: Доступ к данным в сервере ZooPARK

- Динамически загружаемая библиотека шлюза *Z39.50-HTTP*. Загрузка явная.
- Динамически загружаемая библиотека вызова внешних консольных приложений (*PHP*, *PERL*, *TCL*, *SH*, *CMD* и т.п.). Загрузка явная.

На сегодняшний день сервер *ZooPARK* может комплектоваться следующими динамически загружаемыми библиотеками провайдеров данных:

Z-ISIS – провайдер обеспечивает работу с данными СУБД *CDS/ISIS* без привлечения последней.

Z-REMOTE – провайдер удаленного сервера *Z39.50* обеспечивает прозрачный доступ к другим серверам по протоколу *Z39.50*, что позволяет реализовать механизмы переадресации запросов в распределенной информационной системе.

Z-MSSQL – провайдер данных для серверов *MS SQL* обеспечивает доступ к реляционным базам данных серверов *MS SQL*.

Z-MySQL – провайдер данных для серверов *MySQL* обеспечивает доступ к реляционным базам данных серверов *MySQL*.

Z-MSADO – провайдер данных для интерфейсов *MS ADO* обеспечивает доступ к данным через интерфейсы *MS ADO* и *ODBC* (только для *Win32*).

Z-MMARK – провайдер для данных АИБС *MARK-SQL* (Информсистема) на основе *MS SQL Server*.

Z-AMARK – провайдер для данных АИБС *MARK-SQL* (Информсистема) на основе *MS OLE DB* и *ODBC* (только для *Win32*).

Z-SHELL – провайдер данных, обеспечивающий взаимодействие ядра *ZooPARK* с внешними консольными приложениями, например, скриптовыми приложениями *PHP*, *PERL* и т.п.

Модульность сервера *ZooPARK* позволяет расширять его функциональные возможности без перекомпоновки ядра. Эта возможность используется пользователями, эксплуатирующими сервер. В частности, зафиксированы факты создания дополнительных провайдеров данных для СУБД *Oracle*, *Sybase*, *LiberMedia* и успешная их эксплуатация в РИС.

5.5 Расширения Z39.50 сервера ZooPARK

Для адаптации *Z39.50* к задачам построения РИС сервер *ZooPARK* наряду с глобально стандартизованными объектами *Z39.50* оперирует с локально определенными объектами. В Табл. 5.1 приведен список таких объектов (см. [38]), при этом использован *id=155* глобального реестра разработчиков.

Таблица 5.1: Локальные объекты сервера *ZooPARK*

Класс	OID	Описание
ATTSET	Z.3.1000.155.1	ZSTAT-attset – набор поисковых атрибутов для статистической информации о работе РИС по протоколу <i>Z39.50</i>
RECSYN	Z.5.1000.155.1	Rtf – EXTERNAL, для формата представления данных RTF
USERINFO	Z.10.1000.155.3	SearchResultHits – EXTERNAL, определяющий структуру информации о результатах поиска в многобазовом окружении
USERINFO	Z.10.1000.155.4	ModulesInfo – EXTERNAL, определяющий Explain структуру данных о модулях сервера <i>ZooPARK</i>
USERINFO	Z.10.1000.155.5	RemoteInfo – EXTERNAL, определяющий структуру данных о маршруте прохождения запроса в РИС
SCHEMA	Z.13.1000.155.1	UIGGM-schema – схема данных для записей, повторяющая семантику МЕКОФ
SCHEMA	Z.13.1000.155.2	UIGGMe-schema – схема данных для записей, повторяющая семантику МЕКОФ
SCHEMA	Z.13.1000.155.4	PERSONS-schema – схема данных для представления информации о персонах
SCHEMA	Z.13.1000.155.5	ZSTAT-schema – схема данных для представления статистической информации о работе РИС по протоколу <i>Z39.50</i>
TAGSET	Z.14.1000.155.4	PERSONS-tagset – набор меток для схемы данных PERSONS-schema
TAGSET	Z.14.1000.155.5	ZSTAT-tagset – набор меток для схемы данных ZSTAT-schema

5.6 Обеспечение ”распределенности” сервером ZooPARK

Технологические решения, описанные в разделах 3.3 и 3.4, реализованы в сервере ZooPARK.

5.6.1 Объединение баз данных

Сервер ZooPARK позволяет объединять различные базы данных в логические группы с присвоением им уникальных имен. Все операции над логической группой (поиск, извлечение данных, просмотр индексов) для пользователя не отличаются от операций над физическими базами данных. При этом в логической группе могут присутствовать ссылки на базы данных, расположенные на других серверах Z39.50.

5.6.2 Переадресация запросов

Сервер ZooPARK переадресовывает запросы к базам данных, которые зарегистрированы как ссылки на другие серверы Z39.50, на соответствующие серверы. Для пользователя база данных, представленная в виде ссылки, не отличается от обычной базы данных. Если ссылка указывает на базу данных, которая тоже является ссылкой, запрос перенаправляется дальше. Конечной точкой маршрута запроса может выступать не только ZooPARK, но любой другой сервер Z39.50.

5.6.3 Контроль маршрутизации запросов

Сервер ZooPARK в момент инициализации сеанса всегда сохраняет информацию об IP-адресе непосредственного клиента в сеансовых переменных. Эта информация используется сервером при переадресации запросов этого клиента, а именно, сохраненный в момент инициализации сеанса IP-адрес дописывается в маршрут запроса при его переадресации провайдером данных Z-Remote. Таким образом, при многократной переадресации маршрут запроса содержит полный список серверов, через которые он прошел. Самым верхним адресом является адрес истинного клиента. Информация о маршруте запроса помещается в поле *otherInfo APDU initRequest* в локально определенную структуру типа *EXTERNAL (OID Z.10.1000.155.5)*. Анализ сервером маршрута запроса позволяет определить в нем превышение допустимого количества переадресаций, задаваемое в конфигурационных файлах, и наличие петель.

Следует также заметить, что маршрут запроса формирует не только сервер через провайдер данных Z-REMOTE, но и встроенный шлюз Z39.50-HTTP. В этом случае самым верхним адресом является адрес WEB-клиента.

5.6.4 Пополнение, модификация и удаление информации

Ядро сервера ZooPARK обрабатывает запросы расширенного сервиса и передает их соответствующему провайдеру данных. Таким образом, ответственность за исполнение функций расширенного сервиса возлагается на провайдер данных. Однако следует отметить, что реализация существующих провайдеров данных,

как правило, не позволяет производить обработку запросов расширенного сервиса, все провайдеры данных при этом фиксируют состояние штатной ошибки в соответствии с диагностикой *Vib-1*.

5.6.5 Учет ресурсов РИС

Сервер *ZooPARK* поддерживает базу данных о своих ресурсах *IR-Explain-1*. Эта база данных представляет собой набор файлов в формате *XML*, проиндексированных при помощи прилагаемой программы. Структура *XML*-файлов определяется структурой записей *Explain* в соответствии с *Z39.50*. Некоторые элементы записи *Explain* формируются автоматически в момент извлечения записи.

В текущей версии сервер *ZooPARK* не поддерживает распределенную базу данных *Explain*, но работы в этом направлении ведутся.

5.6.6 Сбор статистической информации о работе РИС

Сервер *ZooPARK* может собирать информацию о своей работе в текстовых *log*-файлах (информация о работе подсистем с несколькими уровнями детализации, протоколирование работы *WEB*-сервера, протоколирование всех входящих и исходящих *APDU*). Однако возможен и другой режим работы – передача этой информации по сети на специально сконфигурированный сервер *ZooPARK* для сохранения ее в реляционной базе данных (входящие и исходящие *APDU*) с возможностью доступа к ней по *Z39.50*. Передача информации осуществляется асинхронно по протоколу *HTTP* (метод *PUT*). Асинхронность передачи информации обеспечивает минимизацию времени отклика для клиентов и не вносит каких-либо заметных задержек.

Для обеспечения доступа к статистической информации для сервера *ZooPARK* локально определены три объекта:

Схема данных <i>ZSTAT-schema</i>	Z.13.1000.155.5
Набор меток <i>ZSTAT-tagset</i>	Z.14.1000.155.5
Набор поисковых атрибутов <i>ZSTAT-attset</i>	Z.3.1000.155.1

Информация о запроотоколированных *APDU* доступна по *Z39.50* в форматах *SUTRS*, *GRS-1*, *XML*. Следует отдельно заметить, что, во-первых, сервер *ZooPARK* протоколирует не только *APDU*, обработанные собственно сервером (*Z-service provider*), но и *APDU*, обработанные провайдером *Z-Remote* (переадресация запросов) и встроенным шлюзом *Z-GW*. При этом "происхождение" *APDU* фиксируется в отдельном поле записи *ZSTAT*. Во-вторых, в отдельное поле записи *ZSTAT* заносится идентификатор сервера *ZooPARK*, ее породившего. Для РИС это позволяет как вести централизованную базу статистики работы всех серверов *Z39.50*, так и формировать распределенную базу статистики в виде логического объединения локальных.

Для обеспечения групповых операций (только в случае хранения информации в реляционных СУБД) над базой данных *ZSTAT* (суммирование, подсчет количества, вычисление средних и т.д.) сервер *ZooPARK* использует систему запросов *Z-SQL* [82, 94] и предоставляет записи в табличном формате *RS-SQL*. Это позволяет возложить обработку групповых операций на реляционные СУБД, используя, с одной стороны, выразительные средства языка *SQL*, а, с другой стороны, остаться в рамках стандарта *Z39.50*.

Пользовательские интерфейсы доступа к групповым операциям над базой данных *ZSTAT* реализованы в шлюзе *Z-GW* (см. Рис. 5.3) и доступны по адресу:

http://z3950.uiggm.nsc.ru:210/zgw/zstat_1.htm.

5.6.7 Аутентификация пользователей РИС

Сервер *ZooPARK* позволяет производить аутентификацию пользователей двумя способами:

1. Локальная аутентификация по имени и паролю пользователя. Вся информация о пользователе при этом находится в локальных настроечных текстовых файлах сервера *ZooPARK*.
2. Аутентификация через *LDAP*. Вся информация о пользователе при этом находится в каталогах указанного сервера *LDAP*.

5.6.8 Ограничение доступа к ресурсам РИС

Сервер *ZooPARK* обеспечивает ограничение доступа к базам данных по группам имен пользователей и по группам *IP*-адресов клиентов. При работе через встроенный шлюз *Z39.50-HTTP* и при переадресации запросов на другие серверы в РИС сервер *ZooPARK* контролирует маршрут запроса, что позволяет всегда определить истинный *IP*-адрес клиента (см. раздел 5.6.3). Каждая база данных, доступная через сервер *ZooPARK*, содержит свои списки групп для доступа.

5.6.9 Мониторинг целостности РИС

В текущей версии сервера *ZooPARK* механизмы мониторинга целостности РИС не реализованы.

5.7 Оптимизация распределенных запросов

При работе сервера *ZooPARK* в составе РИС очень важным становится вопрос о минимизации времени выполнения запросов пользователей. Как отмечалось в разделе 3.4, на это время влияют многие факторы и только некоторые из них связаны с архитектурой серверов. Тем не менее, существуют общие правила, выполнение которых позволяет существенно уменьшить время ожидания выполнения запроса в РИС. При этом, естественно, серверы РИС должны быть способными удовлетворить этим правилам. Именно эти свойства сервера *ZooPARK* обсуждаются ниже.

5.7.1 Распараллеливание поисковых запросов

Поиск – наиболее частая операция, выполняемая в РИС, но именно она является наиболее длительной по времени выполнения особенно при многобазовом распределенном поиске.

Общепринятый способ уменьшения времени многобазового поиска – организация параллельного поиска. Это распараллеливание наиболее эффективно в случае, когда базы данных находятся на разных серверах в РИС, т.к. в этом случае поисковый запрос обрабатывает одновременно группа серверов.

Сервер *ZooPARK* может быть настроен на параллельный поиск в базах данных. При этом распараллеливание происходит не для логических, а для физических баз данных, для каждой из которых порождается отдельный поток в рамках

Статистика работы сервера ZooPARK

Сервер: z3950.uiggm.nsc.ru БД: statg_mssql. Записей 22 (SQL-RS). [Покажи SQL](#)

type	way	dbName	errorCode	countItem	byteCount	recordCount
0	S	IR-Explan-1	0	4	5393	4
0	S	ch	0	8	22994	8
0	S	geocat	0	8	6453	10
0	S	egg	0	2	4532	2
0	S	greffh	0	8	93468	52
						3
						194
						110
						6
						2
						59
						8
						10
						2
						52
						23
						14
						86
						24
						3
						194
						110

Статистика работы сервера ZooPARK

Сервер: z3950.uiggm.nsc.ru Порт: 210
 База данных статистики: statg_mssql
 Статистика для сервера: z3950.uiggm.nsc.ru:210

Источник данных: Дата события
 Сервер Z39.50: от 20031207 00:00:00
 Провайдер Z-Remote: до 20031231 23:59:59
 Шлюз Z-GW:

Общая статистика: Info Help

Init: Request Response
 Search: Request Response
 Present: Request Response
 DeleteResultSet: Request Response
 ResourceReport: Request Response
 Scan: Request Response

[Возврат на начало](#)

Copyright © ОИГТМ.СО.РАН, 2000-2003

Powered by ZooPARK
 ZooPARK Version 4.05 Copyright © ОИГТМ.СО.РАН, 2000-2003

Рис. 5.3: Доступ к статистической информации через шлюз Z-GW сервера ZooPARK

потока или процесса сеанса с асинхронным завершением. Под физической базой данных здесь понимается или локальная база данных или ссылка на базу данных другого сервера.

Встроенный в *ZooPARK* шлюз *Z-GW* при поиске в списке баз данных также организует отдельный поток для каждой базы данных.

5.7.2 Обеспечение режимов переадресации

Сервер *ZooPARK* реализует три режима переадресации запросов в РИС в соответствии с принципами, изложенными в разделе 3.3.2. Выбор текущего режима переадресации осуществляется через конфигурационный файл провайдера данных *Z-REMOTE*.

При работе в режиме 2 (см. раздел 3.4) сервер *ZooPARK* может быть сконфигурирован на параллельную асинхронную инициализацию вторичных сеансов, что существенно уменьшает время инициализации первичного сеанса [30].

5.7.3 Контроль петель переадресации и инициализации

Для защиты от избыточной переадресации, инициализационных петель и петель переадресации в сервере *ZooPARK* предусмотрены ограничения, задаваемые в конфигурационном файле провайдера данных *Z-REMOTE* как на максимально допустимое количество переадресаций запроса, так и на максимально допустимое количество петель.

5.8 Работа сервера *ZooPARK* в РИС

Сервер *ZooPARK* предназначен для работы в распределенных информационных системах на основе протоколов *Z39.50* и *HTTP*. РИС на основе серверов *ZooPARK* удовлетворяют критериям, изложенным в разделе 2.2:

Работа с распределенными данными – сервер *ZooPARK* функционирует на различных программно-аппаратных платформах. Обработка данных при этом в *ZooPARK* логически отделена от сервисных функций *Z39.50* и *HTTP*, что позволяет разделить программные модули *service provider* и *data provider* и оформить последние в виде динамически загружаемых библиотек. Для каждой используемой СУБД можно создать свой *data provider* и включить его в состав сервера через документированный интерфейс. Сервер *ZooPARK* содержит *data provider*, который обеспечивает прозрачную переадресацию запросов на другие серверы *Z39.50*.

Логическая группировка данных – сервер *ZooPARK* может использовать в качестве имен баз данных их групповые имена. Все запросы при этом будут обрабатываться на всем списке баз данных, скрытым под групповым именем. В логической группе могут присутствовать различные базы данных, в том числе находящиеся на других серверах.

Абстрактная модель данных – сервер *ZooPARK* обрабатывает данные на основе абстрактной модели в полном соответствии с идеологией *Z39.50*.

Абстрактная система запросов – сервер *ZooPARK* поддерживает стандартную для *Z39.50* систему запросов *RPN*, основанную на стандартизованных

наборах поисковых атрибутов. Дополнительно могут поддерживаться запросы *CQL* [109], *Z-SQL* [82] и запросы в синтаксисе используемых СУБД (Type 0).

Метаинформация о ресурсах – сервер *Zoo*PARK обеспечивает доступ к базе данных *IR-Explain-1* в соответствии с требованиями *Z39.50*. При доступе к этой базе данных многие элементы записи *Explain* заполняются автоматически. Для конечных пользователей сервер *Zoo*PARK в составе шлюза *Z-GW* содержит надстройку просмотра информации *Explain* через *WEB*.

Разграничение доступа – сервер *Zoo*PARK обеспечивает ограничение доступа к базам (см. раздел 5.6.8).

Учет и контроль – сервер *Zoo*PARK собирает информацию о своей работе как в текстовых log-файлах (информация о работе подсистем с несколькими уровнями детализации, протоколирование работы *WEB*-сервера, протоколирование всех входящих и исходящих *APDU*), так и в специальной реляционной базе данных (входящие и исходящие *APDU*) с возможностью доступа к ней по *Z39.50* (см. раздел 5.6.6).

Открытость – сервер *Zoo*PARK функционирует в соответствии со стандартными протоколами *Z39.50* и *HTTP*.

Связь с другими системами – РИС на основе серверов *Zoo*PARK могут легко интегрировать свои ресурсы с другими системами, т.к., во-первых, функционируют на основе открытых протоколов, и, во-вторых, оперируют данными, представленными в стандартных схемах.

Легкость в общении – сервер *Zoo*PARK включает в себя встроенный шлюз *Z39.50-HTTP*, что позволяет организовать точку доступа в РИС на основе *Z39.50* через *WEB* без привлечения специализированного клиентского программного обеспечения *Z39.50* (см. Рис. 5.9).

5.9 Сервер *Zoo*PARK и метаданные

В соответствии с общей идеологией работы с метаданными в *Z39.50*, сервер *Zoo*PARK поддерживает различные глобально и локально определенные схемы данных. Список поддерживаемых схем данных приведен в Табл.5.2.

В качестве примера использования локально определенной схемы данных *UIGGM* на Рис. 5.9 приведено представление шлюзом *Z-GW* библиографической записи в этой схеме.

5.10 Поддержка *Z-SQL* и *SQL-RS*

Для взаимодействия с реляционными СУБД сервер *Zoo*PARK поддерживает упоминавшиеся выше спецификации *Z+SQL* (см. раздел 4.5).

Для установления соответствий между реальными таблицами и поисковыми атрибутами в *Zoo*PARK используются статические конфигурационные файлы. В этих файлах, которые уникальны для каждой базы данных, присутствует информация о соответствии полей таблиц поисковым атрибутам и элементам схем данных, о типах полей и о характере связей между таблицами (см. раздел 6.2.4). В общем случае структура конфигурационных файлов довольно сложна. Но для ряда приложений подобные файлы можно существенно упростить, применяя хранимые процедуры и псевдотаблицы (*view*).

**Добро пожаловать
в Распределенную Информационную Систему
Сибирского Отделения РАН**

Базы данных для внутреннего использования

Простой поиск Расширенный поиск Поиск для эксперта

Home Page

РЖ Все имеющиеся выпуски
 РЖ "Геология и геофизика" (с 1985 г.)
 РЖ "Химия" (с 1981 г.)
 РЖ "Физика" (с 1995 г.)

Expert Search - Microsoft Internet Explorer

Address: http://z3950.uiggm.nsc.ru:210/zgw/se.htm

Поиск для эксперта Z39.50

Term: geology Elements: Full Format: grs1 Operator: and

Use: 4 Title Relation: Default Truncation: right

Position: Default Structure: Default Completeness: Default

Query: @attr 1=4 @attr 5=1 (geology)

Начать с: 1 Количество: 1 Бланк: Rusmarc - rus

Назад Очистить Добавить Поиск

**Результат поиска по запросу:
@attr 1=4 @attr 5=1 {geology}**

РЖ Геология и геофизика (с 1985 г.)	gg	4420	Просмотр	SRW	SRU
РЖ Физика (с 1995 г.)	fi	19	Просмотр	SRW	SRU
РЖ Механика (с 1985 г.)	mx	44	Просмотр	SRW	SRU
CC Proceedings	proc	1996	Просмотр	SRW	SRU

Всего по запросу найдено: 6479 записей

Рис. 5.4: Доступ к базам данных через шлюз Z-GW сервера ZooPARK

Назад на Поиск Формат: grs1 Схема: UIGGM Полный список << < > >>

Сервер: z3950-3.uiggm.nsc.ru База данных: gg. Вывод: 1 , начиная с 2 из 144781

Сменить представление

Запись 2 из 144781

Индекс: gg01-2
Номер записи: 64303010100000022699530
Авторы:
 Автор: Chen Ming
 Автор: Xie Xian-de
 Список авторов: Chen Ming, Xie Xian-de

Заглавие:
Перевод заглавия: Прогресс в изучении природных минералов высокого давления и их значение для минералогии мантии Земли

Издание:
 Сокр заглавие издания: Gaoxiao dizhi xuebao
 Дата издания: 2000
 Том: 6
 Номер выпуска: 2
 Объем публикации: 121-125
 ISSN: 1006-7493
 Место издания страна: 156

Реферат: Изучение шоковых воздействий в наземных астроблемах и в метеоритах позволило по-новому подойти к изучению природных минералов высокого давления. К настоящему времени в метеоритах установлено наличие таких минералов высокого давления как 'альфа'-PbO[3], постстивовит, (Na, K, Ca) AlSi[3]O[8] (голандит) и полиморфные разновидности апатита высокого давления, которые могут быть выявлены и в мантии Земли, что является большим вкладом в минералогию нашей планеты

Номер выпуска РЖБД: 01
Ключевые слова:
 Ключевое слово: ГЕОЛОГИЯ ОБЩИЕ ВОПРОСЫ
 Ключевое слово: МИНЕРАЛОГИЯ МАНТИИ ЗЕМЛИ
 Список ключевых слов: ГЕОЛОГИЯ ОБЩИЕ ВОПРОСЫ, МИНЕРАЛОГИЯ МАНТИИ ЗЕМЛИ

Вид документа: 203
Дата записи: 20010416
Номер выпуска РЖБД: 01
Индекс УДК: 55.001
Коды рубрики ГРНТИ: 38.01.05
Код рубрики ВИНТИ: 383.01.05
Отр реферат службами: 01.01-08A.2
Оснащение:
Библиографические ссылки: 15

Короткая запись: Chen Ming, Xie Xian-de [Прогресс в изучении природных минералов высокого давления и их значение для минералогии мантии Земли]Gaoxiao dizhi xuebao = Geol. J. China Univ. - 2000. - Vol. 6. - N 2. - P. 121-125. - ISSN 1006-7493

Назад на Поиск << < > >>

Done Internet

Рис. 5.5: Библиографическая запись в схеме UIGGM

Z.13.1	WAIS-schema – схема данных WAIS, используемой в устаревшем стандарте Z39.50 v.1 (1988)
Z.13.2	GILS-schema – см. 4.3.2
Z.13.3	Collections-schema – см. 4.3.5
Z.13.4	Geo-schema – см. 4.3.4
Z.13.5	CIMI-schema – см. 4.3.6
Z.13.6	Update ES
Z.13.7	Holdings-schema
Z.13.8	Zthes-schema – см. 4.4.2
Z.13.1000.81.2	Explain-schema
Z.13.1000.155.1	UIGGM-schema – схема данных для записей, повторяющая семантику МЕКОФ
Z.13.1000.155.2	UIGGMе-schema – схема данных для записей, повторяющая семантику МЕКОФ
Z.13.1000.155.4	PERSONS-schema – схема данных для представления информации о персонах
Z.13.1000.155.5	ZSTAT-schema – схема данных для представления статистической информации о работе РИС по Z39.50

Таблица 5.2: Схемы данных, поддерживаемые сервером ZooPARK

В качестве примера использования *Z+SQL* в сервере ZooPARK можно привести технологию доступа к базе данных статистики работы сервера (см. разделы 3.5.3 и 5.6.6).

5.11 Сервер ZooPARK и другие разработки

ZooPARK – модульный сервер Z39.50 v3 разработки ОИГГМ СО РАН (ID=155) общего назначения. Сервер ориентирован на работу как с библиографическими данными, так и с данными в других схемах (*GILS*, *GEO*, *CIMI* и др.). Поддерживаемые платформы: *Windows NT/2000/2003*, *FreeBSD*, *Linux*, *Solaris-Intel*, *Solaris-Sparc*, *SCO OpenServer*. Сервер построен по модульному принципу с динамической загрузкой провайдеров данных (*CDS/ISIS*, *MS SQL*, *MySQL*, *Zebra*, *Remote Z39.50*, *Oracle* и др.) через документированный интерфейс, что позволяет создавать дополнительные провайдеры данных. Форматы представления: *SUTRS*, *GRS-1*, *USmarc*, *UNIMARC*, *RUSmarc*, *HTML*, *XML*, *RTF*. Поддерживается сервис *Explain*. Функциональные возможности сервера: Поиск, Просмотр, Удаление наборов, Просмотр словаря, Расширенный сервис, Именованные наборы, Переговоры при инициализации, Запросы SQL. В настоящее время сервер ZooPARK является наиболее популярным сервером Z39.50 в России, что подтверждается статистикой использования серверов (см. Табл. 5.3, в которой приводится статистика, собираемая по России на сервере *RUSLAN* [31], и по всему миру на сервере компании *IndexData* [6]).

Z-IRBIS – модульный сервер Z39.50 v3 разработки ОИГГМ СО РАН (ID = 155). *Z-IRBIS* отличается от сервера ZooPARK тем, что он оптимизирован для работы с базами данных *IRBIS* и *CDS/ISIS*. Сервер доступен для плат-

Сервер	Источник	27.03.02	2.03.03	29.01.04
ZooPARK	IndexData	4(1%) из 303	7(2%) из 431	14(2%) из 613
ZooPARK	RUSLAN	14(30%) из 46	13(23%) из 56	14(23%) из 62
Z-IRBIS	RUSLAN	13(28%) из 46	12(21%) из 56	10(16%) из 62

Таблица 5.3: Распространенность серверов ZooPARK и Z-IRBIS

форм *Windows NT/2000/2003* и *Linux*. Сервер *Z-IRBIS* распространяется ГПНТБ России (г. Москва) в составе программного обеспечения *IRBIS*. По распространенности в России занимает второе место после *ZooPARK* (см. Табл. 5.3).

Следует заметить, что по данным вышеуказанных источников количество действующих серверов *ZooPARK* и *Z-IRBIS* в России суммарно превышает количество инсталляций серверов других типов, в том числе и свободно распространяемых.

Ruslan Z39.50 server – Государственный Технический Университет, г. С-Петербург, Центр "Открытые Библиотечные Системы" (ID=148). Декларированные функциональные возможности: Поиск, Извлечение, Удаление наборов, Просмотр словаря, Сортировка, Расширенный сервис, Конкурирующие операции, Именованные наборы, Переговоры при инициализации, Обнаружение дублирования. Формально сервер поддерживает *Explain*, однако уровень этой поддержки настолько примитивен, что сводит на нет всю пользу от нее. Сервер *Ruslan* разработан группой, которая одной из первых в России стала внедрять протокол *Z39.50*. Однако детали архитектуры, очень узкая направленность и высокая стоимость сервера не позволяют клонировать порой хорошие, но слишком специфические решения. Аргументом в пользу последнего тезиса служит статистика, собираемая этой группой разработчиков: сервер *RUSLAN* существуют всего в восьми экземплярах, при этом один экземпляр установлен у самих разработчиков. Нам кажется, что наличие других экземпляров сервера в России обусловлено только политикой явного лоббирования финансирования по проектам ИОО в направлении распространения сервера *RUSLAN*, не будь этих проектов, наверное, сервер существовал бы только в одном экземпляре. Впрочем, не будь проектов ИОО, не было бы всплеска в России активности по применению протокола *Z39.50*...

D3Zserver – сервер *Z39.50 v3* разработки ЗАО *Liber*. Сервер доступен на платформах *Windows NT* и *Linux*. Он предназначен для работы с библиографическими базами данных системы *LiberMedia*. *D3Zserver* предназначен для эксплуатации в среде СУБД ДЗ, как самодостаточный сервер *Z39.50* со встроенным провайдером к системе *LiberMedia* версии 4.1. Сервер имеет модульную структуру и написан на языке *PICK/BASIC* (встроенный в СУБД ДЗ язык высокого уровня), без использования библиотек третьих фирм. Функциональные возможности: Поиск, Просмотр, Просмотр словаря, Переговоры при инициализации. Форматы представления данных: *SUTRS*, *USmarc*, *UNIMARC*, *RUSmarc*, *OPAC*. По данным *RUSLAN* в настоящее время в России активны три таких сервера. Информация о сервере можно найти в [110].

BKS Z3950 Server – ООО Библиотечная компьютерная сеть (МГУ) – [111]. К сожалению, отсутствие четкой опознавательной информации и иденти-

фикатора разработчика не позволяют корректно ссылаться на эту разработку. Однако, в рекламных проспектах *Библиотека-2000* и в литературе этот сервер упоминается. Функциональные возможности: Поиск, Просмотр, Просмотр словаря, Сортировка. На начало 2004 года зарегистрированы две инсталляции этого сервера. Ничего большего о данном сервере авторам неизвестно.

ГИВЦ 'АС Библиотека-3' – GIPER 'A-Elite' – программный комплекс, внедряемый в различные, как правило небольшие, библиотеки при поддержке Министерства культуры РФ. На наш взгляд этот продукт является результатом искусственного симбиоза разработок ЗАО GIPER (ID=172) (см. [112]) и ГИВЦ МК РФ (см. [113]).

Tver Regional Library RXC Server – имеет две инсталляции в России. Прослеживаются родственные связи с сервером ГИВЦ. Декларированные функциональные возможности (Поиск, Просмотр, Просмотр словаря, Сортировка, Удаление наборов, Расширенный сервис) проверить не удается.

Все остальные серверы *Z39.50*, разработанные в России, представлены в информационном пространстве в единственном экземпляре.

Приведенный выше список исчерпывает серверы *Z39.50* российских разработчиков. Кроме них в России эксплуатируются серверы *IndexData/YAZ Generic Frontend Server* (около 20 инсталляций) и *CNIDR* (около 5 инсталляций).

5.12 Выводы

1. На сегодняшний день сервер *ZooPARK* является единственным, по крайней мере, в России, способным обеспечить функционирование РИС по протоколу *Z39.50*.
2. Эксплуатационные характеристики сервера *ZooPARK* определяются гибкостью его архитектуры.
3. Сервер *ZooPARK* удовлетворяет основным требованиям, предъявляемым к ПО для РИС, и может быть использован как универсальный сервер для построения РИС. Последнее подтверждается количеством инсталляций *ZooPARK* в России и количеством РИС, функционирующих в России на его основе.

Глава 6

Установка и настройка сервера *ZooPARK*

6.1 Как установить сервер

6.1.1 Состав дистрибутива

В зависимости от платформы сервер поставляется в виде архивного файла:

Windows 2k/2k3: ZooPARK_Win32_version.rar
UNIX: ZooPARK_unix_version.tgz

Этот архивный файл имеет следующую структуру каталогов (для UNIX относительно каталога, например, /usr/local/ZooPARK:

Таблица 6.1: Состав дистрибутива *ZooPARK*

bin/	ZooPARK yazm.so dprv_remote.so dprv_isis.so dprv_mssql.so dprv_mysql.so dprv_mmark.so dprv_shell.so dprv_test.so	каталог исполняемых файлов исполняемый файл сервера динамическая библиотека сервера <i>Z-REMOTE</i> – динамический провайдер удаленного сервера <i>Z39.50</i> <i>Z-ISIS</i> – динамический провайдер данных <i>CDS/ISIS</i> <i>Z-MSSQL</i> – динамический провайдер данных <i>MS SQL Server</i> <i>Z-MySQL</i> – динамический провайдер данных <i>MySQL Server</i> <i>Z-MMARK</i> – динамический провайдер данных <i>MAPK-SQL</i> для <i>MS SQL Server</i> <i>Z-SHELL</i> – динамический провайдер данных для внешнего консольного приложения (php, perl, tcl, cmd, sh и др.) <i>Z-TEST</i> – тестовый динамический провайдер данных
------	--	---

	test_dprv zebraidx zgw.so z_origin.so zclient	программа эмуляции сервера для тестирования динамических провайдеров данных программа индексации текстовых, <i>SGML</i> и <i>MARC</i> -файлов (<i>Zebra/IndexData</i>). встроенный шлюз <i>Z-GW</i> <i>Z-ORIGIN</i> – динамическая библиотека клиента <i>Z39.50</i> консольный клиент <i>Z39.50</i>
etc/	zserverd	системный каталог скрипт запуска даемона для Unix
tab/	*.abs *.att *.tag *.est *.var *.map *.mar, *.mrk *.chr *.tab	каталог настроечных файлов файлы описания схем данных файлы описания поисковых атрибутов файлы описания наборов меток файлы описания наборов элементов файлы описания вариантов файлы описания взаимосвязи схем описание <i>MARC</i> -файлов описание спецсимволов настроечные файлы провайдера <i>Z-ISIS</i>
data/	IR-Explain-1/ tgils/ lock/ register/ shadow/ doc/ tmp/ irbis/ irbisw/ gash/ tgeo/ tcimi/ trusm/ tusm/ mypubs/ test/ testr/ mspubs/ access.ini providers.ini netgroup.ini group.ini users.ini passwd.ini zserver.cfg	каталог тестовой базы данных каталог базы данных <i>IR-Explain-1</i> каталог тестовой БД <i>tgils</i> для <i>Zebra</i> каталог служебных файлов <i>Zebra</i> каталог служебных файлов <i>Zebra</i> каталог служебных файлов <i>Zebra</i> каталог файлов <i>WEB</i> -сервера каталог служебных файлов <i>Zebra</i> каталог тестовой БД <i>irbis</i> (<i>Z-ISIS</i>) каталог тестовой БД <i>irbisw</i> (<i>Z-ISIS</i>) каталог тестовой БД <i>gash</i> (<i>Z-ISIS</i>) каталог тестовой БД <i>tgeo</i> (<i>Zebra</i>) каталог тестовой БД <i>tcimi</i> (<i>Zebra</i>) каталог тестовой БД <i>trusm</i> (<i>Zebra</i>) каталог тестовой БД <i>tusm</i> (<i>Zebra</i>) файлы настройки БД <i>mypubs</i> (<i>Z-MySQL</i>) каталог тестовой БД <i>test</i> (<i>Z-TEST</i>) каталог тестовой БД <i>testr</i> (<i>Z-TEST</i>) файлы настройки БД <i>mspubs</i> (<i>Z-MSSQL</i>) настройка доступа к БД по IP. настройка провайдеров данных. логические группы хостов по IP логические группы пользователей файл пользователей файл паролей пользователей конфигурационный файл сервера

	stat.cfg zgw.cfg shell.cfg srv_start.sh index.sh	конфигурационный файл сбора статистики конфигурационный файл шлюза <i>Z-GW</i> конфигурационный файл вызова внешних (PHP, PERL, TCL, и т.д.) приложений для сервера <i>WEB</i> скрипт консольного запуска сервера скрипт для индексации данных
include/	*.h	каталог заголовочных файлов заголовочные файлы
src/	dprv_test.c Makefile	каталог тестового провайдера данных исходный код тестового динамического про- вайдера данных мак-файл для сборки провайдера данных

6.1.2 Установка сервера для *Windows NT*

Для установки сервера *ZooPARK* на *Windows NT/2000* необходимо скопировать дистрибутивный архивный файл *ZooPARK_nt.zip* в рабочий каталог, например, *C:*, и разархивировать его. При этом будет создан каталог *C:\ZooPARK* со структурой, описанной выше.

Для запуска сервера в качестве службы *NT* и регистрации его в реестре, необходимо исполнить задачу сервера со следующим ключем (следует обратить внимание, что запуск сервера должен всегда производиться из каталога *data*):

```
C:
cd \ZooPARK\data
ZooPARK -install
```

После этого управление запуском и остановкой сервера может производиться через панель управления службами *NT*. Единственное, что требуется сделать – это отредактировать в системном реестре параметры командной строки запуска сервера:

```
[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\ZooPARK]
"ImagePath"=c:\ZooPARK\bin\ZooPARK.exe -w . . . . .
```

что можно сделать при помощи редактора реестра *rgedit*. Начиная с версии 2.36 редактирование необязательно. Описание параметров командной строки запуска сервера приведено ниже.

Для операционных систем *Windows 2000/XP/2003* настоятельно рекомендуется изменить правило восстановления сервиса *ZooPARK* при сбое. Для этого в свойствах сервиса *ZooPARK* необходимо на закладке *Recovery* установить для параметров *First failure*, *Second failure* и *Subsequent failures* значение *Restart the Service* как показано на Рис. 6.1.

Для деинсталляции сервера необходимо остановить службу *ZooPARK*, затем в каталоге *data* набрать на системной консоли команду:

```
cd \ZooPARK\data
..\bin\ZooPARK -remove
```

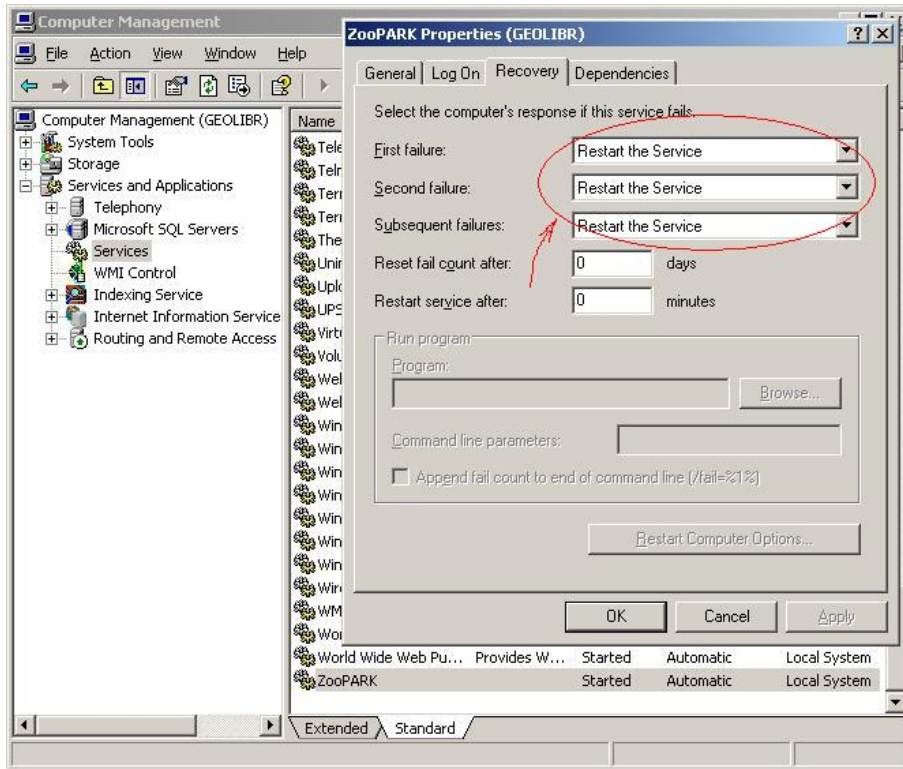


Рис. 6.1: Свойства сервиса ZooPARK при сбое.

При этом будут очищены все относящиеся к серверу записи системного реестра. Содержимое рабочего каталога сервера не изменится.

6.1.3 Установка сервера для UNIX

Для установки сервера на любой версии UNIX необходимо скопировать дистрибутивный архивный файл `ZooPARK_unix_version.tgz` в рабочий каталог, например, в каталог `/usr/local`, и распаковать его:

```
cd /usr/local
gzip -d < ZooPARK\_unix\_version.tgz | tar xfv -
```

В результате будет создан каталог `/usr/local/ZooPARK` с необходимой структурой. Для автоматического запуска сервера в момент загрузки операционной системы необходимо из каталога `ZooPARK/etc` скопировать в соответствующий системный каталог, например, `/etc/init.d`, скрипт, отредактировать его и запустить с параметром `start`.

Для деинсталляции сервера необходимо остановить сервер и удалить скрипт автоматического запуска. Затем можно удалить рабочий каталог сервера.

6.1.4 Настройка конфигурационных файлов

Сервер имеет несколько уровней настройки. Для первого запуска достаточно изменить значения некоторых параметров в файле `data/zserver.cfg`

```
# Пример конфигурационного файла
#
# Где находятся таблицы?

profilePath: ../tab

# Описание сервера

server.name:      ServerName
server.ip:        127.0.0.1
server.tabpath:   ../tab
server.passwd:    ./passwd.ini
server.access:    ./access.ini
server.netgroup:  ./netgroup.ini
server.group:     ./group.ini
server.users:     ./users.ini
server.providers: ./providers.ini
server.threads:   1
server.numoper:   10000

server.docroot:   ../data/doc
server.docfile:   index.htm
server.gwdir:     /zgw/:/zgw/:/zgw/:/zgw/
server.gwcfg:     ../data/zgw.cfg

server.shell:     ../data/shell.cfg
server.shellext:  .php:.asp:.pl:.tcl

# Наборы атрибутов для провайдера Zebra

attset:  bib1.att
attset:  gils.att
attset:  geo.att
attset:  explain.att
attset:  xdl.att
attset:  zthes.att
attset:  util.att
attset:  cimi.att
attset:  collections.att

# Настройка подсистемы Zebra

register: register:500M
lockDir: lock
setTmpDir: tmp
keyTmpDir: tmp
memMax: 4
shadow: shadow:500M
threads: 0

# Правила индексирования Zebra

tgils.database:  tgils
tgils.recordType: grs.sgml
tgils.storeKeys: 1
```

```

tgils.storeData: 1
tgils.recordId: file
tgils.srcCodePage: 0

IR-Explain-1.database: IR-Explain-1
IR-Explain-1.recordType: grs.shtml
IR-Explain-1.storeKeys: 1
IR-Explain-1.storeData: 1

# Описание баз данных

tgils.type: 7
tgils.codepage: 1
tgils.path: ./tgils/tgils.cfg

IR-Explain-1.type: 7
IR-Explain-1.codepage: 1
IR-Explain-1.path: ./IR-Explain-1/IR-Explain-1.cfg

gash.type: 4
gash.index: 1
gash.codepage: 0
gash.path: ./gash/gash.z39

testr.type: 4
testr.index: 10
testr.codepage: 0
testr.path: ./testr/testr.ini

#-----

```

Другие конфигурационные файлы для первого запуска тестовой базы `tgils` можно оставить без изменения.

6.1.5 Подготовка тестовых баз данных

Тестовые базы данных провайдера *Zebra* в виде текстовых *XML*-файлов (`tgils`, `tgeo`, `tcimi`) и в виде файлов формата *ISO-2709* (`tusm`, `trusm`) находятся в каталоге `data`. Для работы необходимо их проиндексировать, выполнив скрипт `index.sh` в соответствующих каталогах или скрипт `index.sh` в каталоге `data`, что приведет к построению в служебных каталогах индексных файлов (`*.mf`). Заметим, что если исполнить скрипт `index.sh` в каталоге `data`, будет проиндексирована база данных `IR-Explain-1`, файлы которой находятся в каталоге `data/IR-Explain-1`.

6.1.6 Запуск сервера

Для запуска сервера в качестве консольного приложения необходимо выполнить следующие команды:

UNIX:

```

# cd /usr/local/ZooPARK/data
# ../bin/ZooPARK tcp:@:210

```

Windows:

```
c:\> cd \ZooPARK\data
c:\> ..\bin\ZooPARK tcp:@:210
```

6.1.7 Проверка работоспособности сервера *Z39.50*

Для проверки работоспособности сервера *ZooPARK* можно воспользоваться консольным клиентом, входящим в состав дистрибутива. Предполагая, что тестовая база данных настроена и сервер стартован, исполним следующие команды (для *UNIX*):

```
# cd /usr/local/ZooPARK/bin - установим рабочий каталог
# ./zclient - выполним программу
z3950> codepage koi - установим кодовую таблицу
z3950> open tcp:localhost:210 - соединимся с сервером
. . . . .
z3950> base tgils - установим базу данных
z3950> find @attr 5=1 geolog - выполним поиск
. . . . .
z3950> format sutrs - установим формат выдачи
z3950> show - посмотрим одну запись
. . . . .
z3950> format grs1 - установим формат выдачи
z3950> show - посмотрим следующую запись
. . . . .
z3950> format usmarc - установим формат выдачи
z3950> show - посмотрим следующую запись
z3950> show - посмотрим следующую запись
. . . . .
z3950> close - завершим сеанс связи
z3950> quit - завершим задачу
#
```

6.1.8 Проверка работоспособности сервера *WEB*

Для проверки работоспособности сервера *ZooPARK* по протоколу *HTTP* необходимо после запуска *ZooPARK* в любом браузере набрать *URL* (см. Рис. 6.2):

```
http://myserver.domain.ru:210
```

6.1.9 Проверка *PHP*-модуля сервера *WEB*

Для проверки работоспособности *PHP*-модуля *WEB* сервера *ZooPARK* необходимо после запуска *ZooPARK* в любом браузере набрать *URL* (см. Рис. 6.4):

```
http://myserver.domain.ru:210/test.php
```

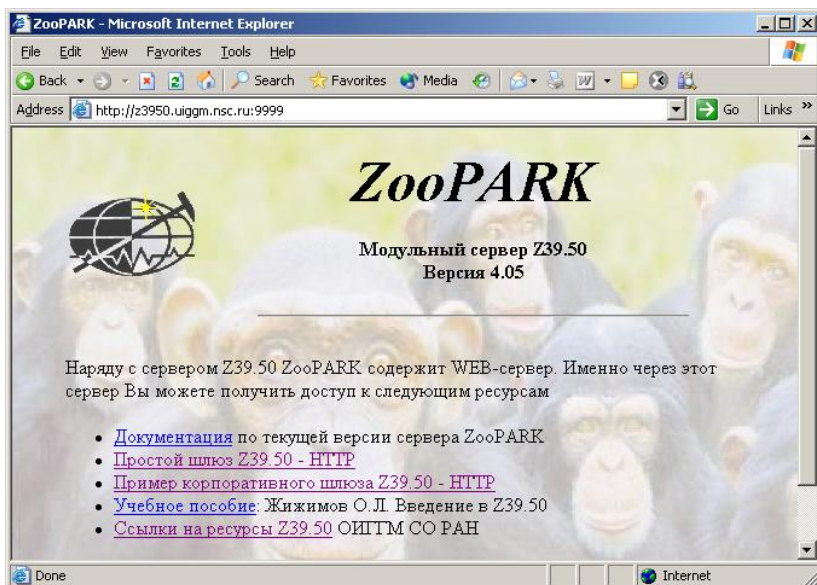


Рис. 6.2: Первая страница WWW сервера

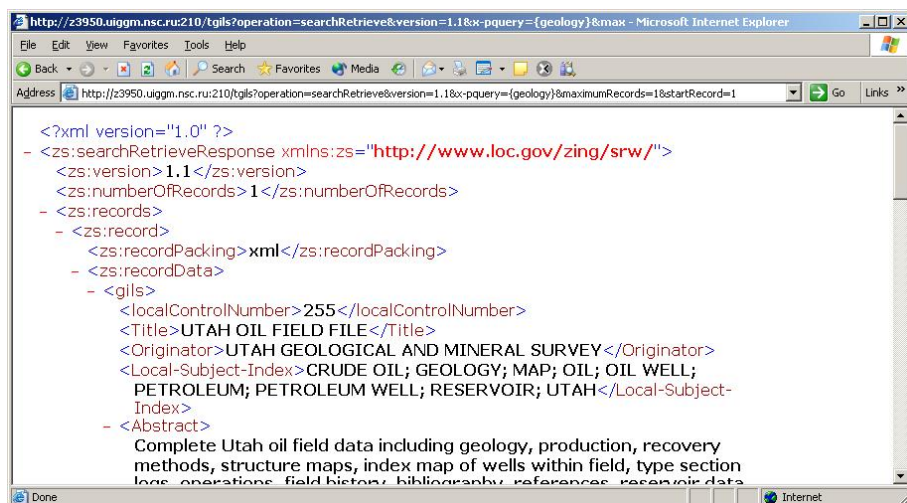


Рис. 6.3: Проверка сервера SRW/SRU

6.1.10 Проверка PERL-модуля сервера WEB

Для проверки работоспособности PERL-модуля WEB сервера ZooPARK необходимо после запуска ZooPARK в любом браузере набрать URL (см. Рис. 6.5):



Рис. 6.4: Проверка модуля поддержки PHP

<http://myserver.domain.ru:210/test.pl>

6.1.11 Проверка сервера *SRW/SRU*

Для проверки работоспособности сервера *ZooPARK* по протоколу *SRU* над *HTTP* необходимо после запуска *ZooPARK* в любом браузере набрать *URL*:

```

http://myserver.domain.ru:210/tgils?operation=searchRetrieve&
version=1.1&x-pquery={geology}&maximumRecords=1&startRecord=1

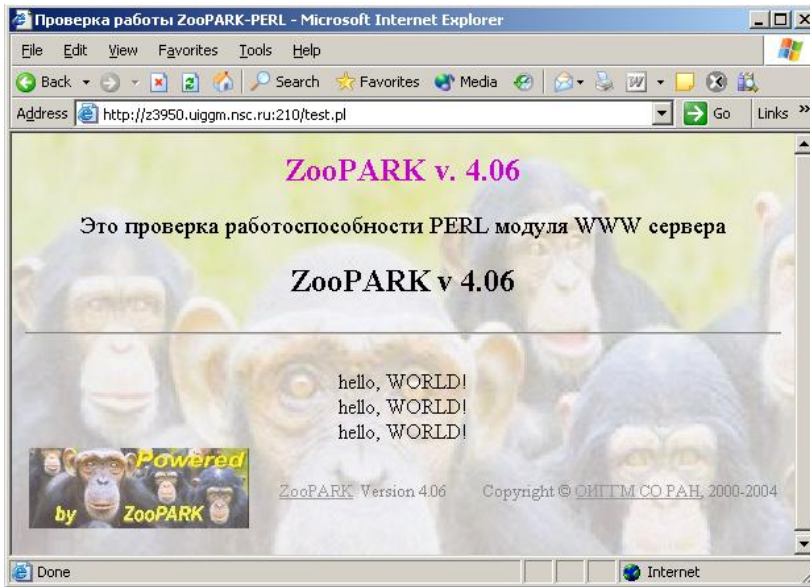
```

Мы должны увидеть две записи в схеме *GILS* в формате *XML* (см. Рис. 6.3).

Проверка работоспособности сервера по *SOAP/SRW* требует применения консольного клиента *Z39.50* от *IndexData* или входящего в комплект *ZooPARK*.

6.1.12 Проверка шлюза *Z39.50-HTTP*

Для проверки работоспособности шлюза *Z-GW* необходимо после запуска *ZooPARK* в любом браузере набрать *URL*:

Рис. 6.5: Проверка модуля поддержки *PERL*

`http://myserver.domain.ru:210/zgw/index.htm`

Если после этого видна начальная страница простого шлюза (Рис.6.6), можно нажать на кнопку со значком "?" против базы `tgils`. При правильной работе шлюза должна быть получена и отображена запись *Explain* в категории *databaseInfo* для БД `tgils` (см. Рис. 6.6).

6.2 Конфигурирование и настройка

6.2.1 Конфигурационный файл сервера

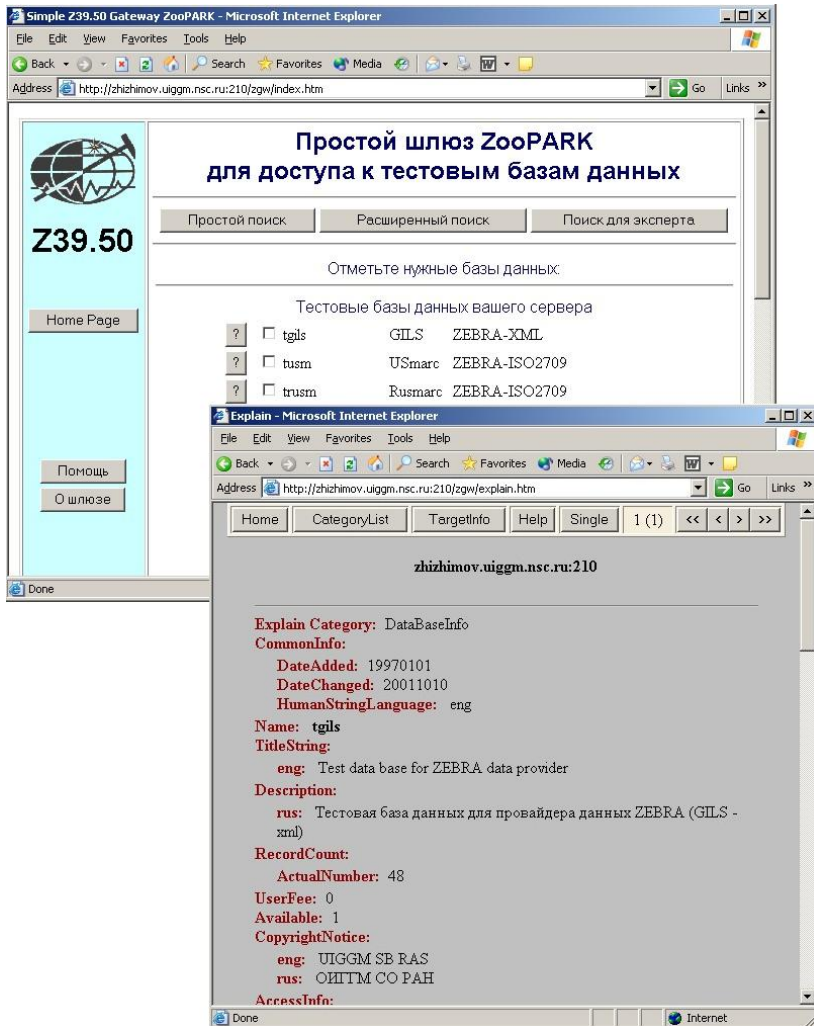
Конфигурационный файл сервера `zserver.cfg` состоит из логических блоков: блок общей настройки, блок настройки параметров сервера, блок настройки провайдера данных *Zebra* и блок настройки баз данных.

Блок общей настройки требует определения параметра *profilePath* – абсолютный или относительный от каталога `data` путь к каталогу таблиц `tab/`. Например:

```
profilePath: ../tab
```

Блок настройки параметров сервера требует определения параметров

```
server.name      - имя сервера
server.ip        - IP-адрес сервера
server.tabpath   - путь к файлам провайдеров данных
server.passwd    - имя файла пользователей и паролей
```


Рис. 6.6: Шлюз Z-GW и Explain-запись *databaseInfo* для tgils

```

server.access      - имя файла прав доступа
server.netgroup   - имя файла логических групп по IP
server.users      - имя файла пользователей
server.group      - имя файла логических групп пользователей
server.providers  - имя файла настройки провайдеров
server.threads    - разрешение (запрещение)
                   распараллеливания поиска в отдельных
                   потоках для сервера
server.numoper    - максимальное количество операций в
                   сеансе
server.docroot:   - путь к каталогу данных
server.docfile:   - имя начального html документа по
                   умолчанию

```

```
server.gwdir:    - список каталогов шлюза в каталоге WEB
                  сервера
server.gwcfg:   - путь и имя файла конфигурации шлюза
```

Например:

```
server.name:     geolibr
server.ip:       194.226.185.16
server.tabpath:  ../tab
server.passwd:   ./passwd.ini
server.access:   ./access.ini
server.netgroup: ./netgroup.ini
server.group:    ./group.ini
server.users:    ./users.ini
server.providers: ./providers.ini
server.threads:  1
server.numoper:  10000
server.docroot:  ../data/doc
server.docfile:  index.htm
server.gwdir:    /zgw/:/zgw/:/zgw/:/zgw/
server.gwcfg:    ../data/zgw.cfg

server.shell:    ../data/shell.cfg
server.shellext: .php:.asp:.pl:.tcl
```

Блок настройки баз данных требует для каждой базы данных указания четырех параметров, префиксированных через точку именем базы данных:

type – Номер статического провайдера данных:

Loadable=4, Group=5, Zebra=7

codepage – Кодовая таблица для этой базы данных:

DOS=0, WIN=1, KOI=2, ISO=3, MAC=4

path – Дополнительные настройки (зависит от провайдера)

index – Номер динамического провайдера (см. файл `provider.ini`)

Ниже приведен пример блока настройки баз данных для IR-Explain-1 (*Zebra*), *gash* (*Z-ISIS*), *tremote* (*Z-Remote*) и *tgroup* (*Group*):

```
# Пример настройки баз данных

IR-Explain-1.type: 7
IR-Explain-1.codepage: 1
IR-Explain-1.path:  ./IR-Explain-1/IR-Explain-1.cfg

tremote.type: 4
tremote.codepage: 0
tremote.path:  z3950.uiggm.nsc.ru:210:geocat
tremote.index: 2

gash.type: 4
gash.codepage: 0
gash.path:  ./gash/gash.z39
gash.index: 1
```

```
tgroup.type: 5
tgroup.codepage: 0
tgroup.path: gash;tremote;

# Конец примера
```

Блок настройки провайдера данных *Zebra* необходим для индексирования данных. Большая часть параметров этого блока описана в документации по *Zebra/IndexData* [114].

6.2.2 Ограничение прав доступа

Правила ограничения доступа к базам данных по *IP*-адресам сформулированы в файле `access.ini`. Этот файл перечитывается сервером в момент инициализации сеанса, поэтому изменения, внесенные в него, на уже открытые сеансы не распространяются.

Внимание! Каждая база данных и каждое групповое имя должны иметь точку входа в файле `access.ini`.

На примере пояснен принцип управления доступом:

```
#----- file access.ini

# Создадим точку входа для БД tgils
[tgils]
# разрешим доступ всем
all=allow

# Создадим точку входа для БД db1
[db1]
# запретим доступ всем
all=denied
# разрешим доступ для групп Group_1, Group_99
Group_1=allow
Group_99=allow

#----- file netgroup.ini

194.226.184: Group_1
194.202.15: Group_99
```

6.2.3 Аутентификация

Сервер допускает работу не только с анонимными клиентами. В момент инициализации сеанса связи возможно указание имени пользователя и пароля в соответствии с протоколом. Сервер проверяет полученные данные по таблице паролей, имеющей простой формат:

```
имя1:пароль1
имя2:пароль2
```

В настоящей версии сервера пароли хранятся в незашифрованном виде. Имя файла паролей указано в конфигурационном файле сервера. Заметим,

что функции расширенного сервиса, как правило, не допускают исполнения анонимными клиентами.

Начиная с версии 4.07 сервер поддерживает группы пользователей и авторизацию через *LDAP*. Группы пользователей описываются в файле `group.ini`. Пример описания групп:

```
[nsc]
type=local
users=nsc_users

[nsc_users]
zhizhimov
mazov
fedotov
skibin

[arbicon]
type=ldap
dll=rldap.dll
servers=arbicon_srv

[arbicon_srv]
l=ldap://www.ruslan.ru/dc=ruslan,dc=ru?uid?sub
```

Здесь описаны две группы. Первая группа `[nsc]` авторизуется локально и содержит список пользователей, перечисленных в секции `[nsc_users]`. Вторая группа `[arbicon]` авторизуется на сервере *LDAP*, указанном в секции `[arbicon_srv]`. При этом используется динамическая библиотека `rldap.dll`.

Указание ограничений доступа к базам данных по пользователям и группам проиллюстрировано на примере:

```
# ----- file access.ini

[tgils]
all=denied
g:arbicon=allow
u:zhizhimov=allow
```

в котором к базе данных `tgils` разрешается доступ только членам группы `arbicon` и пользователю `zhizhimov`.

Следует заметить, что при открытии сеанса имя группы может быть указано в имени пользователя через символ `/`. Например, имя `arbicon/zhizhimov` будет интерпретироваться как пользователь `zhizhimov` в группе `arbicon`.

Наконец, существует возможность авторизации без явного указания группы пользователя. Для этого в файле `users.ini` в секции `default` должны быть в качестве параметров перечислены пользователи. При этом для каждого пользователя значением должен быть список групп, разделенный запятой. Проверка происходит в каждой группе последовательно и при первом успехе завершается. Пример файла `users.ini` приведен ниже

```
# ----- file users.ini
```

```
[default]
zhizhimov=uiggm,nsc,arbicon,local
test=arbicon,local
mazov=uiggm,nsc,arbicon,local
fedotov=nsc,local
```

6.2.4 Настройка провайдеров данных

Конфигурационный файл провайдеров

Конфигурационный файл динамических провайдеров данных (по умолчанию он называется `providers.ini`) имеет структуру

```
[names]
name1=index1
name2=index2
. . .

[name1]
file=dll_file_name
thread=1
. . .
```

где в секции `[names]` перечисляются все динамические провайдеры данных с указанием уникального номера и имени секции, в которой указывается имя файла динамической библиотеки. Уникальный номер провайдера используется в конфигурационном файле сервера при описании параметров баз данных. Параметр `thread` указывает на разрешение (1) или запрещение (0) доступа к провайдеру данных в отдельном потоке для каждой базы данных.

Провайдер данных *Zebra*

Провайдер данных *Zebra* не требует специальных настроечных файлов. Вся настройка сводится к указанию соответствующего номера провайдера в блоке настройки баз данных файла конфигурации сервера (см. раздел 6.2.1) и заполнению в том же файле блока настройки провайдера данных *Zebra*.

Существует несколько отличий провайдера данных *Zebra* от ее исходной реализации *IndexData* [114].

1. При работе с файлами форматов *ISO-2709* в конфигурационном файле `zserver.cfg` необходимо указывать кодовую страницу исходных файлов, например

```
trusm.database:      trusm
trusm.recordType:   grs.marc.rusmarc
trusm.storeKeys:    1
trusm.storeData:    1
trusm.srcCodePage:  0
. . .

trusm.type:         7
trusm.codepage:     1
trusm.path:         ./trusm/trusm.cfg
```

при этом в общем описании БД кодовая страница должна совпадать с внутренней кодовой страницей сервера (`trusm.codepage: 1`).

2. Необходимо указание дополнительного настроечного файла для каждой базы данных, например, файла `trusm.cfg` для БД `trusm`:

```
[base]
name=trusm
schema=rusmarc

[update]
folder=./trusm/data
key=<controlIdentifier>
ignore=<local-control-number>
group=trusm
index=./trusm/index.cmd
```

содержащего название базы данных и ссылку на программу индексации. Следует заметить, что вторая секция в этом файле используется только функциями расширенного сервиса.

3. При обработке *XML*-файлов следует обратить внимание на возможность работы с файлами, отображающими структуру файлов *ISO-2709 (MARC)*. При этом существует возможность описать не только общую структуру файлов, но и содержимое маркера. В качестве примера подобных записей можно обратиться к файлу `testr_data.sgml` в каталоге `data/testr`. Например, *XML*-запись для версии 3.00 и выше

```
<rusmarc>
  <mrk len="24">
    <m_0_4 len="5">#####</m_0_4>
    <m_5_5 len="1">n</m_5_5>
    <m_6_6 len="1">a</m_6_6>
    <m_7_7 len="1">m</m_7_7>
    <m_8_8 len="1">#</m_8_8>
    <m_9_9 len="1">#</m_9_9>
    <m_10_10 len="1">2</m_10_10>
    <m_11_11 len="1">2</m_11_11>
    <m_12_16 len="5">#####</m_12_16>
    <m_17_17 len="1">#</m_17_17>
    <m_18_18 len="1">#</m_18_18>
    <m_19_19 len="1">#</m_19_19>
    <m_20_23 len="4">450#</m_20_23>
  </mrk>
  <001>RU\Novosibirsk\NSU\Library\38285</001>
  <010><##><a>588782218X</a></##></010>
  <010><##><a>0201634554</a></##></010>
  <100>
    <##><a len="36">
      <p100_0_7 len="8">20000429</p100_0_7>
      <p100_8_8 len="1">#</p100_8_8>
      <p100_9_12 len="4">2000</p100_9_12>
      <p100_13_16 len="4">#####</p100_13_16>
      <p100_17_19 len="3">|||</p100_17_19>
```

```

        <p100_20_20 len="1">|</p100_20_20>
        <p100_21_21 len="1">|</p100_21_21>
        <p100_22_24 len="3">rus</p100_22_24>
        <p100_25_25 len="1">у</p100_25_25>
        <p100_26_29 len="1">0102</p100_26_29>
        <p100_30_33 len="4">####</p100_30_33>
        <p100_34_35 len="2">ca</p100_34_35>
    </a></##>
</100>
<101><##><a len="3">rus</a></##></101>
<200>
    <1#>
        <a>Язык программирования Java</a>
        <f>[Пер.с англ.Е.Матвеев]</f>
    </1#>
</200>
<210><##><a>СПб.и др.</a>
        <c>Питер</c><d>1997</d></##></210>
<215><##><a>304 с.</a></##></215>
<300><##><a>Библиогр.:с.291-294.</a></##></300>
<686><1#><a>В185.29(Java)</a><2>rubbk</2></1#></686>
<801><#0>
    <a>ru</a>
    <b>Novosibirsk State University Scientific Library</b>
    <c>20000429</c></#0></801>
</rusmarc>

```

будет однозначно отображена в формат *RUSmarc* с генерацией соответствующего маркера и полей фиксированной длины. Знак # будет заменен пробелом. Следует обратить внимание на способ задания индикаторов, подполей и символьных позиций в полях фиксированной длины. Также следует заметить, что такие файлы следует индексировать по типу *grs.sgml*.

- Начиная с версии 3.00 для корректной обработки и отображения полей фиксированной длины в форматах *ISO-2709*, введены *XML*-тэги, характеризующие символьные позиции. Эти тэги должны быть описаны в файлах *.mrk, которые заменили файлы *.mar предыдущих версий сервера *ZooPARK* и оригинального варианта *Zebra*.

```

# File rusmarc.mrk
name rusmarc
reference rusmarc

fld mrk - 24
pos m_0_4 0 5 00000
pos m_5_5 5 1 n
pos m_6_6 6 1 a
pos m_7_7 7 1 m
pos m_8_8 8 1 @
pos m_9_9 9 1 @
pos m_10_10 10 1 2
pos m_11_11 11 1 2
pos m_12_16 12 5 00000
pos m_17_17 17 1 @

```

```

pos m_18_18 18 1 @
pos m_19_19 19 1 @
pos m_20_23 20 4 450@

fld 100 a 36
pos p100_0_7 0 8 ||| ||| |||
pos p100_8_8 8 1 d
pos p100_9_12 9 4 ||| |
pos p100_13_16 13 4 ||| |
pos p100_17_19 17 3 k||
pos p100_20_20 20 1 y
pos p100_21_21 21 1 0
pos p100_22_24 22 3 rus
pos p100_25_25 25 1 y
pos p100_26_29 26 4 0102
pos p100_30_33 30 4 ||| |
pos p100_34_35 34 2 ca

```

Здесь описана внутренняя структура маркера записи и некоторых полей на примере поля 100 подполя а. Для поля (fld) указывается: имя тэга, символ подполя или '-', если подполя нет, длина данных подполя. Для данных (pos) указывается: имя тэга, смещение символьной группы, длина символьной группы, значение по умолчанию. Следует заметить, что в этом файле должны быть описаны все поля и подполя фиксированной длины с позиционно значащими символами.

Провайдер данных удаленного сервера *Z39.50*

Провайдер данных удаленного сервера *Z39.50 (Z-REMOTE)* не требует специальных настроечных файлов. Вся настройка сводится к указанию соответствующего номера провайдера в блоке настройки баз данных файла конфигурации сервера (см. раздел 6.2.1) и его индекса в соответствии с конфигурационным файлом провайдеров (см. раздел 6.2.4). Для динамического варианта номер провайдера данных равен 4. Параметр *path* имеет значение

<DNS-имя сервера>:<port>:<имя базы данных>

Пример правильной настройки:

```

tremote.type:      4
tremote.codepage:  0
tremote.path:      z3950.uiggm.nsc.ru:210:geocat
tremote.index:     2

```

Для более тонкой настройки провайдера *Z-REMOTE* используется конфигурационный файл *provider.ini*. Пример настройки

```

# File providers.ini (UNIX)
. . .
[remote]
file=./dprv_remote.so
thread=1
load=1
loop=2
max=3

```



```
list=remote_list

[remote_list]
1=mx.uiggm.nsc.ru:210
2=z3950.uiggm.nsc.ru:210
3=camelot.spsl.nsc.ru:2100
```

здесь уникальные для этого провайдера данных настроечные параметры предписывают: *loop* – максимально допустимое количество петель при переадресации запросов, *max* – максимально допустимое количество переадресаций, *load* – разрешение (1) / запрещение (0) открытия вторичных сеансов при инициализации, *list* – имя секции со списком серверов, для которых устанавливаются вторичные сеансы при инициализации первичного сеанса.

Общие правила настройки провайдеров данных для СУБД

Провайдеры данных для конкретных СУБД для настройки требует специального настроечного файла для каждой базы данных. Имя этого файла указывается в параметре *path* блока настройки баз данных файла конфигурации сервера. Пример правильной настройки:

```
gash.type:      4
gash.codepage:  0
gash.path:      ./gash/gash.z39
gash.index:     1
```

При этом параметр *index* для динамических провайдеров должен соответствовать номеру в конфигурационном файле провайдеров данных.

Настроечный файл (в примере *gash.z39*) имеет структуру

```
[section1]
parametr1=value1
parametr1=value2

[section2]
. . .
```

Название секций и параметров настроечных файлов для различных провайдеров различаются. Однако существуют общие для всех провайдеров данных секции. В следующих таблицах приведены имена секций и параметров настроечных файлов, которые являются общими для всех провайдеров.

Наборы поисковых атрибутов и их допустимые значения описываются в секциях:

AttSet – в этой секции в качестве параметров перечисляются допустимые наборы поисковых атрибутов. В качестве названия параметра выступает текстовое представление соответствующего *OID*, а в качестве значения – название секции с описанием набора.

Параметр	Значение	Пример
Bib-1 (имя набора)	Имя секции, описывающей набор атрибутов.	Attr_Bib1 ¹

Attr_Bib1¹ — описание набора поисковых атрибутов Bib-1:

Параметр	Значение	Пример
0	Имя секции, описывающей значения атрибутов по умолчанию	Bib1_def ¹
1	Имя секции, описывающей значения атрибутов Use	Bib1_1 ¹
2	Имя секции, описывающей значения атрибутов Relation	Bib1_2 ¹
3	Имя секции, описывающей значения атрибутов Position	Bib1_3 ¹
4	Имя секции, описывающей значения атрибутов Structure	Bib1_4 ¹
5	Имя секции, описывающей значения атрибутов Truncation	Bib1_5 ¹
6	Имя секции, описывающей значения атрибутов Completeness	Bib1_6 ¹
d	Имя секции, описывающей тип данных	Bib1_d ¹

Bib1_def¹ — в данной секции перечисляются значения атрибутов по умолчанию, например

```
[Bib1_def]
1=1035
2=3
3=3
4=108
5=100
6=3
```

Bib1_1¹ — в данной секции каждому значению поискового атрибута USE ставится в соответствие

- префикс поля (для ISIS) или символ ”~”. Если после префикса указан текст в фигурных скобках, то при поиске по этому атрибуту этот текст будет удален из поискового термина.
- перечисленные в квадратных скобках *USE* атрибуты, которые являются комбинацией для текущего.

¹Название секции зависит от значений параметров других секций

¹Название секции зависит от значений параметров других секций

Например,

```
[Bib1_1]
4=~
8=.is.
12=.no.{aa\bb}
31=.yr.
1035=[4][8][31]
```

Bib1_2, Bib1_3, Bib1_4, Bib1_5, Bib1_6 — в секциях¹ перечисляются допустимые значения соответствующих атрибутов или символ ~

Bib1_d¹ — `<typ>[<list>]` — В данной секции каждому значению поискового атрибута Use ставится в соответствие тип данных — с (текст), d (data), y(год), i (целое число). По умолчанию все поля имеют тип с. Необязательный параметр list содержит список допустимых атрибутов в виде `typ:value;[typ:value;]`, например

```
[Bib1_d]
4=c
31=y 4:4;2:1;
1011=d
```

Для описания допустимых схем данных, форматов внешнего представления и способа формирования структуры экспортируемой записи:

Schema — в этой секции в качестве параметров перечисляются допустимые схемы данных или форматы. В качестве названия параметра выступает текстовое представление соответствующего *OID* (см. Приложение), а в качестве значения — название секции с описанием схемы.

Параметр	Значение	Пример
default	Имя секции, описывающей выбор схемы по умолчанию	rusmarc
RUSmarc	Имя секции, описывающей формат <i>RUSMARC</i>	rusmarc
USmarc	Имя секции, описывающей формат <i>USmarc</i>	usmarc
UIGGM-schema	Имя секции, описывающей схему <i>UIGGM</i>	uiggm

RecordSyntax — в этой секции в качестве параметров перечисляются допустимые форматы. В качестве названия параметра выступает текстовое представление соответствующего *OID*, а в качестве значения — название секции с описанием способа формирования записи.

¹Название секции зависит от значений параметров других секций

Параметр	Значение	Пример
default	Имя секции, описывающей выбор формата по умолчанию	rusmarc
RUSmarc	Имя секции, описывающей формат RUSMARC	rusmarc
USmarc	Имя секции, описывающей формат USmarc	usmarc
SUTRS	Имя секции, описывающей формат SUTRS	sutrs
GRS-1	Имя секции, описывающей формат GRS-1	~
rtf	Имя секции, описывающей формат rtf	rtf
html	Имя секции, описывающей формат html	html

В секциях, описывающих конкретную схему или формат, указывается способ формирования записей. При этом допускается четыре способа (параметр `Type`):

Type=shablon – запись формируется как целое в соответствии с указанным файлом шаблона

Type=table – запись формируется поэлементно в соответствии с указанной таблицей

Type=procedure – запись формируется как целое не провайдером данных, а СУБД. Как правило, указывается имя хранимой процедуры.

Type=inner – запись формируется автоматически сервером *ZooPARK* на основе таблиц конвертирования схем.

Название файла шаблона должно быть указано в параметре *Shablon* или параметрах, имена которых совпадают с именами затребованных наборов элементов, например, `B=shablon_b.txt`. В таблице приведен пример:

rusmarc – секция, описывающая формат RUSmarc

Параметр	Значение	Пример
Type	Тип формирования записи	Type=table
Shablon	Имя файла шаблона	Shablon= ./gash/gash_rmarc.tab
B	Имя файла шаблона для набора элементов "B"	B= ./gash/gash_rmarc_b.tab

uiggm – секция, описывающая схему uiggm

Параметр	Значение	Пример
Type	Тип формирования записи	Type=shablon
Shablon	Имя файла шаблона	Shablon= ./gash/gash_uiggm.txt
B	Имя файла шаблона для набора элементов "B"	B= ./gash/gash_uiggm_b.txt

При описании форматов *RTF*, *HTML* может быть указаны параметры, уточняющие детали автоматической конвертации записей в соответствующий формат:

rtf – секция, описывающая формат *RTF*

Параметр	Значение
rec_beg	Текст, вставляемый перед началом записи, или символ ~
rec_end	Текст, вставляемый после записи, или символ ~
tag_beg	Текст, вставляемый перед названием тэга, или символ ~
tag_end	Текст, вставляемый после названия тэга, или символ ~
dat_beg	Текст, вставляемый перед выводом поля, или символ ~
dat_end	Текст, вставляемый после вывода поля, или символ ~
lev_beg	Текст, вставляемый при переходе на более высокий уровень вложения, или символ ~
lev_end	Текст, вставляемый при переходе на более низкий уровень вложения, или символ ~

html – секция, описывающая формат *HTML*. Параметры соответствуют параметрам секции [rtf].

Провайдер данных *Z-ISIS*

Провайдер данных *CDS/ISIS (Z-ISIS)* требует специального настроечного файла для каждой базы данных. Имя этого файла указывается в параметре *path* блока настройки баз данных файла конфигурации сервера (см. раздел 6.2.4). Пример правильной настройки:

```
gash.type:      4
gash.codepage: 0
gash.path:     ./gash/gash.z39
gash.index:    1
```

Настроечный файл (в примере *gash.z39*) имеет структуру

```
[section1]
parametr1=value1
parametr1=value2
```

[section2]

...

В таблице ниже приведены имена секций и параметров настроечного файла. При этом опущены все секции и параметры, значения которых соответствуют общим правилам (см. выше):

DataBase – секция, описывающая базу данных

Параметр	Значение	Пример
name	Имя базы данных <i>CDS/ISIS</i>	cat
path	Полный путь к файлам БД	/usr/cat/
description	Описание БД (необязательно)	~

Для вывода записей в некоторых форматах *ISO-2709* могут понадобиться файлы отображения меток полей *ISIS* в структуры внешнего представления. Блок формирования полей представляет собой список инструкций для заполнения каждого элемента структуры внешнего представления записи:

```
tag <element_path> <element_name> <typ> <from> <num> <fmt>
```

Здесь `tag` – признак блока, `<element_path>` – положение элемента в иерархической схеме, `<element_name>` – уникальное имя элемента, `<typ>` – тип формирования поля, причем обрабатываются по формату `<fmt> <num>` повторений полей *ISIS*, начиная с повторения `<from>`. Значения `<typ>` могут быть: 0 – форматная обработка, повторяющиеся подполя, 1 – форматная обработка, повторяющиеся поля, 2 – без формата с повторением на текущем уровне, 3 – без формата ветвление, 4 – без формата с повторением на корневом уровне. В качестве примера приведен фрагмент мапирования базы данных *CDS/ISIS* на *RUSmarc* (версия 3.00 и выше):

```
# Marker
tag mrk mrk 3 0 0 -
tag mrk/m_0_4 m_0_4 0 0 0 '?????'
tag mrk/m_5_5 m_5_5 0 0 0 'n'
tag mrk/m_6_6 m_6_6 0 0 0 'a'
tag mrk/m_7_7 m_7_7 0 0 0 'm'
tag mrk/m_8_8 m_8_8 0 0 0 ' '
tag mrk/m_9_9 m_9_9 0 0 0 ' '
tag mrk/m_10_10 m_10_10 0 0 0 '2'
tag mrk/m_11_11 m_11_11 0 0 0 '2'
tag mrk/m_12_16 m_12_16 0 0 0 '?????'
tag mrk/m_17_17 m_17_17 0 0 0 '1'
tag mrk/m_18_18 m_18_18 0 0 0 'i'
tag mrk/m_19_19 m_19_19 0 0 0 ' '
tag mrk/m_20_23 m_20_23 0 0 0 '450@'

# ISBN
tag 010 010 3 0 0 -
tag 010/a a 0 0 0 v010

# Local
tag 091 091 3 0 0 -
```

```

tag 091/a a 0 0 0 v501

# Publication
tag 210 210 3 0 0 -
tag 210/a a 0 0 0 v210
tag 210/b b 0 0 0 v211
tag 210/c c 0 0 0 v207

# Subject
tag 606 606 3 0 0 -
tag 606/@0 @0 3 0 0 -
tag 606/@0/a a 1 1 99 (v640/)

# ББК
tag 686 686 3 0 0 -
tag 686/a a 0 0 0 v400

# Author
tag 700 700 3 0 0 -
tag 700/@1 @1 3 0 0 -
tag 700/@1/a a 0 1 1 v705[1]^a
tag 700/@1/b b 0 1 1 v705[1]^b

# Other authors
tag 701 701 3 0 0 -
tag 701/@1 @1 3 0 0 -
tag 701/@1/a a 1 2 99 (v705[1..]^a,| b|v705[1..]^b/)

```

Следует обратить внимание на способ формирования ветвления и индикаторов.

Внимание! Кодовая таблица файла мапирования должна совпадать с кодовой таблицей базы данных CDS/ISIS (IRBIS).

Провайдер данных Z-MSSQL

Провайдер данных *Z-MSSQL* для *MS SQL Server* (`dprv_mssql`) требует специального настроечного файла для каждой базы данных. Имя этого файла указывается в параметре `path` блока настройки баз данных файла конфигурации сервера. Пример правильной настройки:

```

mspubs.type: 4
mspubs.codepage: 0
mspubs.path: ./mspubs/mspubs.z39
mspubs.index: 3

```

Настроечный файл (в примере `mspubs.z39`) имеет структуру

```

[section1]
parametr1=value1
parametr1=value2

[section2]
. . .

```

В следующей таблице приведены имена секций и параметров настроечного файла, отличающиеся от общих настроек:

DataBase – секция, описывающая базу данных:

Параметр	Значение
Name	Имя базы данных
Server	Имя сервера MS SQL
IP_address	IP-адрес SQL сервера
Port	TCP-порт SQL сервера
TDS_version	Версия протокола (только для UNIX)
Login	Имя входа на SQL сервер
Passwd	Пароль
SqlName	Имя БД для SQL сервера
AnsiToOem	Признак конвертирования данных на уровне сетевой библиотеки (только для Windows NT)
QueryCount	Запрос полного количества записей в БД
QuerySearch	Запрос для поиска данных
QueryPresent	Запрос для представления данных

Ниже приведен пример этой секции

```
[DataBase]
Server=geolibr
IP_address=194.226.185.16
Port=1433
TDS_Version=ver4.2
Login=www
Passwd=www
Name=mspubs
SqlName=pubs
QuerySearch=select COUNT(*) from pubs.dbo.titles
             where 1=1 and
QueryPresent=select title_id, pub_id from pubs.dbo.titles
             where 1=1 and
QueryCount=select COUNT(*) from pubs.dbo.titles
AnsiToOem=OFF
```

Bib1_1 – в данной секции каждому значению поискового атрибута *USE* ставится в соответствие название поля в таблице, например

```
[Bib1_1]
4=title
12=title_id
```

Bib1_d – <val> [<list>] – В данной секции каждому значению поискового атрибута *Use* ставится в соответствие тип данных – с (текст), d (data), y(год), i (целое число). По умолчанию все поля имеют тип с. Необязательный параметр *list* содержит список допустимых атрибутов в виде

typ:value; [typ:value;], например

```
4=c
12=i
31=y 2:1;2:2;2:3;4:4;5:100;
```


Следует обратить внимание на способ формирования *SQL*-запроса провайдером данных:

- Получив *RPN*-запрос, например,

```
@attr 1=4 @attr 5=3 {unix}
```

провайдер в соответствии с параметрами настроечного файла сформирует строку

```
title like '%unix%'
```

- Сформированная таким образом строка будет добавлена к значению параметра *QuerySearch* секции *[DataBase]*, который и будет передан *SQL* серверу.
- При запросе на представление данных совершенно аналогично используется параметр *QueryPresent*.
- Параметр *QueryCount* используется сервисом *Explain* для определения количества записей в базе данных.

Провайдер допускает два варианта настройки для представления данных.

Вариант 1: При запросе на представление данных провайдер, получив данные от *SQL* сервера, производит дополнительную обработку в соответствии с шаблоном, имя которого указано в настроечном файле. Ниже приведен пример такого шаблона для отображения данных из базы *pubs* в формат *RUSmarc* (см. файл `./data/mspubs/mspubs_rusmarc_f.txt`):

```
#
<rusmarc>
  <mrk len="24">
    <m_0_4 len="5">?????</m_0_4>
    <m_5_5 len="1">n</m_5_5>
    <m_6_6 len="1">a</m_6_6>
    <m_7_7 len="1">m</m_7_7>
    <m_8_8 len="1">#</m_8_8>
    <m_9_9 len="1">#</m_9_9>
    <m_10_10 len="1">2</m_10_10>
    <m_11_11 len="1">2</m_11_11>
    <m_12_16 len="5">?????</m_12_16>
    <m_17_17 len="1">#</m_17_17>
    <m_18_18 len="1">#</m_18_18>
    <m_19_19 len="1">#</m_19_19>
    <m_20_23 len="4">450#</m_20_23>
  </mrk>
  <001>Microsoft\pubs\$r.title_id$</001>
  <010><##><d>$r.pric$</d></##></010>
  <101><0#><a>eng</a></0#></101>
  <200><1#><a>$r.title$</a></1#></200>

<$ sql pu select a.* from pubs.dbo.publishers a
      where a.pub_id='$r.pub_id$' $>
<$ for pu $>
```

```

<102><##><a>$pu.country$</a></##></102>
<210><##><a>$pu.city$</a>
    <c>$pu.pub_name$</c>
    <d>$r.pubdate$</d>
</##></210>
<$ close pu $>

<330><##><a>$r.notes$</a></##></330>

<$ sql au select aa.* from pubs.dbo.titleauthor ta,
    pubs.dbo.authors aa
    where ta.au_id=aa.au_id and
        ta.title_id='$r.title_id$' and ta.au_ord=1 $>
<$ for au $>
    <700><##>
        <a>$au.au_fname$ $au.au_lname$</a>
        <p>$au.address$ $au.city$ $au.state$ $au.zip$</p>
    </##></700>
<$ close au $>

<$ sql au select aa.* from pubs.dbo.titleauthor ta,
    pubs.dbo.authors aa
    where ta.au_id=aa.au_id and
        ta.title_id='$r.title_id$' and ta.au_ord>1 $>
<$ for au $>
    <701><##>
        <a>$au.au_fname$ $au.au_lname$</a>
        <p>$au.address$ $au.city$ $au.state$ $au.zip$</p>
    </##></701>
<$ close au $>

<801><#0>
    <a>US</a>
    <b>Microsoft</b>
    <2>rusmarc</2>
</#0></801>
</rusmarc>

```

Файл шаблона представляет собой файл типа *XML* для формирования иерархической структуры для одной записи. Кроме элементов, типа

```
<700><##><a>. . </a></##></700>
```

которые описывают теги *ISO-2709* в цепочке поле-индикаторы-подполе, он содержит команды для встроенного интерпретатора и значения переменных. Команды интерпретатора имеют вид:

```

КомандаИнтерпретатора ::= '<$ ' Команда ' $>'
Команда ::= Тип Имя Тело | Тип Имя
Тип ::= 'sql' | 'for' | 'close'
Имя ::= text
Тело ::= [text] Переменная [text]
Переменная ::= '$Имя'. 'text'$' | '$r.'text'$'

```

Как видно, в процессе обработки шаблона провайдер данных может выполнять дополнительные запросы к *SQL* серверу, порождая каждый раз именованный нескроллируемый курсор. Имя первого курсора предопределено (*r*), он порождается запросом, сформированным из настроечного

файла (см. выше). Дополнительные курсоры должны быть закрыты после обработки. Прокрутка курсора допускается только вперед, при этом обрабатывается следующий после команды *for* полный элемент шаблона.

Вариант 2: В этом варианте формирование XML-подобной структуры переносится на SQL сервер. Для этого на SQL сервере создается хранимая процедура с функциями описанного выше шаблона. В качестве входного параметра эта процедура получает SQL-запрос на представление данных. Процедура должна вернуть таблицу с одной колонкой (тип данных – *text*), каждая строка которой представляет собой XML-представление записи. Ниже приведен пример такой процедуры (см. файл в каталоге тестовой базы данных mspubs ./data/mspubs/mk_xml_rusmarc.sql):

```

use pubs
if exists (select * from dbo.sysobjects
where id = object_id(N'[dbo].[mk_xml_rusmarc]') and
OBJECTPROPERTY(id, N'IsProcedure') = 1)
drop procedure [dbo].[mk_xml_rusmarc]
GO

SET QUOTED_IDENTIFIER OFF
GO
SET ANSI_NULLS OFF
GO

CREATE PROCEDURE [dbo].[mk_xml_rusmarc] @sQuery varchar(2048)
WITH RECOMPILE AS
begin
    declare @k1 int, @k2 int,
            @title_id varchar(10) ,
            @pub_id varchar(10) ,
            @out varchar(8000)
    create table #t1 ( rec text )
    exec("declare cur1 scroll cursor for " + @sQuery)
    open cur1
    select @k1=@@cursor_rows

    if @k1>0
        begin
            fetch next from cur1 into @title_id,@pub_id
loop1:
            if (@@fetch_status <> -2)
                begin
                    select @out =
-----
'<rusmarc>' +
'<mrk len="24">' +
'<m_0_4 len="5">00000</m_0_4>' +
'<m_5_5 len="1">n</m_5_5>' +
'<m_6_6 len="1">a</m_6_6>' +
'<m_7_7 len="1">m</m_7_7>' +
'<m_8_8 len="1">#</m_8_8>' +
'<m_9_9 len="1">#</m_9_9>' +
'<m_10_10 len="1">2</m_10_10>' +
'<m_11_11 len="1">2</m_11_11>' +
'<m_12_16 len="5">00000</m_12_16>' +
'<m_17_17 len="1">#</m_17_17>' +
'<m_18_18 len="1">#</m_18_18>' +

```

```

'<m_19_19 len="1">#</m_19_19>' +
'<m_20_23 len="4">450#</m_20_23>' +
'</mrk>' +
'<001>Microsoft/pubs/' + @title_id + '</001>' +
'<010><##><d>' + (select CONVERT(varchar,price) from titles
                    where title_id=@title_id) + '</d></##></010>' +
'<100><##>' +
'<a len="36">' +
'<p100_0_7 len="8">' +
    CONVERT(varchar,getdate(),112) + '</p100_0_7>' +
'<p100_8_8 len="1">d</p100_8_8>' +
'<p100_9_12 len="4">' + (select
    SUBSTRING(CONVERT(varchar,pubdate,112),1,4) from titles
    where title_id=@title_id) + '</p100_9_12>' +
'<p100_13_16 len="4">|||</p100_13_16>' +
'<p100_17_19 len="3">k||</p100_17_19>' +
'<p100_20_20 len="1">y</p100_20_20>' +
'<p100_21_21 len="1">0</p100_21_21>' +
'<p100_22_24 len="3">eng</p100_22_24>' +
'<p100_25_25 len="1">y</p100_25_25>' +
'<p100_26_29 len="4">0102</p100_26_29>' +
'<p100_30_33 len="4">|||</p100_30_33>' +
'<p100_34_35 len="2">ca</p100_34_35>' + '</a></##></100>' +
'<101><0#><a>eng</a></0#></101>' +
'<102><##><a>' +
    (select country from publishers where pub_id=@pub_id) +
'</a></##></102>' +
'<105><##>' + '<a len="13">' +
'<p105_0_3 len="4">y|||</p105_0_3>' +
'<p105_4_7 len="4">|||</p105_4_7>' +
'<p105_8_8 len="1">0</p105_8_8>' +
'<p105_9_9 len="1">0</p105_9_9>' +
'<p105_10_10 len="1">0</p105_10_10>' +
'<p105_11_11 len="1">y</p105_11_11>' +
'<p105_12_12 len="1">y</p105_12_12>' + '</a></##></105>' +
'<200><1#><a>' +
    (select title from titles where title_id=@title_id) +
'</a></1#></200>' +
'<210><1#>' +
    (select '<a>' + city + '</a>'
     from publishers where pub_id=@pub_id) +
    (select '<c>' + pub_name + '</c>'
     from publishers where pub_id=@pub_id) + (select '<d>' +
    SUBSTRING(CONVERT(varchar,pubdate,112),1,4) + '</d>'
     from titles where title_id=@title_id) + '</1#></210>' +
'<330><##><a>' + (select notes from titles
                    where title_id=@title_id) + '</a></##></330>'

-- For Author
declare @au_ord tinyint, @au_fname varchar(20),
        @au_lname varchar(40), @address varchar(40),
        @city varchar(20), @state char(2), @zip char(5)
declare cur2 scroll cursor for select ta.au_ord,au.au_fname,
        au.au_lname, au.address, au.city, au.state, au.zip
from titleauthor ta, authors au
where ta.au_id=au.au_id and ta.title_id=@title_id
open cur2
select @k2=@@cursor_rows
if @k2>0
begin

```

```

        fetch next from cur2 into @au_ord, @au_fname,
            @au_lname, @address, @city, @state, @zip
loop2:
        if (@@fetch_status <> -2)
            begin
        if @au_ord = 1
            begin
                select @out=@out+'<700><##>' +
                    '<a>'+ @au_lname +'</a>'+
                    '<b>'+@au_fname+'</b>'+
                    '<p>'+@address+' '+@city+' '+@state+' '+@zip+'</p>'+
                    '</##></700>'
                end
            else
                begin
                    select @out=@out+'<701><##>' +
                        '<a>'+ @au_lname +'</a>'+
                        '<b>'+@au_fname+'</b>'+
                        '<p>'+ @address + ' '+@city + ' '+@state + ' ' +@zip +'</p>'+
                        '</##></701>'
                    end
                end
                fetch next from cur2 into @au_ord, @au_fname, @au_lname,
                    @address, @city, @state, @zip
                if (@@fetch_status <> -1) goto loop2
            end
            close cur2
            deallocate cur2
-- End Authors
select @out=@out+
    '<801><#0>' +
    '<a>US</a>' +
    '<b>Microsoft</b>' +
    '<c>' +CONVERT(varchar,getdate(),112)+ '</c>' +
    '<2>rusmarc</2>' +
    '</#0></801>' +
    '<899><##>' +
    '<b>Тестовая база для проверки провайдера'+
    ' Z-MSSQL для ZooPARK v 3.00 (ОИГМ СО РАН, 2002)</b>'+
    '</##></899>'+
    '</rusmarc>'
-----
    if (@out is not null) insert #t1 select rec=@out
        end
        fetch next from cur1 into @title_id,@pub_id
        if (@@fetch_status <> -1) goto loop1
    end
    close cur1
    deallocate cur1
    select * from #t1
    drop table #t1
end
GO

SET QUOTED_IDENTIFIER OFF
GO
SET ANSI_NULLS ON
GO

```

Для настройки формирования представления данных по варианту 2

необходимо:

1. Создать на *MS SQL* сервере хранимую процедуру, для тестовой БД `msspubs` достаточно исполнить скрипт `./msspubs/mk_xml_rusmarc.sql`.
2. В секции определения формата настроечного файла следует заменить настройки варианта 1

```
[rusmarc]
Type=shablon
Shablon=./msspubs/msspubs_rusmarc_f.txt
B=./msspubs/msspubs_rusmarc_b.txt
F=./msspubs/msspubs_rusmarc_f.txt
```

на настройки варианта 2:

```
[rusmarc]
Type=procedure
Shablon=pubs.dbo.mk_xml_rusmarc
```

Следует заметить, что вариант 2 намного более эффективен, чем вариант 1. В заключение следует подчеркнуть, что в текущей версии провайдера данных его реализации для *UNIX* и *Windows NT* отличаются между собой.

Провайдер данных Z-MSADO

Провайдер данных *Z-MSADO* (`dprv_ado`) для доступа к данным через интерфейсы *MS ADO* требует специального настроечного файла для каждой базы данных. Структура этого файла аналогична структуре настроечных файлов для провайдера данных *MS SQL Server*. Имя этого файла указывается в параметре *path* блока настройки баз данных файла конфигурации сервера. Пример правильной настройки:

```
tfox.type:      4
tfox.codepage:  1
tfox.path:      ./tfox/tfox.z39
tfox.index:     7
```

Здесь параметр *index* соответствует номеру провайдера данных *Z-MSADO* в файле `provider.ini`. Настроечный файл (в примере `amspubs.z39`) имеет структуру

```
[section1]
parametr1=value1
parametr1=value2

[section2]
. . .
```

В следующей таблице приведены имена секций и параметров настроечного файла:

DataBase – секция, описывающая базу данных:

Параметр	Значение
Name	Имя базы данных
ConnStr	ConnectionString для соответствующего <i>ADO</i> или <i>ODBC</i> драйвера
QuerySearch	Запрос для поиска данных
QueryPresent	Запрос для представления данных
QueryCount	Запрос для определения количества записей

Например,

```
[DataBase]
Name=tfoc
ConnStr=Driver={Microsoft Visual FoxPro Driver};UID=;
        SourceDB=E:\z3950\data\tfox\data\BOOKS.DBC;
        SourceType=DBC;Exclusive=No;BackgroundFetch=Yes;
        Collate=Machine;Null=Yes;Deleted=Yes;
QuerySearch=select COUNT(*) AS a from titles where 1=1 and
QueryPresent=select title_id,title,pub_id, price,notes,
        YEAR(pubdate) pdate from titles where 1=1 and
QueryCount=select COUNT(distinct *) AS a from titles
```

Bib1_1 – в данной секции каждому значению поискового атрибута *USE* ставится в соответствие название поля в таблице, например

```
[Bib1\_1]
4=title
12=title\_id
```

Bib1_d – `<val> [<list>]` – в данной секции каждому значению поискового атрибута *USE* ставится в соответствие тип данных – с (текст), d (data), y(год), i (целое число). По умолчанию все поля имеют тип с. Необязательный параметр `list` содержит список допустимых атрибутов в виде `typ:value;[typ:value;]`, например

```
4=c 5:1;5:2;5:3;5:100;
12=i
31=y 2:1;2:2;2:3;4:4;5:100;
```

Структура файла-шаблона *XML* для этого провайдера данных полностью соответствует соответствующему шаблону провайдера данных *Z-MSSQL*.

Провайдер данных *Z-MMARK*

Провайдер данных *Z-MMARK* предназначен для доступа к данным АИБС *MARK-SQL*, хранящимся на сервере *MS SQL Server*. Провайдер данных *Z-MMARK* (`dprv_mmark.dll`, `dprv_mmark.so`) требует специального настроечного файла для каждой базы данных. Структура этого файла аналогична структуре настроечных файлов для других провайдеров данных. Имя настроечного файла указывается в параметре `path` блока настройки баз данных файла конфигурации сервера. Пример правильной настройки:

```
mmark.type: 4
mmark.codepage: 1
mmark.path: ./mmark/mmark.z39
mmark.index: 9
```

Здесь параметр *index* соответствует номеру провайдера данных *dprv_mmark* в файле *provider.ini*. Настроечный файл (в примере *mmark.z39*) имеет структуру

```
[section1]
parametr1=value1
parametr1=value2

[section2]
. . .
```

В следующей таблице приведены имена секций и параметров настроечного файла:

DataBase – секция, описывающая базу данных:

Параметр	Значение
Name	Имя базы данных
Server	Имя сервера
IP_address	IP-адрес сервера
Port	TCP-порт SQL сервера
TDS_version	Версия протокола (только для UNIX)
Login	Имя входа на SQL сервер
Passwd	Пароль
SqlName	Имя БД для SQL сервера
AnsiToOem	Признак конвертирования данных на уровне сетевой библиотеки (только для Windows)
QueryCount	Запрос полного количества записей в БД
QuerySearch	Запрос для поиска данных
QueryPresent	Запрос для представления данных

Пример приведен ниже.

Bib1_1 – `<q1> SCAN: <q2>` – в данной секции каждому значению поискового атрибута *USE* ставится в соответствие два запроса – для поиска (извлечения) и для просмотра индексов. Запросы могут содержать макросы (см. пример ниже).

Bib1_2 – в качестве значения параметров указываются значения макроса `$att2$` для подстановок в секции *Bib1_4* (см. пример ниже *Bib1_4 <value>*). В качестве значения параметров указываются значения макроса `exp_att4` для подстановок в секции *Bib1_1* см. пример ниже).

Bib1_Type – `<value> [<list>]` – в данной секции каждому значению поискового атрибута *Use* ставится в соответствие тим данных – с

(текст), d (data), y(год), i (целое число). По умолчанию все поля имеют тип с. Необязательный параметр `list` содержит список допустимых атрибутов в виде `typ:value;[typ:value;]` (см. пример ниже).

Schema – секция, описывающая схемы данных:

Параметр	Значение
absyn	Описание формата хранения данных
rusmarc	Ссылка на таблицу конвертации usmarc-rusmarc
unimarc	Ссылка на таблицу конвертации usmarc-unimarc

Для макроподстановок в секции `Bib1_1` используются макросы:

Макрос	Секция	Комментарий
<code>\$term\$</code>	Bib1-1	Подставляется значение поискового термина
<code>\$exp_att4\$</code>	Bib1-1	Подставляется значение параметра для соответствующего поискового атрибута поискового термина
<code>\$att2\$</code>	Bib1-4	Подставляется значение параметра для соответствующего поискового атрибута поискового термина

Ниже приведен пример настроечного файла `mmark.z39` для тестовой базы `mmark` (значения параметров должны быть записаны одной строкой)

```
[DataBase]
Server=geolibr.uiggm.nsc.ru
IP_address=194.226.185.16
Port=1433
TDS_Version=ver7.0
Login=www
Passwd=www
Name=mmark
SqlName=marc
QuerySearch=select a=COUNT(c.DOC_ID) from marc.dbo.DOC c
QueryPresent=select c.* from marc.dbo.DOC c
QueryTag=select TAG,SUBTAG,CAPTION from marc.dbo.TAG
QueryNumrec=select COUNT(c.DOC_ID) from marc.dbo.DOC c
AnsiToOem=NONE
```

```
[AttSet]
Bib-1=Attr_Bib1
```

```
[Attr_Bib1]
0=Attr_Bib1_default
1=Attr_Bib1_1
2=Attr_Bib1_2
3=Attr_Bib1_3
4=Attr_Bib1_4
5=Attr_Bib1_5
6=Attr_Bib1_6
d=Attr_Bib1_1_Type
```

```
[Attr_Bib1_default]
1=1035
2=3
```

3=3
 4=2
 5=100
 6=3

[Attr_Bib1_1]

4=(c.DOC_ID in (select b.DOC_ID from marc.dbo.IDX245a a,
 marc.dbo.IDX245aX b where a.IDX_ID=b.IDX_ID and \$exp_att4\$))
 SCAN:select UPPER(TERM),CNT from marc.dbo.IDX245a where
 UPPER(TERM)>=UPPER('\$term\$')
 31=(c.DOC_ID in (select b.DOC_ID from marc.dbo.IDX260c a,
 marc.dbo.IDX260cX b where a.IDX_ID=b.IDX_ID and \$exp_att4\$))
 SCAN:select UPPER(TERM),CNT from marc.dbo.IDX260c
 where UPPER(TERM)>=UPPER('\$term\$')
 1003=(c.DOC_ID in (select b.DOC_ID from marc.dbo.IDX100a a,
 marc.dbo.IDX100aX b where a.IDX_ID=b.IDX_ID and \$exp_att4\$))
 SCAN:select UPPER(TERM),CNT from marc.dbo.IDX100a where
 UPPER(TERM)>=UPPER('\$term\$')
 1035=(c.DOC_ID in (select b.DOC_ID from marc.dbo.IDX245a a,
 marc.dbo.IDX245aX b where a.IDX_ID=b.IDX_ID and \$exp_att4\$))
 SCAN:select UPPER(TERM),CNT from marc.dbo.IDX245a where
 UPPER(TERM)>=UPPER('\$term\$')

[Attr_Bib1_1_Type]

4=c 5:1;5:2;5:3;5:100;
 31=y 4:1;4:2;4:4;4:5;4:109;2:1;2:2;2:3;2:4;5:1;5:2;5:3;5:100;
 1003=c 5:1;5:2;5:3;5:100;
 1035=c 5:1;5:2;5:3;5:100;

[Attr_Bib1_2]

1=<
 2=<=
 3==
 4=>=
 5=>
 6!=

[Attr_Bib1_3]

3=any

[Attr_Bib1_4]

1=PATINDEX(UPPER('\$term\$'),UPPER(a.TERM)) > 0
 2=PATINDEX(UPPER('\$term\$'),UPPER(a.TERM)) > 0
 4=CONVERT(int,CASE WHEN ISNUMERIC(a.TERM)=1
 THEN a.TERM END) \$att2\$ \$term\$
 5=CONVERT(int,CASE WHEN ISNUMERIC(a.TERM)=1
 THEN a.TERM END) \$att2\$ \$term\$
 108=UPPER(a.TERM) like UPPER('\$term\$')
 109=CONVERT(int,CASE WHEN ISNUMERIC(a.TERM)=1
 THEN a.TERM END) \$att2\$ \$term\$

[Attr_Bib1_5]

1=%
 2=%
 3=%
 100=~

[Attr_Bib1_6]

3=complete

```
[Schema]
absyn=usmarc.abs
rusmarc=./mmark/Us2Rus.cnv
unimarc=./mmark/Us2Uni.cnv

[rtf]
tabl=~
rec_beg=~
rec_end=~
tag_beg=\\pard\\li%u\\widctlpar\\adjustright{\\cf2\\b{
tag_end=: }}
dat_beg=~
dat_end=\\par
lev_beg=\\par
lev_end=~

[html]
tabl=~
rec_beg=<table border=0 width="100%"><tr><td>
rec_end=</td></tr></table><br>
tag_beg=<FONT color=#8b0000><b>
tag_end=:&nbsp;</b></FONT>
dat_beg=&nbsp;<br>
dat_end=<br>
lev_beg=<table border=0><tr><td width=15>&nbsp;</td><td>
lev_end=</td></tr></table>
```

6.2.5 Параметры запуска сервера

Задача сервера может быть запущена с различными параметрами. Ниже приведена таблица допустимых параметров командной строки:

```
# ../bin/ZooPARK [parameters] tcp:@:port
```

Таблица 6.8: Параметры запуска сервера *ZooPARK*

Параметры	Комментарий	Default
-a <file>	Указывает направлять dump APDU в файл <file>. Специальное имя '-' перенаправляет вывод на stderr	Нет
-c <config-file>	Указывает использовать файл конфигурации сервера <config-file>	./zserver.cfg
-l <log-file>	Указывает использовать log-файл <log-file>	stderr
-n <stat-file>	Указывает использовать файл Z-статистики <stat-file>	stderr
-g <wstat-file>	Указывает использовать файл WEB-статистики <wstat-file>	stderr
-v <log-level>	Указывает уровень записи log-файла. Допустимы значения (fatal, log, debug, warn, prov, all, none). Допустим список, разделенный запятой.	log

-w <directory>	Устанавливает рабочий каталог <directory>	./
-i <minutes>	Запуск через inetd	Нет
-t <timeout>	Устанавливает timeout для сессии	60
-u <username>	Устанавливает UID серверного процесса	
-e	Отменяет обработку Exceptions (для Win NT,2000)	Обработка разрешена

Пример командной строки запуска сервера:

```
# ../bin/ZooPARK -l none -n ../stat.txt -g ../wstat.txt
-w /usr/local/ZooPARK/data -t 30 tcp:@:210
```

6.2.6 Сбор статистической информации

Начиная с версии 4.05, сервер протоколирует все входящие и выходящие *APDU* в файле статистики, а с версии 4.06 может передавать информацию на специально сконфигурированный сервер *ZooPARK* (Log-сервер). Передача происходит по протоколу *HTTP* (*PUT*) в асинхронном режиме.

Настройка сбора статистики

Конфигурация сбора статистики осуществляется при помощи текстового файла *stat.cfg*, который находится в каталоге *data*. Файл статистики может быть указан в командной строке при запуске сервера (-n имя_файла). По умолчанию статистика пишется в стандартный выходной поток *stderr*.

Таблица 6.9: Параметры конфигурационного файла сбора статистики *stat.cfg*

Параметры	Комментарий	Default
default	serverId Srvname Srv_ip Type File Log_srv	Числовой идентификатор сервера в РИС Имя сервера IP-адрес сервера Тип сбора статистики: 0 – запись в локальный файл; 3 – передача на Log-сервер Имя локального файла для Type=0 Адрес и порт Log-сервера для Type=3
DataBase	Provider Server IP_address	Динамическая библиотека провайдера данных, которому Log-сервер передает информацию для сохранения в БД Имя SQL сервера Адрес SQL сервера

	Port TDS_Version ² Login ² Passwd ² Name ² SqlName ²	Порт SQL сервера Версия протокола SQL сервера Login для SQL сервера Пароль для SQL сервера Имя БД статистики для Z39.50 Имя БД статистики для SQL сервера
--	--	--

Пример конфигурационного файла статистики:

```
[default]
serverId=100
Srvname=zhizhimov
Srv_ip=194.226.185.131
Type=3
File=./zstat.txt
Log_srv=zhizhimov.uiggm.local:210

[DataBase]
Provider=dprv_mssql.dll
Server=geolibr
IP_address=194.226.185.16
Port=1433
TDS_Version=ver7.0
Login=wwwwww
Passwd=wwwwww
Name=stat_mssql
SqlName=zstat
```

где в секции `[default]` перечисляются идентификационное имя сервера, его *IP*-адрес, тип протоколирования (для сбора в простой файл или выходной поток *stderr* параметр *Type* должен быть равен 0) и имя файла относительно каталога *data*. В приведенном примере конфигурационный файл соответствует настройке сервера *ZooPARK* для передачи статистической информации на другой сервер *ZooPARK* (*Log_srv=geolibr.uiggm.local:210*). Секция `[DataBase]` должна присутствовать только в случае, когда *ZooPARK* конфигурируется как сервер сбора статистики (*Log_srv*) с возможностью сохранения этой статистики в реляционной СУБД (в приведенном примере – в СУБД *MS SQL Server*, БД *zstat*).

Формат файла статистики

Файл статистики представляет собой обычный текстовый файл со строками переменной длины, каждая строка файла описывает один *APDU*. Поля в строке разделены знаком табуляции. Количество полей для описания *APDU* разного типа различно, но одинаково для *APDU* одного типа. Возможное отсутствие значений в поле обозначается знаком ”-”. Заполняемые поля делятся на общие для *APDU* любого типа и специфические для *APDU* каждого типа. Ниже перечислены по порядку все поля записи файла статистики.

²Параметр актуален для провайдера данных Z-MSSQL

Общие поля :

- 1 – Объект. Указывает протоколируемый модуль: 0 – сервер, 1 – провайдер *Z-REMOTE*, 2 – шлюз *Z-GW*. Тип: число. Пример: 0.
- 2 – Идентификатор сессии. Уникальная последовательность символов, идентифицирующая сессию. Обычно имеет формат: время-ид. потока (процесса). Тип: текст. Пример: 20031030101441-fa4
- 3 – Текущее время. Тип: норм. дата-время. Формат: *YYYYMMDDhhmmss*. Пример: 20031030101441
- 4 – Сессионное время. Количество секунд от начала сессии. Тип: число. Пример: 11
- 5 – Тип *APDU*. Числовой идентификатор *APDU*:

1 = initReq	13 = triggerResCntrReq
2 = initResp	14 = resourceReportReq
3 = searchReq	15 = resourceReportResp
4 = searchResp	16 = scanReq
5 = presentReq	17 = scanResp
6 = presentResp	18 = sortReq
7 = delResultSetReq	19 = sortResp
8 = delResultSetResp	20 = segmentReq
9 = accessCntrReq	21 = extendServicesReq
10 = accessCntrResp	21 = extendServicesResp
11 = resourceCntrReq	23 = close
12 = resourceCntrResp	24 = duplicatDetectReq
	25 = duplicatDetectResp
- Тип: число. Пример: 3.
- 6 – Размер *APDU*. Количество байт в *APDU*. Тип: число. Пример: 352
- 7 – Направление и название *APDU*. Направление обозначается буквами S (Send) и R (Receive). Знак ":" отделяет эту букву от названия *APDU*. Тип: текст. Пример: S:present

Начиная с поля 8 следуют специфические для каждого типа *APDU* поля.

initRequest :

- 8 – implementationName. Тип: текст.
Пример: Zoo/PARK (with IndexData/YAZ ToolKit)
- 9 – implementationId. Тип: текст. Пример: 155 (UIGGM SB RAS)
- 10 – implementationVersion. Тип: текст. Пример: 4.05
- 11 – maximumRecordSize. Тип: число. Пример: 1048576
- 12 – preferredMessageSize. Тип: число. Пример: 524288
- 13 – user. Имя пользователя при авторизованном входе. Тип: текст.
Пример: zhizhimov

- 14 – Список языков. Названия языков в списке разделены ”;”. Тип: текст. Пример: rus;eng;
- 15 – Список кодировок символов. Названия кодировок в списке разделены ”;”. Тип: текст. Пример: dos;koi;
- 16 – recordsInSelectedCharSets. Тип: число. Пример: 1
- 17 – IP адрес клиента. Тип: текст. Пример: 194.226.185.2
- 18 – Маршрут прохождения запроса. IP-адреса серверов в списке разделены ”;”. Тип: текст. Пример: 194.226.185.2; 194.226.184.14;
- 19 – Максимальное количество переадресаций запроса. Тип: число. Пример: 10
- 20 – Максимальное количество петель переадресаций. Тип: число. Пример: 3
- 21 – Текущая переадресация. Тип: число. Пример: 1
- 22 – Текущая петля переадресации. Тип: число. Пример: 0

initResponse :

- 8 – implementationName. Тип: текст.
Пример: ZooPARK (with IndexData/YAZ ToolKit)
- 9 – implementationId. Тип: текст. Пример: 155 (UIGGM SB RAS)
- 10 – implementationVersion. Тип: текст. Пример: 4.05
- 11 – maximumRecordSize. Тип: число. Пример: 1048576
- 12 – preferredMessageSize. Тип: число. Пример: 524288
- 13 – Выбранный язык. Название языка. Тип: текст. Пример: rus
- 14 – Выбранная кодировка символов. Название кодировки. Тип: текст. Пример: dos;
- 15 – recordsInSelectedCharSets. Тип: число. Пример: 1

searchRequest :

- 8 – databaseNames. Имена БД в списке разделены ”;”. Тип: текст.
Пример: geocat;tgils;
- 9 – preferredRecordSyntax. Тип: текст. Пример: RUSmarc
- 10 – smallSetUpperBound. Тип: число. Пример: 0
- 11 – largeSetLowerBound. Тип: число. Пример: 1
- 12 – mediumSetPresentNumber. Тип: число. Пример: 0
- 13 – smallSetElementSetNames. Тип: текст. Пример: F
- 14 – mediumSetElementSetNames. Тип: текст. Пример: B
- 15 – resultSetName. Тип: текст. Пример: Set_1
- 16 – queryType. Тип запроса. Тип: текст. Пример: RPN
- 17 – attributeSetId. Тип: текст. Пример: Bib-1

18 – query. Строка запроса, PQF-представление RPN-запроса. Тип: текст. Пример: @and 5=1 "mineral" 1=4 "geo"

searchResponse :

8 – errorCode. Номер ошибки по Vib-1. Тип: число. Пример: 114

9 – searchStatus. Тип: число. Пример: 1

10 – resultCount. Тип: число. Пример: 1230

11 – presentStatus. Тип: число. Пример: 0

12 – numberOfRecordsReturned. Тип: число. Пример: 10

13 – recordCountbyDB. Список БД с указанием количества извлеченных записей. Тип: текст. Пример: geocat(5);tgils(6);

presentRequest :

8 – resultSetId. Тип: текст. Пример: Set_1

9 – resultSetStartPoint. Тип: число. Пример: 1

10 – numberOfRecordsRequested. Тип: число. Пример: 12

11 – preferredRecordSyntax. Тип: текст. Пример: GRS-1

12 – recordSyntax. Список допустимых форматов . Тип: текст. Пример: USmarc;RUSmarc;

13 – elementSetName. Тип: текст. Пример: F

14 – schema. Тип: текст. Пример: CIMI-schema

presentResponse :

8 – errorCode. Номер ошибки по Vib-1. Тип: число. Пример: 238

9 – presentStatus. Тип: число. Пример: 0

10 – numberOfRecordsReturned. Тип: число. Пример: 10

11 – recordCountbyDB. Список БД с указанием количества извлеченных записей. Тип: текст. Пример: geocat(5);tgils(6);

close :

8 – closeReason. Тип: число. Пример: 0

9 – diagnosticInformation. Тип: текст.

6.3 Explain

В соответствии со стандартом *Z39.50* база метаданных *IR-Explain-1* является обычной базой данных. Поиск в этой базе осуществляется по набору атрибутов *Exp-1*, а извлечение записей из нее возможно, как минимум, в формате *Explain*.

6.3.1 Категории

Текущая версия сервера поддерживает следующие категории *Explain*:

CategoryList – список категорий *Explain*, поддерживаемых конкретным сервером. Этот список может быть лишь частью списка категорий в этой таблице.

TargetInfo – информация о конкретном сервере. Должна существовать одна запись этой категории.

DatabaseInfo – информация о каждой базе данных, поддерживаемой конкретным сервером. Количество записей этой категории совпадает с количеством баз данных.

SchemaInfo – информация о каждой схеме данных, поддерживаемой конкретным сервером. Количество записей этой категории совпадает с количеством поддерживаемых схем.

TagSetInfo – информация о каждом наборе меток, поддерживаемом конкретным сервером. Количество записей этой категории совпадает с количеством поддерживаемых наборов.

RecordSyntaxInfo – информация о каждом формате внешнего представления, поддерживаемом конкретным сервером. Количество записей этой категории совпадает с количеством форматов.

AttributeSetInfo – информация о каждом наборе поисковых атрибутов, поддерживаемом конкретным сервером. Количество записей этой категории совпадает с количеством поддерживаемых наборов.

AttributeDetails – информация о каждом поисковом атрибуте. Количество записей этой категории совпадает с количеством баз данных.

ElementSetDetails – информация о каждом наборе элементов, поддерживаемом конкретным сервером. Запись этой категории существует для каждого формата внешнего представления и для каждой базы данных.

RetrievalRecordDetails – информация об элементах в каждом формате внешнего представления. Запись этой категории существует для каждого формата внешнего представления, для каждой базы данных и каждой схемы.

6.3.2 Как создать базу данных *IR-Explain-1*

Формально, база данных *IR-Explain-1* может существовать под управлением любой СУБД. Соответственно, доступ к этой базе может осуществляться через любого провайдера данных. Основное требование – база данных *IR-Explain-1* должна отражать реальную конфигурацию сервера и его возможностей, с одной стороны, а с другой – провайдер данных должен

поддерживать набор поисковых атрибутов *Exp-1* и формат внешнего представления *Explain*. Поддержка других форматов для базы *IR-Explain-1* является необязательной. Тем не менее, мы рекомендуем на начальном этапе следовать правилу поддержки базы данных *IR-Explain-1* в виде текстовых *SGML*-файлов, проиндексированных при помощи инструментов *Zebra* [114] компании *IndexData*. При этом доступ к этой базе данных со стороны сервера будет осуществляться через провайдер данных *Zebra*.

Внимание! В отличие от оригинального комплекта *Zebra* компании *IndexData* текущая версия описываемого сервера *ZooPARK* не поддерживает автоматическую генерацию базы данных *IR-Explain-1* даже для данных провайдера *Zebra*. Создание и модификация данных *Explain* должны производиться вручную. Автоматическое наполнение некоторых полей, тем не менее, существует (см. ниже).

При создании данных *Explain* в виде *SGML*-файлов следует применять схему данных *Explain*. Сервер поддерживает эту схему данных с локально определенным *OID* (*Z.13.1000.81.2*). Описание схемы данных *Explain* можно найти в каталоге `tab/` в файлах `explain.abs`, `explain.att`, `explain.tag`. В дистрибутиве сервера данные *Explain* расположены в каталоге `data/IR-Explain-1/data`. Можно отредактировать данные обычным текстовым редактором и произвести индексацию данных:

```
# cd /usr/local/ZooPARK/data/IR-Explain-1
# ./index.sh
```

Внимание! Перед выполнением индексации данных убедитесь, что в каталоге отсутствуют резервные и старые копии файлов данных. В противном случае данные из этих файлов также будут проиндексированы.

6.3.3 Пример базы данных

Файл `CategoryList.sgm1`

```
<explain>
  <categoryList> categoryList
    <commonInfo>
      <dateAdded>19990322115051</dateAdded>
      <dateChanged>19990322115051</dateChanged>
      <languageCode>EN</languageCode>
    </commonInfo>
    <categories>
      <category><name>CategoryList</name></category>
      <category>
        <name>TargetInfo</name>
        <description>Информация о сервере</description>
      </category>
    </categories>
  </categoryList>
</explain>
```

Файл TargetInfo.sgml

```

<explain>
  <targetInfo> targetInfo
    <welcomeMessage> Hello! This is GEOLIBR (UIGGM
      SB RAS)!</welcomeMessage>
    <commonInfo>
      <dateAdded>19981218</dateAdded>
      <dateChanged>19981218</dateChanged>
      <languageCode>eng</languageCode>
      <otherInfo> otherInfo </otherInfo>
    </commonInfo>
    <name> Geolibr </name>
    <recentNews> No </recentNews>
    <namedResultSets> 1 </namedResultSets>
    <multipleDbSearch> 1 </multipleDbSearch>
    <maxResultSets> 10 </maxResultSets>
    <maxResultSize> 64000 </maxResultSize>
    <maxTerms>10</maxTerms>
    <description>GEOLIBR is Z39.50-95 server for Windows NT
      (Z-WinNT-Srv/UIGGM SB RAS Id=155 with YAZ/IndexData
      Toolkit). Data providers: CDS/ISIS, Remote Z-data
      provider (use YAZ/IndexData Toolkit). </description>
    <nicknames>
      <name>GEOLIBR</name>
    </nicknames>
    <contactInfo>
      <name> Zhizhimov Oleg L. </name>
      <description>Head of Sector of Informational
        Technology</description>
      <email>zhizhim@uiggm.nsc.ru</email>
      <address>Koptuyug pr.3, Novosibirsk, 630090,
        Russia</address>
      <phone>+7(383-2) 33-20-05</phone>
    </contactInfo>
    <addresses>
      <internetAddress>
        <host>geolibr.uiggm.nsc.ru</host>
        <port>210</port>
      </internetAddress>
    </addresses>
    <languages>
      <language> eng </language>
      <language> rus </language>
    </languages>
  </targetInfo>
</explain>

```

При этом элемент

```
<otherInfo> otherInfo </otherInfo>
```

в `targetInfo/commonInfo` является фиктивным и заполняется динамически в момент запроса.

Эти два файла образуют базу данных с категориями `CategoryList` и `TargetInfo`. После индексации проверить ее работоспособность можно так:

```
# cd /usr/local/ZooPARK/bin - установим рабочий каталог
```

```

# ./zclient - выполним программу
z3950> codepage dos - установим кодовую таблицу
z3950> open tcp:localhost:210 - соединимся с сервером
. . . . .
z3950> base IR-Explain-1 - установим базу данных
z3950> attributeset Exp-1 - установим набор атрибутов
z3950> find @attr 1=1 CategoryList - выполним поиск
. . . . .
z3950> format sutrs - установим формат выдачи
z3950> show 1+1 - посмотрим запись
. . . . .
z3950> format explain - установим формат выдачи
z3950> show 1+1 - посмотрим запись
. . . . .
z3950> format grs1 - установим формат выдачи
z3950> show 1+1 - посмотрим запись
. . . . .
z3950> find @attr 1=1 TargetInfo - выполним поиск
. . . . .
z3950> format explain - установим формат выдачи
z3950> show 1+1 - посмотрим запись
. . . . .
z3950> close - завершим сеанс связи
z3950> quit - завершим задачу
#

```

6.3.4 Категории *Explain* для сервера ZooPARK

Сервер *ZooPARK* поддерживает следующие категории *Explain*: *categoryList*, *targetInfo*, *databaseInfo*, *attributeSetInfo*, *recordSyntaxInfo*, *attributeDetails*, *schemaInfo*, *tagsetInfo*. Файлы, соответствующие этим категориям могут присутствовать в каталоге *data/expl* и индексироваться в соответствии с общими правилами. Вывод записей может быть осуществлен в форматах: *SUTRS*, *Explain*, *GRS-1*, *RTF*, *XML*, *HTML*.

6.3.5 Автоматически заполняемые поля в записях *Explain*

Некоторые элементы в записях *Explain* могут быть заполнены автоматически, вернее, их заполнение происходит в момент извлечения записи из базы данных. Поскольку технология извлечения информации тесно связана с архитектурой соответствующего провайдера данных, каждый провайдер данных может иметь точку входа (экспортируемую функцию), ответственную за наполнение некоторых полей. Например, информация о динамических провайдерах данных (элемент *commonInfo/otherInfo* категории *targetInfo*) фиксируется в момент инициализации соответствующего провайдера данных. При этом заполняется три поля: название провайдера данных, номер его версии и дата генерации. Детали рассмотрены в документации [38] в разделе, посвященном разработке динамических провайдеров данных.

В текущей версии автоматически заполняются в категории *databaseInfo*

элементы: *commonInfo/dataChanged* и *recordCount/recordCountActual* для провайдеров данных *Zebra* и динамических, а также элемент *attributesBySet* в категории *attributeDetails* для динамических провайдеров данных. При этом динамический провайдер данных обязан экспортировать функцию *dll_explain*, получая в соответствующем параметре тип категории *Explain*. Детали механизма описаны в документации [38] в разделе, посвященном разработке динамических провайдеров данных. Что касается статического провайдера данных *Zebra*, то наполнение соответствующих элементов записи для него происходит автоматически. Провайдер данных *Zebra* также автоматически заполняет элементы *attributesBySet* категории *attributeDetails*. Для активизации этого механизма, достаточно наличия "заглушки" в соответствующем XML-файле. Дистрибутив поставки включает несколько баз данных, демонстрирующих описанные функции.

В последующих версиях сервера предполагается расширить возможность автоматического наполнения элементов записей *Explain*.

6.4 WEB сервер ZooPARK

Начиная с версии 4.00, сервер *ZooPARK* содержит встроенный сервер *WEB*. Функциональность этого сервера, конечно, не может сравниться с такими серверами как *Apache* или *MS IIS*, однако, она вполне достаточна для решения задач просмотра *HTML*-документов и организации пользовательских *WEB*-интерфейсов доступа к ресурсам *Z39.50*. В качестве эксперимента в версию 4.01 сервера *ZooPARK* включена частичная поддержка технологии *XML/SOAP/SRW*.

Начиная с версии 4.06, *WEB*-сервер *ZooPARK* включает поддержку внешних консольных приложений, применяемых для генерации *HTML*-страниц. Это позволяет использовать технологии *PHP*, *PERL* и т.п.

Ниже приведены основные сведения по конфигурированию *WEB*-сервера *ZooPARK*.

6.4.1 Настройка WEB-сервера

Для настройки *WEB*-сервера для предоставления доступа к *HTML*-документам, которые находятся, например, в каталоге */ZooPARK/data/doc*, необходимо в конфигурационном файле сервера *ZooPARK* (*zserver.cfg*) указать:

```
server.docroot:  ../data/doc - путь к каталогу данных
server.docfile:  index.htm - имя html документа по умолчанию
```

После внесения изменений в указанные параметры в соответствии с вашей конфигурацией в *WEB*-браузере по URL: *http://server.domain:210* вы должны увидеть начальную страницу, изображенную на Рис. 17 (см. раздел 6.1.8).

6.4.2 Настройка внешних модулей WEB-сервера

Если в вашей операционной системе поддерживается *PHP*, *PERL* или любой другой консольный интерпретатор, вы можете настроить *WEB*-сервер *ZooPARK* для обработки соответствующих скриптов. Для этого необходимо, во-первых, включить в файл `zserver.cfg` следующие строки:

```
server.shell: ../data/shell.cfg
server.shellext: .php:.asp:.pl:.tcl
```

первая из которых указывает на файл конфигурации внешних консольных приложений (`shell.cfg`), а вторая содержит перечисления расширений файлов, которые будут обрабатываться внешними приложениями. Во-вторых, необходимо создать упомянутый файл `shell.cfg`, который выглядит следующим образом

```
[php]
dir=/php/
cmd=e:/php/php.exe
arg= -q -c ./doc/php/php.ini

[asp]
dir=/php/
cmd=e:/php/php.exe
arg= -q -c ./doc/php/php.ini

[pl]
dir=/perl/
cmd=perl.exe
```

Здесь имя секции соответствует расширению файла, которые перечислены в `zserver.cfg` (`zserver.shellext`). Параметр `dir` указывает на каталог, в котором находятся соответствующие скрипты (относительно каталога документов *WEB*-сервера, в нашем случае относительно каталога `../data/doc/`), параметр `cmd` указывает на исполняемый файл соответствующего приложения, а параметр `arg` – на список аргументов командной строки.

В дистрибутивной поставке сервера *ZooPARK* присутствуют тестовые скрипты *PHP* (`test.php`, `test.asp`) и *PERL* (`test.pl`) в каталогах `data/doc/php` и `data/doc/perl` соответственно. При наличии у вас установленных *PHP* и *PERL* работоспособность соответствующих модулей можно проверить, как описано в разделах 6.1.9 и 6.1.10.

6.5 Встроенный шлюз Z-GW

Для пользователей информационных систем на основе протокола *Z39.50* сервер *ZooPARK*, начиная с версии 4.01, содержит встроенный шлюз *Z39.50-HTTP* (*Z-GW*), который не требует наличия отдельного *WEB*-сервера.

6.5.1 Основные характеристики шлюза Z-GW

Встроенный шлюз Z-GW представляет собой динамически загружаемую библиотеку (*zgw.dll*, *zgw.so*), которая явно вызывается ядром сервера ZooPARK на стадии выполнения. Такая организация позволяет серверу ZooPARK функционировать и в отсутствие динамической библиотеки шлюза, однако в последнем случае, естественно, функции шлюза будут отсутствовать.

Шлюз Z-GW поддерживает функции клиента Z39.50 версии 3. В текущей версии реализованы следующие функции:

- Генерация HTML-страниц на основе правил интерпретатора встроенного языка
- Поиск
- Извлечение данных в различных форматах внешнего представления (*SUTRS*, *GRS-1*, *USmarc*, *UNIMARC*, *RUSmarc*, *HTML*, *XML*)
- Представление данных формата *GRS-1* в соответствии с определенными шаблонами.
- Сохранение записей в различных форматах на компьютере клиента Web.
- Поиск и просмотр записей БД *IR-Explain-1* в формате *Explain*.
- Поиск в многосерверном окружении.
- Поиск и просмотр записей в схеме *Zthes-1* для тезаурусов и классификационных схем.
- Поддержка различных кодовых таблиц (*DOS*, *WIN*, *ISO*, *MAC*) для кириллических текстов

Входной информацией для шлюза служат:

- переменные, переданные шлюзу ядром сервера
- переменные, полученные по *GET* или *POST*
- переменные, прочитанные из файла конфигурации
- HTML-документы со встроенными командами интерпретатора
- файлы описания шаблонов *GRS-1*

Шлюз обрабатывает запросы методов *GET* и *POST*. Функции клиента Z39.50 локализованы в дополнительных динамически загружаемых библиотеках (*z-origin* и *yazm*). Шлюз использует динамическое распределение оперативной памяти, но не использует временных файлов на дисках. Конфигурирование шлюза производится через конфигурационный текстовый файл.

6.5.2 Настройка шлюза Z-GW

Для активизации шлюза *Z-GW* необходимо наличие в каталоге `bin` дополнительно двух динамических библиотек: `z_origin.dll` (`z_origin.so`) и `zgw.dll` (`zgw.so`). Настройка шлюза производится через два файла: конфигурационный файл сервера `zserver.cfg`) и конфигурационный файл шлюза.

Файл `zserver.cfg` в каталоге `data` должен содержать следующие строки (см. раздел 6.2.1):

```
server.gwdir: /zgw/;/zgwс/ - список каталогов шлюза в каталоге
                               WEB сервера
server.gwcfg: ../data/zgw.cfg - имя файла конфигурации шлюза
```

При этом подразумевается, что в каталоге данных *WEB*-сервера (`../data/doc`) находятся два подкаталога (`zgw` и `zgwс`), которые содержат шаблоны двух различных вариантов оформления шлюза. Именно такая конфигурация шлюза *Z-GW* включена в дистрибутив.

Для конфигурирования шлюза используется конфигурационный текстовый файл с именем `zgw.cfg`, расположенный, например, в каталоге `data`. Структура этого файла совпадает со структурой обычных `*.ini` файлов *Windows*:

```
[Имя\секция]
Переменная1=Значение
Переменная2=Значение
. . .
```

В качестве примера ниже приведен простейший файл конфигурации

```
# В секции [default] перечислены все возможные
# режимы работы шлюза. Имя режима работы шлюза
# должно быть передано через переменную conf.
# Параметр default определяет режим работы по
# умолчанию.

[default]
default=corpvars
zgw=zgwvars
thes=thesvars
corp=corpvars
expl=expvars
rubr=rubrvars
gas=gasvars
geo=geovars

# Список переменных и их значений по умолчанию для
# режима corp

[corpvars]
server=z3950.uiggm.nsc.ru
port=210
base=geocat
format=grs1
elem=F
attset=bib1
```



```
cmd1=none
query1=@attr 5=1
term1=geology
hits=0
pattern=uiggm
bnum=1
nums=1
nums0=1
panel=off
error=

# Список переменных и их значений по умолчанию для
# режима exp

[expvars]
server=z3950.uiggm.nsc.ru
port=210
base=IR-Explain-1
format=explain
elem=F
attset=exp1
cmd1=none
query1=@attr 1=1
query=@attr 1=1 categorylist
term1=categorylist
pattern=expl
hits=0
bnum=1
nums=1
nums0=1
panel=off
error=

# Список переменных и их значений по умолчанию для
# режима rubr

[rubrvars]
server=z3950.uiggm.nsc.ru
port=210
base=geotrn
format=grs1
elem=F
attset=bib1
cmd1=none
query1=@attr 5=1
term1=gas
pattern=uiggm
hits=0
bnum=1
nums=1
nums0=1
tserver=z3950-3.uiggm.nsc.ru
tport=210
tbase=vinitir
tformat=grs1
telem=F
tattset=ZTHES-attset
tcmd1=none
tquery1=@attr XD-1 1=1
```

```

tterm1=реология
tindex=1,14
thits=0
tbnum=1
tnums=1
tnums0=1
panel=off
error=

# Конец файла

```

Список инициализируемых по умолчанию переменных для каждого режима работы шлюза включает в себя обязательные переменные и необязательные. Количество переменных может быть расширено по усмотрению администратора шлюза. Список обязательных переменных приведен далее.

Для активизации гипертекстовых ссылок при выводе записей в различных форматах, в конфигурационном файле шлюза должны присутствовать дополнительные секции. Например, если при выводе в формате *RUSmarc* мы хотим гипертекстить поля 686, 856 и 675, мы должны описать ссылки:

```

# Секция, описывающая список форматов, для которых
# активизируются гипертекстовые ссылки

[links]
rusmarc=rusmarc_links

# Список поле для RUSMARK

[rusmarc_links]
686=rusmarc_link_686
856=rusmarc_url_856
675=rusmarc_link_675

# Описание каждого поля и подполя

[rusmarc_url_856]
type=url
term=u

[rusmarc_link_675]
type=rubr
name=v
term=a
1=ruudc3
2=ruudc3
3=ruudc3
none=ruudc3

[rusmarc_link_686]
type=rubr
name=2
term=a
rugasnti=rugasnti
rubbk=rubbk
ruvak=ruvak
vak=ruvak

```

```

ruagr=ruagr
local=local

[rugasnti]
link=sr4.htm?conf=rubr&
server=tserver=z3950-3.uiggm.nsc.ru&tport=210&tbase=gasnti&
cmd=tcmd1=fullterm&
query1=@attr+Zthes-attset+1=1+

[ruagr]
link=sr4.htm?conf=rubr&
server=tserver=z3950-3.uiggm.nsc.ru&tport=210&tbase=ruagr&
cmd=tcmd1=fullterm&
query1=@attr+Zthes-attset+1=1+

[rubbk]
link=srf.htm?conf=rubr&shablon=srf.htm&
server=tserver=z3950-3.uiggm.nsc.ru&tport=210&tbase=rubbki&
cmd=tcmd1=fullterm&
query1=@attr+Zthes-attset+1=1+

[local]
link=srf.htm?conf=rubr&shablon=srf.htm&
server=tserver=z3950-3.uiggm.nsc.ru&tport=210&tbase=local&
cmd=tcmd1=fullterm&
query1=@attr+Zthes-attset+1=1+

[ruvak]
link=sr4.htm?conf=rubr&shablon=sr4.htm&
server=tserver=z3950-3.uiggm.nsc.ru&tport=210&tbase=ruvak&
cmd=tcmd1=fullterm&
query1=@attr+Zthes-attset+1=1+

[ruudc3]
link=srf.htm?conf=rubr&shablon=srf.htm&
server=tserver=z3950-3.uiggm.nsc.ru&tport=210&tbase=ruudc&
cmd=tcmd1=fullterm&
query1=@attr+Zthes-attset+1=1+

```

6.5.3 Интерпретатор

Шлюз *Z-GW* представляет собой программу-интерпретатор специальных команд. Эти команды встраиваются в обычный *HTML*-файл в виде комментариев. Исполнение команд зависит от переменных, значения которых может быть изменено различными способами.

Команды

Для демонстрации работы интерпретатора ниже приведена простейшая *HTML*-страничка с именем `search_1.htm`, результатом обработки которой будет исполнение запроса поиска на сервере *Z39.50*:

```

<html><head><title>Test Search</title></head>
<!-- !SET endpar end -->
<!-- Если переменная hits отсутствует, заведем ее и присвоим
      значение 0: -->
<!-- !SET hits 0 -->

```

```

<body topmargin=0 bgcolor=#ffffff><center>
<!-- Если значение переменной cmd1 равно
"search", выполним блок до !ELSE и перейдем на строчку после
!FI, иначе перейдем на строчку после !ELSE -->
<!-- !IF cmd1 search -->
<!-- !INIT --> <!-- Создадим сеансовый блок Z39.50 -->
<!-- !OPEN --> <!-- Откроем сеанс связи Z39.50 -->
<!-- !FIND --> <!-- Выполним поиск -->
<!-- !CLOSE --> <!-- Закроем сеанс связи Z39.50 -->
<!-- !END --> <!-- Уничтожим сеансовый блок Z39.50 -->
<hr> Результат поиска на сервере
    $$server$: $$port$ (БД $$base$)<hr>
    Найдено записей:<b>$$hits$</b>. $$error$
<br><br>
<!-- !ELSE cmd1 search -->
<font size=+2> Select TERM and click to SEARCH! </font>
<!-- !FI cmd1 -->
</center></body></html>

```

Таблица 6.10: Команды шлюза Z-GW и переменные

Команда	Функция	Входные переменные	Выходные переменные
!SET aaa bbb	Создание переменной aaa со значением bbb. Игнорируется, если переменная aaa существует.		aaa
!INIT	Создает и инициализирует сеансовый блок	server port	
!OPEN	Открывает сеанс связи		
!FIND	Осуществляет поиск	base attset query	hits error
!SHOWREC	Извлекает записи в требуемом формате и помещает их в выходной поток	format elem pattern bnum nums records_begin records_end record_begin record_end	rnum error
!CLOSE	Закрывает сеанс Z39.50		
!END	Уничтожает сеансовый блок		
!SHOWVAR	Выводит в выходной поток значения всех переменных	все	

!TAR_INIT	Выполняет многосерверный поиск	num_targets targets allhits neg server port base descr attset query targets_find_begin targets_find_end targets_find_block targets_error tar_error_num tar_error_text	server port base neg descr allhits hits tar_error_num tar_error_text
-----------	--------------------------------	--	--

Как видно из приведенного примера, все команды начинаются с символа '!' в комментариях *HTML*. Для исполнения некоторых команд необходимо наличие переменных, например, имя сервера и базы данных. Некоторые команды при исполнении изменяют значения переменных, которые могут быть выведены в выходной поток. Например, команда `!FIND` изменяет значение переменной `$$hits$`. В Таблице 6.10 приведен список команд шлюза с указанием входных и выходных переменных для каждой.

Переменные

Переменные, которыми оперирует шлюз, можно разбить на три группы:

- простые переменные
- переменные-шаблоны
- шаблонные переменные

Простые переменные содержат информацию о конкретных параметрах соединения, например, имя сервера *Z39.50* (*server*), количество найденных записей (*hits*) и т.п. Список простых переменных приведен в таблице 6.11.

Таблица 6.11: Простые переменные шлюза ZGW

Переменная	Команды	Комментарий
conf	–	Имя набора переменных в файле конфигурации
server	!INIT !TAR_INIT	DNS-имя или IP-адрес сервера Z39.50
port	!INIT !TAR_INIT	Номер TCP-порта сервера Z39.50
base	!FIND !TAR_INIT	Имя базы данных или список имен, разделенных пробелом
attset	!FIND !TAR_INIT	Имя набора поисковых атрибутов

query	!FIND !TAR_INIT	Строка запроса
hits	!FIND !TAR_INIT	Количество найденных записей
elem	!SHOWREC	Имя набора элементов
format	!SHOWREC	Имя формата представления данных
schema	!SHOWREC	Имя схемы данных
bnum	!SHOWREC	Номер первой запрашиваемой записи в найденном множестве
nums	!SHOWREC	Количество запрашиваемых записей из найденного множества
rnum	!SHOWREC	Количество выданных записей в запросе на представление данных
error	!SHOWREC	Текст диагностического сообщения
pattern	!SHOWREC	Имя шаблона для дополнительного форматирования записей в формате GRS-1
neg	!INIT !TAR_INIT	Флаг разрешения переговоров о кодовой таблице и языке
descr	!INIT !TAR_INIT	Дополнительное описание target – комбинация значений server: port: base: neg: descr
num_targets	!TAR_INIT	Количество target, участвующих в запросе
targets	!TAR_INIT	Список target, разделенных ”;”
allhits	!TAR_INIT	Общее количество записей, найденных в результате исполнения запроса
tar_error_num	!TAR_INIT	Номер диагностического сообщения
tar_error_text	!TAR_INIT	Текст диагностического сообщения

Переменные-шаблоны предназначены для вывода информации. Они могут содержать ссылки на простые переменные. Например, содержимое переменной *record_begin* всегда выводится перед каждой записью. Если указать

```
<!-- !SET record_begin <hr>Сервер: $$server$,
      база данных: $$base$ -->
```

то перед каждой записью будет выводиться горизонтальная черта, имя сервера и базы данных. Следует обратить внимание, что имя переменной указывается как *\$\$имя\$*.

Список переменных-шаблонов приведен в таблице 6.12.

Таблица 6.12: Переменные-шаблоны шлюза ZGW

Переменная	Команды	Комментарий
records_begin	!SHOWREC	Обрабатывается один раз перед началом вывода записей
records_end	!SHOWREC	Обрабатывается один раз после вывода всех записей
record_begin	!SHOWREC	Обрабатывается перед каждой выводимой записью
record_end	!SHOWREC	Обрабатывается после каждой выведенной записью
targets_find_begin	!TAR_INIT	Обрабатывается один раз перед началом поиска по списку targets
targets_find_end	!TAR_INIT	Обрабатывается один раз после поиска по списку targets
targets_find_block	!TAR_INIT	Обрабатывается после поиска по каждой target
targets_error	!TAR_INIT	Обрабатывается после поиска по каждой target в случае наличия диагностики

На основе приведенной информации методом проб и ошибок администратору предоставляется возможность самостоятельно разрабатывать пользовательские интерфейсы шлюза *Z-GW* с применением встроенных возможностей сервера *ZooPARK*. При этом подспорьем будут являться элементы тестового шлюза, входящего в комплект поставки.

6.6 Построение распределенных систем

Сервер *ZooPARK* предназначен для построения распределенных информационных систем. Основные принципы построения можно проиллюстрировать на следующем примере.

6.6.1 Пример простой распределенной системы

Предположим, установлено три сервера *ZooPARK*: *Z1* (*z1.domain1.net*), *Z2* (*z2.domain2.net*) и *Z3* (*z3.domain3.net*). Каждый сервер имеет по одной базе данных: *db1*, *db2* и *db3* соответственно, доступ к которым осуществляется, например, через провайдер данных *Z-ISIS*. В конфигурационном файле каждого сервера будут присутствовать записи типа:

```
# Для сервера Z1
db1.type:      4
db1.codepage:  0
db1.path:      ../db1.z39
db1.file:      ../bin/dprv_isis.so
```


Используя провайдер данных *Z-REMOTE*, можно обеспечить доступ к базам данных других серверов. В частности, для сервера *Z1* можно указать:

```
# Для сервера Z1

db1.type:      4
db1.codepage:  0
db1.path:      ./db1.z39
db1.file:      ../bin/dprv_isis.so

db2.type:      4
db2.codepage:  0
db2.path:      z2.domain2.net:210:db2
db2.file:      ../bin/dprv_remote.so

db3.type:      4
db3.codepage:  0
db3.path:      z3.domain3.net:210:db3
db3.file:      ../bin/dprv_remote.so
```

При этом все запросы к базам данных *db2* и *db3* сервер *Z1* будет перенаправлять к серверам *Z2* и *Z3* соответственно. Эта процедура для клиента, установившего сеанс связи с сервером *Z1*, будет скрыта, клиент будет работать с базами данных *db2* и *db3*, так как если бы они были расположены на сервере *Z1*. Аналогичным образом можно настроить серверы *Z2* и *Z3*.

Если базы данных *db1*, *db2* и *db3* однотипны, можно объединить их в группу, присвоив ей имя, например, *db*. Для этого для каждого сервера нужно внести в конфигурационный файл следующие строки:

```
db.type:      5
db.path:      db1;db2;db3;
```

Здесь использован фиктивный провайдер с номером 5 для объединения баз данных. Теперь, работая с базой данных *db* и любым из трех серверов, клиент фактически будет осуществлять поиск и извлечение записей из трех баз данных, расположенных на разных серверах. Эта процедура для него останется скрытой для операций *search*, *present* и *scan*.

Наконец, следует заметить, что к построенной таким образом информационной системе могут быть "добавлены" базы данных любого сервера *Z39.50*, например, Библиотеки Конгресса США.

Приведенный пример распределенной информационной системы в достаточной степени иллюстрирует принцип ее построения и может быть обобщен на другие конфигурации.

6.6.2 Оптимизация распределенных систем

Рассмотренный выше пример распределенной системы основан на перенаправлении запросов к базе *db* сервера 1 на серверы 2 и 3, т.е. обработка запроса сервером 1 сводится к независимому обращению к *собственной* базе *db1* и к *чужим* базам *db2* и *db3*. Процесс последовательного обращения

к разным базам данных может занять длительное время, особенно, если баз данных много и они расположены на различных серверах.

Поиск в этой ситуации может быть существенно ускорен при распараллеливании процесса обращения к разным базам данных. При этом задача упрощается с учетом того факта, что ядро сервера *Zoo/PARK* может получить доступ к различным базам данных только через единый для всех провайдеров данных интерфейс. Архитектура сервера допускает два режима вызова провайдеров данных: последовательный вызов каждого провайдера данных с ожиданием завершения его работы и инициализация для каждого провайдера данных отдельного потока с ожиданием асинхронного окончания исполнения всех потоков. Последний вариант соответствует практически одновременному доступу ко всем базам данных. При этом суммарное время обработки запроса существенно снижается не только для многопроцессорных систем, но и для однопроцессорных, т.к. время обработки запроса, как правило, определяется временем отклика внешних устройств и удаленных серверов.

Для активизации механизма порождения отдельного потока для каждого провайдера данных при поиске для сервера *Zoo/PARK* необходимо изменить некоторые параметры в конфигурационных файлах:

```
# file zserver.cfg
# Глобальное разрешение(1)/запрещение(0) порождения потоков
server.threads: 1

# Разрешение(1)/запрещение(0) потоков для Zebra
threads: 0

# file providers.ini

# Разрешение(1)/запрещение(0) потоков для соответствующего
# динамического провайдера данных, например, isis
[isis]
thread=1
```

Дальнейшая оптимизация работы сервера *Zoo/PARK* в составе распределенной информационной системы связана с выбором режима функционирования провайдера данных удаленного сервера (*Z-REMOTE*). В соответствии с изложенным выше (см. раздел 3.4), сервер *Zoo/PARK* может быть сконфигурирован для работы в режиме 2. Для этого необходимо наличие в конфигурационных файлах следующих параметров:

```
# File providers.ini (UNIX)
. . .
[remote]
file=./dprv_remote.so
thread=1
load=1
list=remote_list

[remote\_list]
1=mx.uiggm.nsc.ru:210
2=z3950.uiggm.nsc.ru:210
3=Camelot.spsl.nsc.ru:2100
. . .
```

где параметр *thread* разрешает(1) / запрещает(0) порождение отдельных потоков как при многократных обращениях к провайдеру данных, так и при инициализации вторичных сеансов. Параметр *load* разрешает(1) / запрещает(0) открытие вторичных сеансов, т.е. устанавливает режимы работы провайдера 1 или 2. Параметр *list* указывает на секцию, которая содержит список серверов, с которыми в режиме 2 должен быть открыт сеанс связи.

Петли переадресации возникают при построении следующей конфигурации для сервера 1:

```
# Для сервера Z1

db1.type:      4
db1.codepage:  0
db1.path:      z2.domain2.net:210:db2
db1.file:      ../bin/dprv_remote.so
db1.index:     2
```

и для сервера 2:

```
# Для сервера Z2

db2.type:      4
db2.codepage:  0
db2.path:      z1.domain1.net:210:db1
db2.file:      ../bin/dprv_remote.so
db2.index:     2
```

При обращении к *db1* на сервере *Z1* или *db2* на сервере *Z2* мы получим петлю переадресации с попыткой открытия бесконечного количества сеансов между серверами *Z1* и *Z2*. Интересно, что специальная организация такой петли позволяет определить запас ресурса системы для работы в многопользовательском (многосеансном) режиме.

Для защиты от инициализационных петель и петель переадресации в сервере *ZooPARK* предусмотрены ограничения, задаваемые в конфигурационном файле провайдеров:

```
# File providers.ini (UNIX)
. . .
[remote]
file=../dprv_remote.so
thread=1
load=1
loop=2
max=3
. . .
```

где параметр *loop* указывает на максимальное количество петель при переадресации запросов, а параметр *max* – на максимальное количество переадресаций.

6.6.3 Некоторые рекомендации

Ниже приведен список рекомендаций, которым следует руководствоваться при построении распределенных информационных систем на основе серверов *ZooPARK*.

- Распределенная система должна содержать несколько серверов *ZooPARK*.
- Все серверы должны быть объединены надежными линиями связи. Требуемая ширина полосы пропускания этих линий зависит от величины сторонней загрузки, но сама по себе слабо сказывается на эффективности информационной системы.
- Выбор аппаратной платформы для функционирования серверов *ZooPARK* непринципиален, но предпочтение следует отдавать многопроцессорным системам с быстрыми дисковыми устройствами.
- В качестве программной платформы можно использовать *Windows NT/2000* или *UNIX (Linux, Solaris, FreeBSD)*. Временные характеристики систем на *Windows* обычно лучше, но надежность на *UNIX* выше. Из всех перечисленных *UNIX* наиболее эффективной является *Solaris*, наименее – *Linux*.
- Базы данных в распределенной системе следует группировать по серверам в соответствии с их аппаратной конфигурацией – на более мощном сервере должно находиться большее количество баз данных.
- В распределенной системе должны существовать 1-2 сервера, способные переадресовывать запросы. На остальных серверах провайдер данных *Z-REMOTE* лучше всего заблокировать.
- Выделенные для переадресации серверы должны функционировать в режиме 2 или 3 (см. выше). Список удаленных серверов для этого режима должен включать все серверы распределенной системы, кроме тех, которые работают в режиме 2-3 (точки входа).
- Конфигурация серверов, которые являются точками входа в систему, должна содержать в качестве имен баз данных групповые имена баз данных. Группировать базы данных следует так, чтобы в одну группу попадали базы данных с разных серверов.
- Для пользователей следует открывать доступ только к групповым именам баз данных.
- На каждом сервере должна существовать база данных *IR-Explain-1*, в которой должны быть описаны все базы данных этого сервера, в том числе и фиктивные групповые.

Глава 7

Эксплуатация *ZooPARK* в РИС

В этом разделе будут описаны различные информационные системы, в которых используется сервер *ZooPARK*. Список рассмотренных ниже информационных систем и организаций не исчерпывает все инсталляции сервера *ZooPARK*, но содержит основные. При этом наряду с системами, в которых функционируют серверы *ZooPARK*, рассматриваются системы с серверами *Z-IRBIS*. Сервер *Z-IRBIS* является упрощенной версией сервера *ZooPARK*, ориентированной на работу с библиографическими данными.

7.1 Информационная система ОИГГМ СО РАН

Созданный в ОИГГМ СО РАН сервер *ZooPARK* активно здесь эксплуатируется, являясь программной основой распределенной информационной системы организации.

Основу информационной системы ОИГГМ СО РАН составляет информационно-библиотечная система научно-технической информации (НТИ), которая является сегодня одной из наиболее развитых информационно-библиографических систем Сибирского отделения РАН. Наряду с базами данных внутренней генерации – это труды сотрудников ОИГГМ, электронные каталоги периодики и книг, пользователям предоставляется доступ к отечественным и зарубежным базам данных: машиночитаемые издания ВИНТИ, Current Contents и др. Этому предшествовал многолетний опыт проведения работ по автоматизации доступа к информационным массивам НТИ. Работы, начатые в начале восьмидесятых годов на больших ЭВМ серии ЕС, сегодня успешно продолжаются на компьютерном парке персональных ЭВМ [115–118].

Информационно-библиотечный центр ОИГГМ стал своего рода испытательным полигоном в плане освоения и внедрения новых информационных технологий. В СО РАН успешно используется накопленный опыт, во

многих Институтах СО РАН эксплуатируется программное обеспечение, разработанное в ОИГГМ.

В начале 90-х годов, в связи с переходом от машин серии ЕС на персональные компьютеры, в ОИГГМ СО РАН были успешно проведены работы по переносу накопленной ранее машиночитаемой информации НТИ на ПЭВМ и в качестве программного обеспечения для ее обработки была выбрана некоммерческая СУБД *CDS/ISIS*, разработанная в ЮНЕСКО и предназначенная для обработки информационно-библиотечных массивов.

Следует отметить, что на основе этой системы были разработаны различные автоматизированные информационные и библиотечные системы. Для многих учреждений эта система стала основным программным средством ведения баз данных научно-технической информации, электронных каталогов их библиотек, а также для решения других технологических задач. Собственно факт использования этой системы в ОИГГМ и ГПНТБ СО РАН и определил то, что основные информационные массивы НТИ в СО РАН находятся в файловой структуре *CDS/ISIS*. В настоящее время система эффективно эксплуатируется в ОИГГМ и ряде других организаций СО РАН, используется при подготовке региональных указателей литературы, ведения баз данных отечественной и иностранной периодики, для ведения баз данных трудов научных сотрудников, ведения электронных каталогов библиотек.

С появлением локальной вычислительной сети (ЛВС) в ОИГГМ возникла потребность перехода к новым технологиям предоставления пользователям имеющихся информационных ресурсов. В ОИГГМ одним из первых в Сибирском отделении был реализован "файл-серверный" вариант доступа к НТИ, который фактически ничем не отличался от локальных систем, за исключением возможности обработки больших информационных массивов, расположенных на файловых серверах и возможностью одновременного доступа нескольких пользователей к этим ресурсам. Это был первый наш опыт взаимодействия с базами данных *CDS/ISIS* в сетевом окружении.

Несмотря на то, что применяемая технология была достаточно примитивна, такая схема работы оказалась достаточно живучей в информационно-библиотечной практике и в различных вариантах функционирует и по сей день, поскольку практически все персональные СУБД, в том числе и *CDS/ISIS* могут работать в архитектуре "файл-сервер". Эта технология, отлаженная в ОИГГМ была успешно перенесена и адаптирована для работы в локальной вычислительной сети СО РАН, являющейся в настоящее время одной из самых развитых сетей Академии наук России [119–121], которая насчитывает свыше 2500 станций, управляемых мощными файловыми серверами Институтов СО РАН в средах *UNIX*, *Windows NT* и др. Базы данных и электронные каталоги библиотек, электронные коллекции ряда организаций СО РАН насчитывают свыше 40 млн. записей. Основными держателями ресурсов являются ГПНТБ СО РАН, Информационно-библиотечный Центр ОИГГМ СО РАН, Информационный центр Института катализа СО РАН и ряд других (см. также раздел 7.2).

Несмотря на то, что разработанная система на протяжении длительного времени успешно функционировала и функционирует, в такой архитектуре существуют ряд недостатков, которые резко снижают ее эффективность:

- большой сетевой трафик при монтировании сетевых дисков;
- неэффективность использования ресурсов сервера;
- примитивная обработка транзакций;
- нагрузка на компьютер клиента.

Большинство из перечисленных выше недостатков могут быть устранены при переходе к "клиент-серверной" архитектуре. Исходя из этого, в ОИГГМ была предпринята успешная попытка создания собственного переносимого сервера, позволяющего работать с базами данных *CDS/ISIS*, способного функционировать на различных платформах в средах *OC UNIX*, *Windows*.

В целях обеспечения коллективного использования информационных ресурсов был проведен анализ существующих информационно-поисковых систем для обработки больших информационных массивов – *SPIRS* (*Silver Platter Information Retrieval System*), *ISI CDE*, *CC on Disk* и ряда других. На основе анализа было разработано необходимое программное обеспечение для доступа к информационным массивам, физического расположения которого могло находиться как на стороне клиента, так и на стороне сервера.

При этом предполагалось согласование входных и выходных потоков данных с потоками данных и протоколами серверов *HTTP*, что позволило бы увязать весь комплекс программ с *WEB*-сервисом сети Интернет, а в качестве графического интерфейса конечного пользователя использовать любой *WEB*-браузер, реализованный в средах *OC UNIX* и *Windows* [106–108, 122]. Эта работа явилась предпосылкой для разработки ряда других решений.

Схема взаимодействия при этом выглядит следующим образом: клиент (пользователь) посылает запрос серверу, который им (сервером) исполняется, после чего результат запроса передается обратно клиенту.

Клиент в данном случае – это рабочая станция пользователя с установленным *WEB*-браузером. Сервер – станция со значительными вычислительными и дисковыми ресурсами (для хранения баз данных) с установленным *WEB*-сервером (для взаимодействия с *WEB*-браузером клиента), который через *CGI*-приложение взаимодействует с базами данных *CDS/ISIS*.

Пользователь при этом может использовать любую ОС и *WEB*-браузер. При этом предполагалось, что интерфейс работы с базами данных будет везде иметь один и тот же вид. Этим достигается унификация пользовательского интерфейса системы. Набор элементов экранной формы при этом должен быть жестко типизирован и прост. Это позволяет упростить требования к подготовленности пользователя в плане работы с базами данных. Ему достаточно было уметь работать с *WEB*-браузером и изучить справочный материал системы о правилах составления поискового запроса к базам данных.

Ключевым моментом при создании такой системы доступа к библиографическим ресурсам НТИ был выбор *CGI*-приложения, которое бы обеспечивало полную функциональность и высокое быстродействие. Под функциональностью в данном случае понимались:

- возможность взаимодействия с *WEB*-сервером для обмена информацией при обработке запросов клиента к базам данных и выдаче результатов поиска;
- наиболее полный набор функций для обработки данных баз данных *CDS/ISIS* (поиск, работа со словарем, вывод результатов поиска, ввод, обновление и др.).

Для этих целей было разработано *CGI*-приложение, которое удовлетворяло этой функциональности. Однако, реализованное приложение имело ряд существенных недостатков, которые можно было снять лишь с реализацией полнофункционального *CGI*-приложения. Такая программа – *WWWISIS* была разработана в *BIREME/PAHO/WHO* (Бразилия) [123]. Версия этой программы существует для наиболее распространенных аппаратных платформ. Программа *WWWISIS* взаимодействует с *WEB*-сервером через *CGI*-интерфейс, который в настоящее время поддерживается всеми существующими *WEB*-серверами. Использование *Win32 API* в версии *WWWISIS* для *MS Windows NT* позволило существенно повысить быстродействие одновременной (многопользовательской) обработки запросов к базам данных *CDS/ISIS* (многопоточность при выполнении запроса), по сравнению с быстродействием традиционной СУБД *CDS/ISIS*.

В процессе работ над интерфейсом будущей *WEB*-ориентированной информационно-поисковой системы был детально изучен существующий опыт систем подобного рода и выдвинут 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 различных кодовых таблиц);
- корректная работа системы с доминирующими *WEB*-браузерами;
- выбор оптимального разрешения экрана;
- использование стандартных унифицированных средств *HTML* для отображения запрашиваемой информации;
- в целях экономии сетевого трафика, полностью текстовый интерфейс экранных форм для минимизации времени доступа к информации;
- обработка полноты и корректности данных экранных форм на стороне клиента (возможностями используемого *Web*-браузера) без участия сервера (уменьшения количества транзакций "клиент-сервер");
- эргономичность интерфейса;

- 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, таких как:
 - выбор поисковых терминов из словаря БД;
 - 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 службы администрирования баз данных для осуществления обратной связи с пользователями (внесения замечаний и предложений по дальнейшему развитию системы).

Исходя из вышеперечисленных требований, сотрудниками ОИГГМ СО РАН и ГПНТБ СО РАН была разработана *WEB*-ориентированная информационно-поисковая система [124], которая в настоящий момент находится в промышленной эксплуатации в ряде организаций Сибирского Отделения РАН.

Несмотря на привлекательность построенного *WEB*-решения для работы с однородными базами данных следует отметить и недостатки этого решения. В первую очередь – это то, что такое решение не позволяет в полной мере построить распределенную информационную систему для доступа к разнородным ресурсам в силу того, что специфика протокола *HTTP* такова, что в нем невозможно передать структурированную информацию без привлечения дополнительных механизмов передачи данных. Во-вторых, при таком решении невозможно хранить контекст сессии, т.е. пользователь не может вернуться к результатам ранее выполненного запроса, с целью построения более сложного или отложить просмотр результатов поиска, что является очень существенным в распределенных информационных системах. В-третьих, при таком подходе для перенаправления запроса на другой сервер, требуется еще один промежуточный сервер, который бы занимался диспетчеризацией запросов. И наконец, возможности *JavaScript* ограничены, что существенно ограничивает сложность построения графического интерфейса.

Решение вышеперечисленных проблем возможно лишь в унификации доступа к серверам любых производителей и базам данных, которая может быть реализована при использовании протокола *Z39.50*. В отличие от других протоколов, обеспечивающих поиск и извлечение информации в глобальных информационных системах, *Z39.50* отличается от них тем, что способы формирования поисковых запросов и форма представления результатов поиска не зависят от источника информации. Протокол *Z39.50*

позволяет абстрагироваться от способа организации конкретной базы данных, сделать доступ к ее содержанию независимым от формы хранения ее данных. Использование протокола *Z39.50* позволяет избавить пользователя от необходимости адаптации к различным источникам информации, что в свою очередь позволяет сократить непроизводительное время на обучение пользователя работе с различными базами данных и тем самым повысить их производительность труда.

Обеспечение доступа по протоколу *Z39.50* к существующим базам данных, потребовало разработки специализированного программного обеспечения, интегрирующего эти технологии [125–129]. Результатом выполнения этих работ явилось создание комплекта серверного программного обеспечения под общим названием *ZooPARK* (см. раздел 5.1), включающего в себя в одной из первых версий лишь модуль работы по протоколу *Z39.50* с базами данных *CDS/ISIS*. Основными задачами этого провайдера были: отображение запросов на обработку данных для *Z39.50* в логику запросов СУБД *CDS/ISIS* и отображение ответов из баз данных *CDS/ISIS* в логику ответов для *Z39.50*. В более поздних версиях была существенно расширена функциональность и сервера и провайдера данных *CDS/ISIS*, созданы провайдеры данных для других СУБД (см. раздел 5.4). Сегодня программное обеспечение *ZooPARK* является основой корпоративной распределенной информационно-библиотечной системы г. Новосибирска (см. раздел 7.4.1) и других (см. раздел 7.4).

Библиографические ресурсы не исчерпывают ресурсы РИС ОИГГМ СО РАН, доступные по протоколу *Z39.50*. Так на сегодняшний день РИС ОИГГМ СО РАН включает информацию обо всех сотрудниках института. Предоставляемая информация достаточно многогранна и сгруппирована по разделам (см. Рис.7.1):

- Общая информация
- Место работы
- Образование
- Ученые звания
- Ученые степени
- Награды
- Гранты
- Научные интересы
- Участие в конференциях
- Преподавательская деятельность
- Публикации

При этом часть информации (Общая информация, Место работы, Награды, Ученые степени, Ученые звания) берется в момент обработки запроса из баз данных Отдела кадров (*MS SQL Server*), что делает ее всегда актуальной, другая часть (Гранты, Научные интересы, Участие в конференциях, Преподавательская деятельность) – из таблиц, доступных для модернизации каждому пользователю, а часть (Публикации) представлена в виде ссылок на другие базы данных, например, на библиографическую базу данных "Труды сотрудников ОИГГМ СО РАН". Такая организация видимой для пользователей базы данных "Сотрудники ОИГГМ СО РАН" позволяет

- оперировать только с актуальной информацией
- скрыть истинный источник и истинную структуру данных
- разграничить зоны ответственности за данные
- предоставлять доступ через общие для всех ресурсов *Z39.50* интерфейсы
- интегрировать внутреннюю РИС с другими информационными системами

В РИС ОИГГМ СО РАН существуют и другие информационные ресурсы.

Таким образом, информационная система ОИГГМ СО РАН является в настоящее время своеобразным конгломератом современных информационных технологий, мирно сосуществующих и направленных на наиболее полное удовлетворение информационных потребностей ученых и специалистов в области наук о Земле и позволяющих интегрировать базы данных ОИГГМ с другими базами данных в гетерогенных информационных системах как СО РАН, так и России в целом.

7.2 Информационная система СО РАН

Важнейшей частью Информационной среды Сибирского отделения РАН является информационная поддержка научных исследований, проводимых в Отделении, а также создание и развитие собственных информационных ресурсов, управление этими ресурсами и обеспечение использования информационных ресурсов мирового научного сообщества, представляемых сетью Internet, распространение своих достижений в виде электронных коллекций, атласов и информационных систем, а также в виде электронных публикаций и электронных библиографических ресурсов.

В Отделении накоплена и постоянно собирается уникальная научная информация как по различным отраслям наук, так и по природному комплексу. Но, к сожалению, пока отсутствует единая технология сбора информации, разобщенность ее как по месту хранения, так и по форме представления и что самое главное отсутствует информация об информации. Все это не позволяют широко использовать информационные ресурсы для

интенсификации и кооперации междисциплинарных фундаментальных и прикладных исследований. Существенной проблемой, связанной с информационными ресурсами Отделения, является факт "гибели" информации, вызванный постоянным оттоком кадров исследователей и технического персонала, ответственного за ее хранение.

Для решения проблем информационного обеспечения в Отделении принято решение о создании собственной информационной системы, в которой бы аккумулировалась большая часть необходимой для сотрудников информации, включая создание полнофункциональной системы об интеллектуальном потенциале Отделения и "Электронной библиотеки Сибирского отделения РАН", содержащей информационную часть (информационную систему об институтах, сотрудниках, достижениях и др. аспектах, связанных с работой Отделения), систему электронной поддержки сбора и накопления информации (системы электронных коллекций, баз данных и т.п.) и электронную поддержку документооборота в Отделении. В результате в Сибирском отделении РАН была сформирована целевая программа Отделения создания и развития информационных ресурсов под общим названием "Электронная библиотека Сибирского отделения РАН" [130, 131]. Данная программа, как направление деятельности, была одобрена на заседании Президиума СО РАН 25.06.1998.

Основные направления программы связаны с формированием собственных электронных ресурсов по основным отраслям наук (математика, науки о земле, химия, биология, археология и др.), созданию и поддержке электронных коллекций и электронных публикаций, организации удобных систем доступа к библиотечным и библиографическим базам данных ГПНТБ СО РАН и базам данных Институтов Отделения, организации зеркал наиболее значимых мировых информационных ресурсов. Создаваемая информационная система Сибирского отделения РАН призвана обеспечить:

- Единую распределенную информационную среду Отделения.
- Информационную поддержку исследований по фундаментальным и прикладным направлениям, проводимым в институтах Отделения, а также межинститутских междисциплинарных научных исследований.
- Поддержку профессионально-ориентированных систем подготовки и обмена научных документов с элементами удаленной совместной работы.
- Поддержку профессионально-ориентированных систем доступа и интерфейсов с банками данных и автоматизированными библиотеками.
- Поддержку перспективных систем телекооперации исследователей на базе современных технологий.
- Коллективное использование приобретаемой электронной литературы, каталогов, баз данных и библиографических изданий.

- Поддержку электронных версий научных журналов, издаваемых институтами Отделения. Издание собственных электронных журналов, книг, препринтов и дайджестов по различным направлениям научных исследований.
- Поддержку принятия и реализации организационных и управленческих решений в Отделении. Создание среды, обеспечивающей электронный документооборот Отделения.

Если говорить о типах информационных ресурсов, которые должны поддерживаться ИРИС, то можно выделить следующие

- Каталоги информационных ресурсов, т.е. информация об информационных ресурсах, или метаинформация. К этому типу следует отнести всю вторичную информацию об информационных ресурсах, в том числе электронные каталоги традиционных библиотек, электронные версии реферативных журналов, каталоги музеев, архивов и т.п. К этому типу также следует относить описания ресурсов *WEB*-серверов, базы данных по организациям, сотрудникам и т.д.
- Электронные коллекции, т.е. совокупности цифровых объектов, объединенных по каким-либо общим признакам. Электронные коллекции представляют собой первичный информационный ресурс. Несомненно, описания этих коллекций и отдельных цифровых объектов являются ресурсами вторичными, т.е. каталогами.
- Классификационные ресурсы, т.е. различные электронные рубрикаторы, тезаурусы, схемы и справочники, нормирующие правила составления каталогов информационных ресурсов и отношения между цифровыми объектами.
- Хранилища программного обеспечения, которое обеспечивает функционирование ИРИС в целом, ее отдельных компонент и информационных ресурсов.
- Административные информационные ресурсы, включающие полную информацию об актуальном состоянии ИРИС и ее отдельных компонент. Несомненно, каждый из перечисленных типов информационных ресурсов требует собственного подхода при его интеграции в рамках ИРИС и собственных интерфейсов для доступа пользователей.

Работы по программе "Электронная библиотека Сибирского отделения РАН" и более поздним программам "Информационные ресурсы Сибирского отделения РАН" и "Информационно-телекоммуникационные ресурсы СО РАН" привели к созданию прототипа [8, 129, 132] ИРИС на основе протокола *Z39.50* с доступом через шлюз *HTTP-Z39.50*

http://z3950.uiggm.nsc.ru:210/zgw/sbras_ris.htm

с возможностью сквозного поиска в распределенных базах данных (см. Рис. 7.1). Этот прототип включает в себя следующие подсистемы

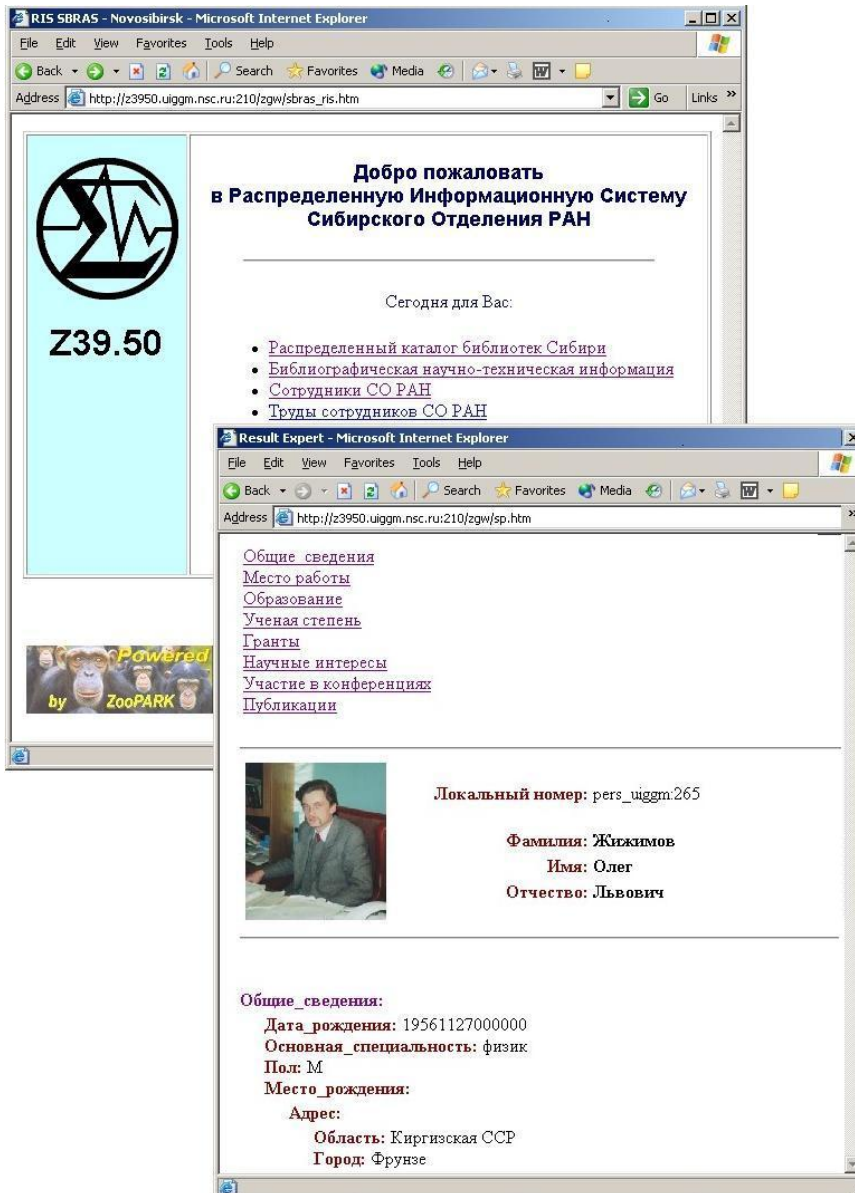


Рис. 7.1: Точка входа в РИС СО РАН

- Информационные ресурсы СО РАН
- Распределенный электронный каталог библиотек г. Новосибирска
- Библиографическая научно-техническая информация
- Труды сотрудников СО РАН
- Сотрудники СО РАН (см. Рис. 7.1)
- Электронные коллекции СО РАН
- Геоинформационные ресурсы СО РАН
- Тезаурусы и классификационные схемы
- Информация о серверах *Z39.50* (Explain)

На сегодняшний день не все подсистемы этого прототипа функционируют в требуемом режиме, некоторые находятся в стадии отладки.

При построении инфраструктуры РИС на основе протокола *Z39.50* используется программное обеспечение *ZooPARK*, включающие в себя все необходимые компоненты для предоставления доступа к информационным ресурсам (сервер *Z39.50*, шлюз *Z39.50-HTTP*). Программное обеспечение *ZooPARK* позволяет построить распределенную информационную систему, содержащую множество серверов *Z39.50* с единой точкой входа, сквозным поиском и возможностью извлекать и просматривать информацию в различных форматах (*SUTRS*, *XML*, *HTML*, *GRS-1*, *RUSmarc* и др.). В качестве схем данных используются принятые в мировой практике стандартные схемы (*GILS* [47], *Collection* [63], *CIMI* [66], *GEO* [60], *ZTHES* [74], *EXPLAIN*, *RUSmarc*, *USmarc*) и локально определенные схемы данных (*UIGGM*, *PERSONS*).

В 2003 году проект создания распределенной информационной системы СО РАН получил дальнейшее развитие в рамках интеграционного проекта СО РАН "Создание виртуального музея СО РАН" [133, 134]. В этом проекте в качестве основы доступа к распределенным ресурсам (метаданным) используется протокол *Z39.50*.

7.3 Информационная система *LibWeb*

Проекты *LibWeb* (1995-2001) и *LibWeb-2* (2001-2003) были направлены на сетевую интеграцию электронных информационных ресурсов ведущих библиотек и информационных центров России на базе программно-аппаратной среды, создаваемой в рамках Федеральной межведомственной программы создания национальной сети компьютерных телекоммуникаций для науки и высшей школы (Миннауки РФ, РФФИ, Госкомвузом РФ и РАН)¹.

Для реализации указанной цели предполагалось (см. [135–140]):

¹Проект *LibWeb* финансировался РФФИ с 1995 года (проекты № 95-07-19417, № 98-07-90394, № 01-07-90445) и с 1997 года РГНФ (проект № 96-05-12025).

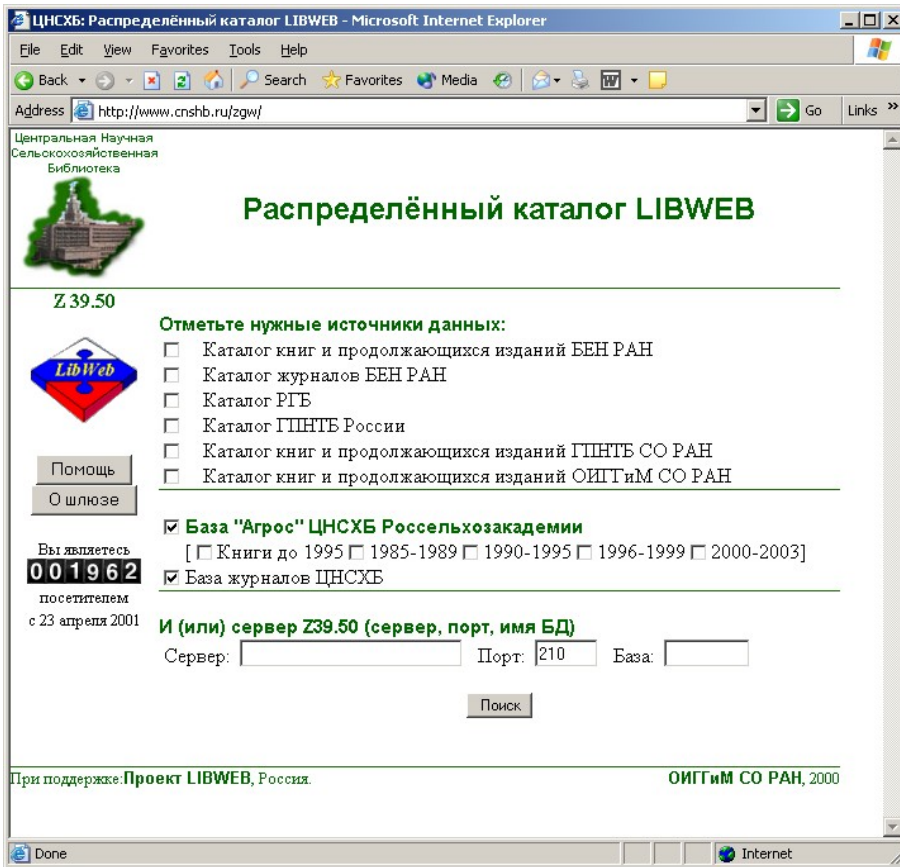


Рис. 7.2: Точка входа в распределенную информационную систему *LibWeb* (ЦНСХБ, шлюз *Z39.50-HTTP*)

1. Сетевая интеграция существующих информационных ресурсов ведущих библиотек и информационных центров России путем широкого использования архитектуры клиент-сервер, современных реляционных баз данных (РСУБД), апробации и выбора типовых РСУБД функционирующих в *UNIX*-среде, использования стандартов языка запросов *SQL* для организации взаимодействия *WWW*-серверов и серверов БД.
2. Широкое использование имеющихся и создание новых телекоммуникационных ресурсов отечественных некоммерческих сетей передачи данных, входящих в Internet. Проект *LibWeb* не предусматривал финансирование подключения библиотек и информационных центров к Internet, закупку коммуникационного оборудования и т.п., оставляя решение этих задач на другие проекты или за счет собственных средств участников.

3. Модернизация технологии ведения и использования электронных каталогов библиотек-участниц проекта и автоматизированной системы Российского Сводного Каталога по научно-технической литературе на базе принципов поддержки информационных продуктов, принятых в Internet.
4. Обеспечение доступа пользователей к электронным каталогам, библиографическим массивам и другим информационным ресурсам *Lib-Web* с использованием стандартного клиентского ПО, принятого в WWW (*Mosaic*, *Netscape*, *MS Internet Explorer* и др.); организация одновременного поиска в нескольких электронных каталогах путем диспетчеризации обращений к сети серверов баз данных через промежуточные интерфейсы (*Web-DBMS Interface*) соответствующих WWW-серверов.
5. Организация ретроспективной конверсии в библиотеках-участницах проекта с помощью современных сетевых и других автоматизированных технологий.

В результате осуществления проекта была создана распределенная библиотечная сеть, обеспечивающая доступ и обмен библиографической, технологической и другой информацией. Подавляющее большинство информационных ресурсов *LibWeb* предоставляется учреждениям и работникам науки, культуры и образования бесплатно.

В рамках *LibWeb* функционирует распределенная информационно-поисковая библиографическая система, обеспечивающая:

- непосредственный доступ работникам науки, культуры и образования к информационным ресурсам крупнейших библиотек России;
- кооперированную каталогизацию для всех библиотек-участниц проекта, что позволяет резко сократить трудоемкость и время обработки изданий, повысить оперативность получения новейшей информации пользователем;
- организацию межбиблиотечного российского и международного книгообмена и кооперированного заказа.

Участниками и разработчиками проекта LibWeb были:

- Библиотека Российской Академии Наук (БАН)
- Библиотека по естественным наукам РАН (БЕН РАН)
- Вычислительный центр РАН (ВЦ РАН)
- Государственная публичная историческая библиотека
- Государственная публичная научно-техническая библиотека России (ГПНТБ России)

- Государственная публичная научно-техническая библиотека СО РАН (ГПНТБ СО РАН)
- Институт научной информации по общественным наукам РАН (ИНИОН РАН)
- Институт прикладной математики им. М.В. Келдыша РАН
- Институт развития информационного общества (ИРИО)
- Научная библиотека Казанского государственного университета
- Научная библиотека Московского государственного университета
- Научная библиотека Санкт-Петербургского государственного университета
- Научная библиотека Южно-Уральского государственного университета
- Объединенный институт геологии, геофизики и минералогии СО РАН (ОИГГМ СО РАН)
- Российская государственная библиотека (РГБ)
- Российский государственный университет нефти и газа им. И.М. Губкина
- Центр по проблемам информатизации сферы культуры Минкультуры РФ (Центр ПИК МК РФ)
- Центральная научная сельскохозяйственная библиотека РАСХН (ЦНСХБ РАСХН)

В ходе выполнения проекта была осуществлена пилотная стадия создания распределенной сети *LibWeb* на основе поддержки протокола *Z39.50* (см. [140]).

Созданы новые электронные информационные ресурсы, содержащие полнотекстовую информацию о наиболее значимых для библиотечного сообщества материалах по созданию и использованию распределенных корпоративных сетей.

В ходе выполнения проекта разработаны типовые решения для представления библиографических информационных ресурсов в Интернет, которые позволили представить в публичный доступ научному и образовательному сообществу России информацию о нескольких миллионах документов из фондов ведущих библиотек и информационных центров:

- типовое решение ГПНТБ России (представление библиографической информации в Интернет)
- типовое решение ВЦ РАН (представление библиографической информации в Интернет)

- типовое решение ОИГГМ СО РАН (сервер *ZooPARK* для работы с библиографическими данными, метаданными, тезаурусами и классификационными схемами по протоколу *Z39.50*)
- кириллица (стандарт представления кириллической информации в электронной форме)

В соответствии с принципами проекта *LibWeb*, все типовые решения есплатно передавались всем участникам проекта.

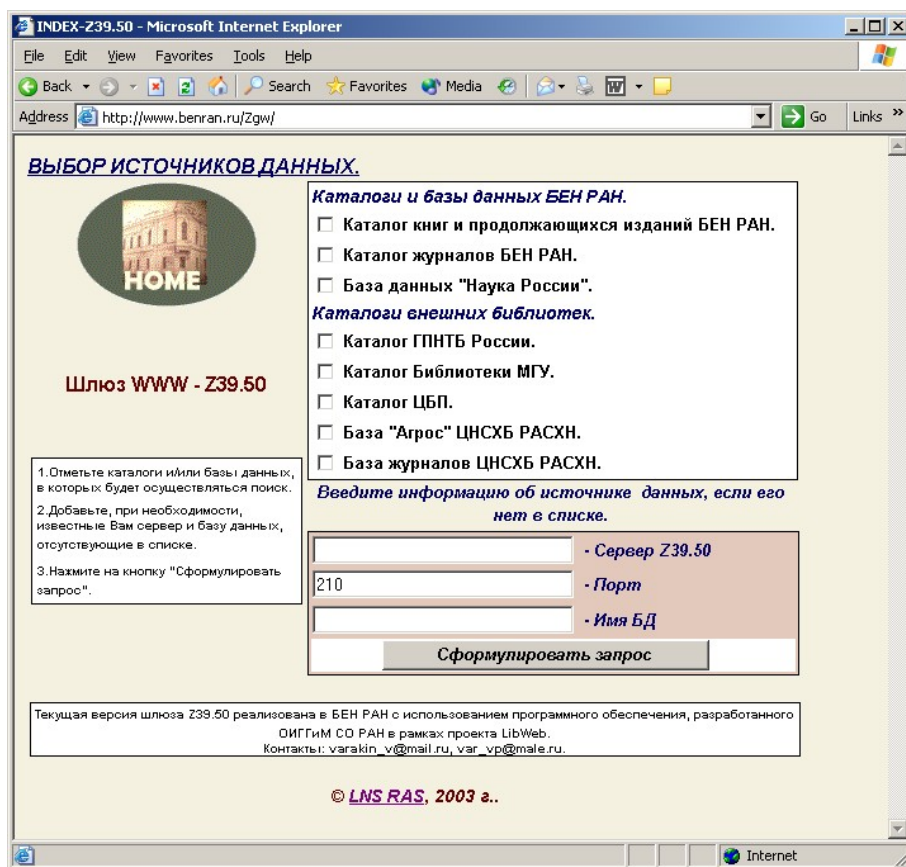


Рис. 7.3: Точка входа в распределенную информационную систему LibWeb (БЕН РАН, шлюз *Z39.50-HTTP*)

7.4 Информационные системы корпоративных проектов ИОО

В 1999 году Институт "Открытое общество" (Фонд Сороса) объявил программу создания в России региональных библиотечных консорциумов для

интеграции информационных ресурсов, в первую очередь электронных каталогов, библиотек. Основой интеграции был провозглашен протокол *Z39.50* и коммуникативный формат *RUSmarc*. В течение трех лет ИОО финансировал создание в России региональных библиотечных информационных систем на основе открытых международных стандартов. Это привело к созданию более десятка информационных систем, некоторые из которых являются распределенными и функционируют на основе протокола *Z39.50*.

Программное обеспечение, которое было использовано при создании корпоративных библиотечных информационных систем, создавалось разными группами разработчиков. Однако многие консорциумы остановили свой выбор на сервере *ZooPARK*.

7.4.1 Корпоративная библиотечная РИС Новосибирска

Целью создания региональной распределенной корпоративной библиотечной системы явилось осуществление работ по объединению имеющихся информационных ресурсов библиотек и обеспечения свободного доступа к ним как пользователей Сибирского региона, так и пользователей Российской Федерации. Система создавалась на основе интенсивного использования технологий Интернет и открытых международных и российских стандартов (см. [141–144]).

Основные задачи проекта:

- обеспечение широкого и свободного доступа читателей, в том числе с физическими ограничениями, к библиотечно-информационным ресурсам региона по полному репертуару изданий, поступающих в фонды библиотек-участниц, как к единому информационному ресурсу;
- реализация технологии корпоративной каталогизации;
- обеспечение удаленного заказа на получение первоисточника (оригинал, ксерокопия, электронная копия и др.) через Интернет;
- обеспечение электронной доставки документов;
- ретроспективная конверсия каталогов библиотек;
- приведение описаний документов в соответствие с российскими и международными стандартами для обеспечения обмена;
- внедрение системы организационного, технического и технологического взаимодействия библиотек-участниц в режиме реального времени, на основе разработанной системы.

Для решения этих задач, являющихся составной частью информационного развития региона, объединились библиотеки, имеющие свои сети библиотек: Государственная публичная научно-техническая библиотека Сибирского отделения Российской академии наук (ГПНТБ СО РАН), Ново-

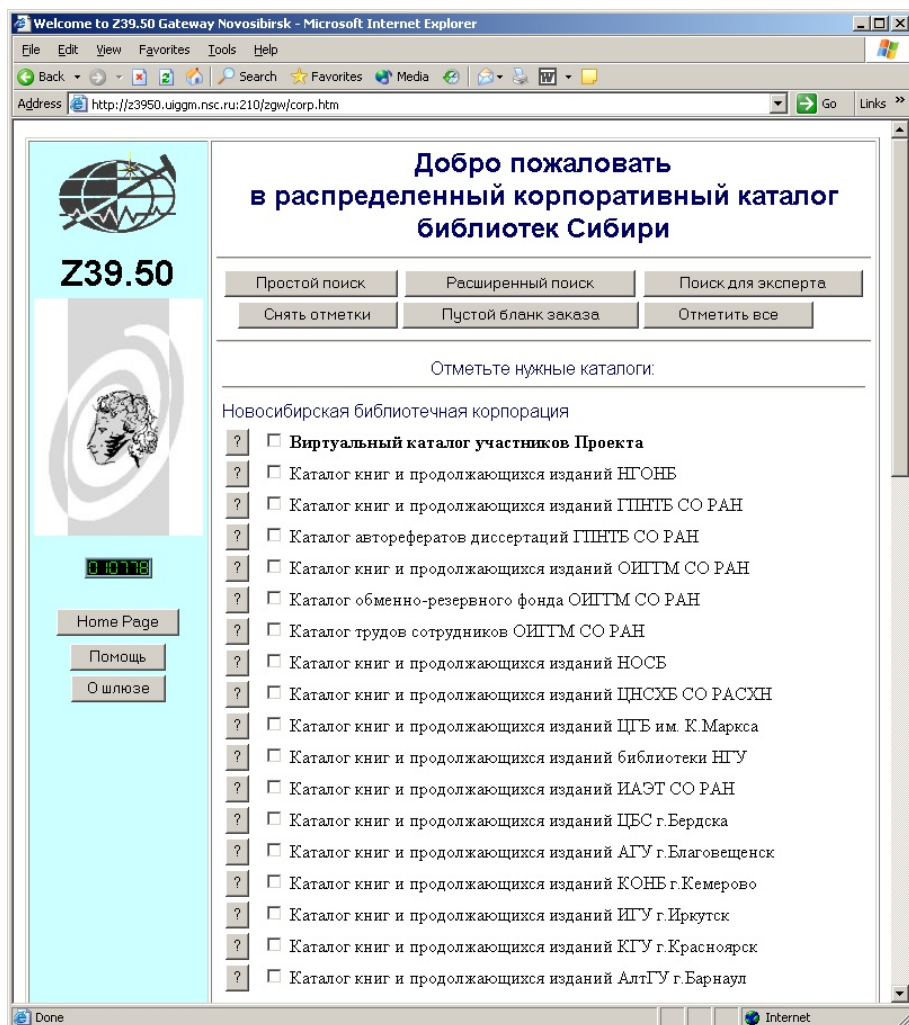


Рис. 7.4: Точка входа в Новосибирскую распределенную библиотечно-информационную систему (шлюз *Z39.50-HTTP*)

Адрес сервера	Организация	Тип
z3950.rstlib.nsc.ru:210	НГОНБ Новосибирск	<i>ZooPARK</i>
camelot.spsl.nsc.ru:2100	ГПНТБ СО РАН Новосибирск	<i>ZooPARK</i>
z3950.uiggm.nsc.ru:210	ОИГГМ СО РАН Новосибирск	<i>ZooPARK</i>
z3950-1.uiggm.nsc.ru:210	ОИГГМ СО РАН Новосибирск	<i>ZooPARK</i>
z3950.nsu.ru:210	НГУ Новосибирск	<i>ZooPARK</i>
www.sati.archaeology.nsc.ru:210	ИАЭТ СО РАН Новосибирск	<i>ZooPARK</i>
libserv.amursu.ru:210	Библиотека АГУ Благовещенск	<i>ZooPARK</i>
z3950.library.isu.ru:210	ИГУ Иркутск	<i>ZooPARK</i>
z3950.asu.ru:210	АлтГУ Барнаул	<i>ZooPARK</i>
62.76.120.166:210	КГУ Красноярск	<i>Liber</i>
213.59.170.24:210	КГТУ Красноярск	<i>ZooPARK</i>
62.76.2.62:210	КОНБ Кемерово	<i>Liber</i>

Таблица 7.1: Серверы *Z39.50* РИБС г. Новосибирска

сибирская государственная областная научная библиотека (НГОНБ), Областная специальная библиотека для незрячих и слабовидящих (НОСБ), библиотека и Интернет Центр Новосибирского государственного университета (НГУ), Центральная городская библиотека им. К. Маркса (ЦГБ). Участником проекта также являлся и Информационно-библиотечный центр Объединенного института геологии, геофизики и минералогии СО РАН (ИБЦ ОИГГМ СО РАН).

В результате выполнения проекта кроме всего прочего была создана распределенная библиотечная информационная система, основанная на технологиях *Z39.50* (см. [145–149]). К концу 2003 года эта система объединила ресурсы ведущих библиотек г. Новосибирска и других городов Сибирского региона на основе серверов *Z39.50*, перечисленных в Табл.7.1. Доминирующими серверами *Z39.50* являются серверы *ZooPARK*.

7.4.2 Корпоративная библиотечная РИС Москва

Проект "Корпоративная сеть московских библиотек" в течение трех лет финансировался Институтом "Открытое Общество" в рамках программы "Корпоративные библиотечные системы" (см. [150–158]). Обязательными требованиями для участников этого проекта были

- Полная совместимость с Российским коммуникативным форматом *RUSmarc* и международными форматами *UNIMARC* и *USmarc* на основе использования средств двухстороннего конвертирования данных;

- Поддержка WWW-технологии, определяемая наличием средств, обеспечивающих доступ к электронным каталогам и другим информационным ресурсам библиотек через WEB-сервер;
- Поддержка протокола Z39-50 (ISO-23950) средствами, реализующими доступ к электронным каталогам по этому протоколу;
- Открытость для развития собственных библиотечных технологий их пользователями.

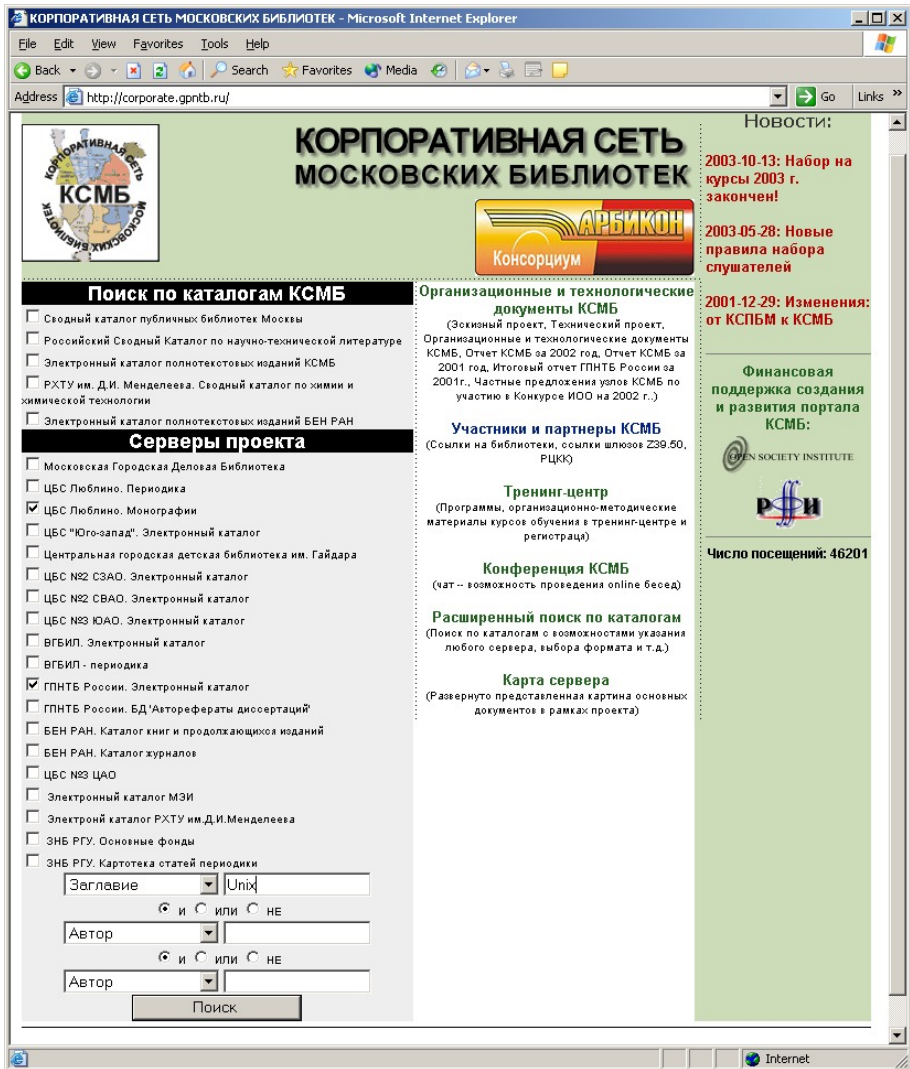


Рис. 7.5: Точка входа в Московскую распределенную библиотечно-информационную систему (шлюз Z39.50-HTTP)

Адрес сервера	Организация	Тип
213.171.40.126:210	Московская Городская Деловая Библиотека	Z-RBIS
www.infokniga.ru:210	ЦБС Люблино	ZooPARK
www.cbs-uz.ru:210	ЦБС "Юго-запад".	ZooPARK
195.178.220.50:210	Центральная городская детская библиотека им. Гайдара	
195.209.131.2:210	ЦБС №2 СЗАО	Z-IRBIS
195.209.135.50:210	ЦБС №2 СВАО	Z-IRBIS
212.34.55.72:210	ЦБС №3 ЮАО	ZooPARK
zserv.libfl.ru:210	ВГБИЛ	Z-IRBIS
193.233.14.7:210	ГПНТБ России	Z-IRBIS
80.250.169.6:210	ЦБС "Гагаринская"	Z-IRBIS
www.benran.ru:210	БЕН РАН	ZooPARK
212.45.9.7:210	ЦБС №3 ЦАО	
lib.muctr.edu.ru:210	РХТУ им.Д.И.Менделеева	Z-IRBIS
www.znb.rsu.ru:210	ЗНБ РГУ	ZooPARK

Таблица 7.2: Серверы ZooPARK и Z-IRBIS в РИС г. Москвы

7.4.3 Корпоративная библиотечная РИС Ярославль

Проект "Ярославская корпоративная библиотечная сеть" [159] финансировался в рамках программы ИОО "Российские корпоративные библиотечные системы".

1. Ярославский государственный университет им. П.Г. Демидова (<http://www.uniyar.ac.ru>)
2. Ярославская областная универсальная научная библиотека им. Н.А. Некрасова (<http://www.rlib.yar.ru>)
3. Ярославский государственный технический университет (<http://library.ystu.yar.ru>)
4. Ярославская государственная медицинская академия (<http://gw.yma.ac.ru>)
5. Ярославский государственный педагогический университет им. К. Д. Ушинского (<http://www.yspu.yar.ru>)
6. Центральная научно-техническая библиотека ЦНТИ (<http://www.csti.yar.ru>)
7. Ярославский государственный историко-архитектурный и художественный музей-заповедник (<http://www.yarmp.yar.ru>)
8. Централизованная библиотечная система города Ярославля (<http://www.clib.yar.ru>)

9. Ярославская областная детская библиотека им. И.А. Крылова
10. Ивановская областная научная библиотека
(<http://www.ionb.interline.ru>)
11. Костромской государственный университет им. Н.А. Некрасова
12. Международный университет бизнеса и новых технологий
(<http://www.mubint.ru>)
13. Ярославская государственная сельскохозяйственная академия
14. Ярославская областная научная медицинская библиотека
15. Ивановский государственный университет
(<http://www.ivanovo.ac.ru>)

Характерной особенностью созданной информационной системы (см. [160–163]) является, с одной стороны, распределенность на уровне баз данных, но, с другой стороны, централизованность доступа по протоколу *Z39.50*. На сегодняшний день существует только один сервер *Z39.50* (*ZooPARK*, версия 4.05, *lib.yar.ru:210*), обеспечивающий сервис *Z39.50* по доступу к базам данных всех участников Проекта. Доступ к ресурсам возможен также через шлюз (<http://lib.yar.ru/z3950.html>), пользовательские интерфейсы которого приведены на Рис. 7.6.

7.4.4 Корпоративная библиотечная РИС Нижний Новгород

Характерной особенностью корпоративной распределенной информационной библиотечной системы (РИБС) г. Нижний Новгород является разнородность серверного программного обеспечения *Z39.50* и используемых АИБС и СУБД (см. [164, 165]). Участники нижегородской корпорации используют АИБС *Liber-Media* (ЗАО *Liber*), *МАРК-SQL* (Информсистема), *Библиотека-4* (ГИВЦ МК РФ) и др. Для предоставления доступа к ресурсам по протоколу *Z39.50* используются как серверы *ZooPARK*, так и серверы *Z39.50* ЗАО *Liber*. Точкой входа в РИБС служит шлюз *Z39.50-HTTP* ЗАО *Liber*, функционирующий на *WEB*-сервере Нижегородской областной универсальной научной библиотеки.

Используемый в корпорации шлюз позволяет производить одновременный поиск по протоколу *Z39.50* в информационных массивах участников в соответствии с *Vib-1* и рекомендациями программы "Российские корпоративные библиотечные системы". Извлечение информации возможно в форматах *RUSmarc*, *USMARC* и *SUTRS*. Графические *WEB*-интерфейсы пользователя приведены на Рис.7.7 (см. также [166]).

7.4.5 Корпоративная библиотечная РИС Омск

Участники проекта создания корпоративной библиотечной РИС сформулировали свои цели так (см. [167–169]):

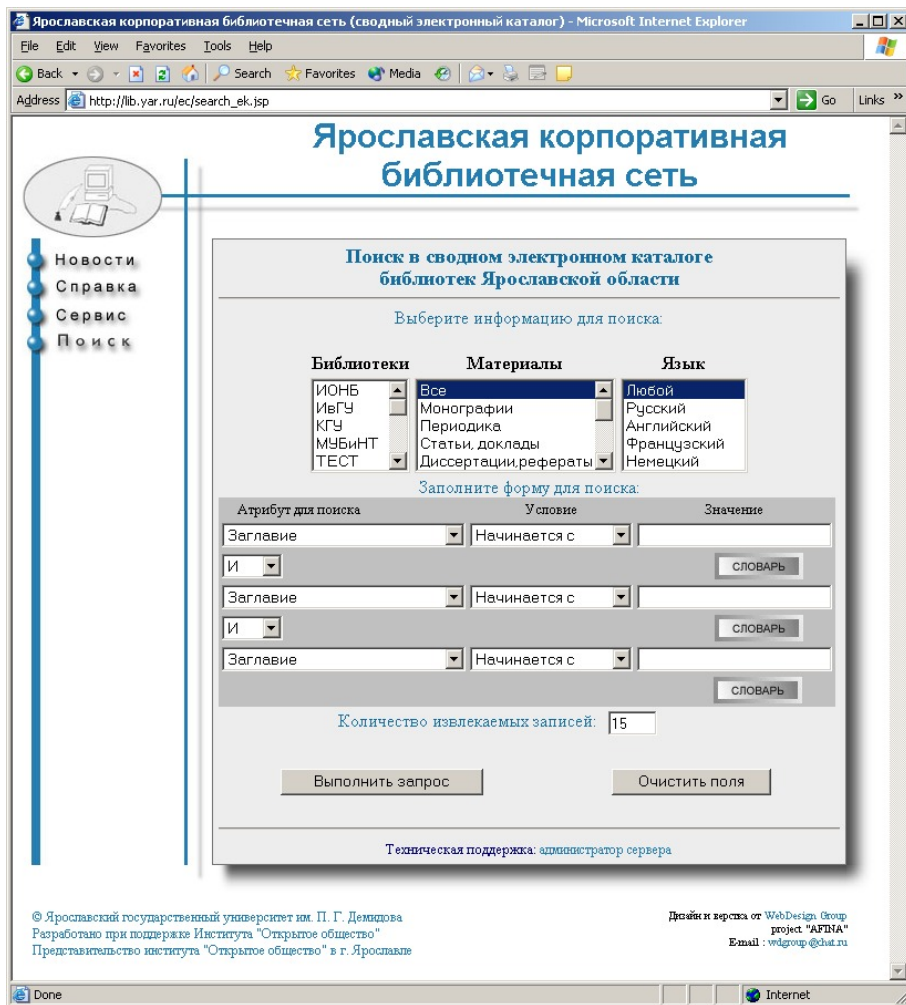


Рис. 7.6: Точка входа в Ярославскую корпоративную библиотечно-информационную систему (шлюз Z39.50-HTTP)

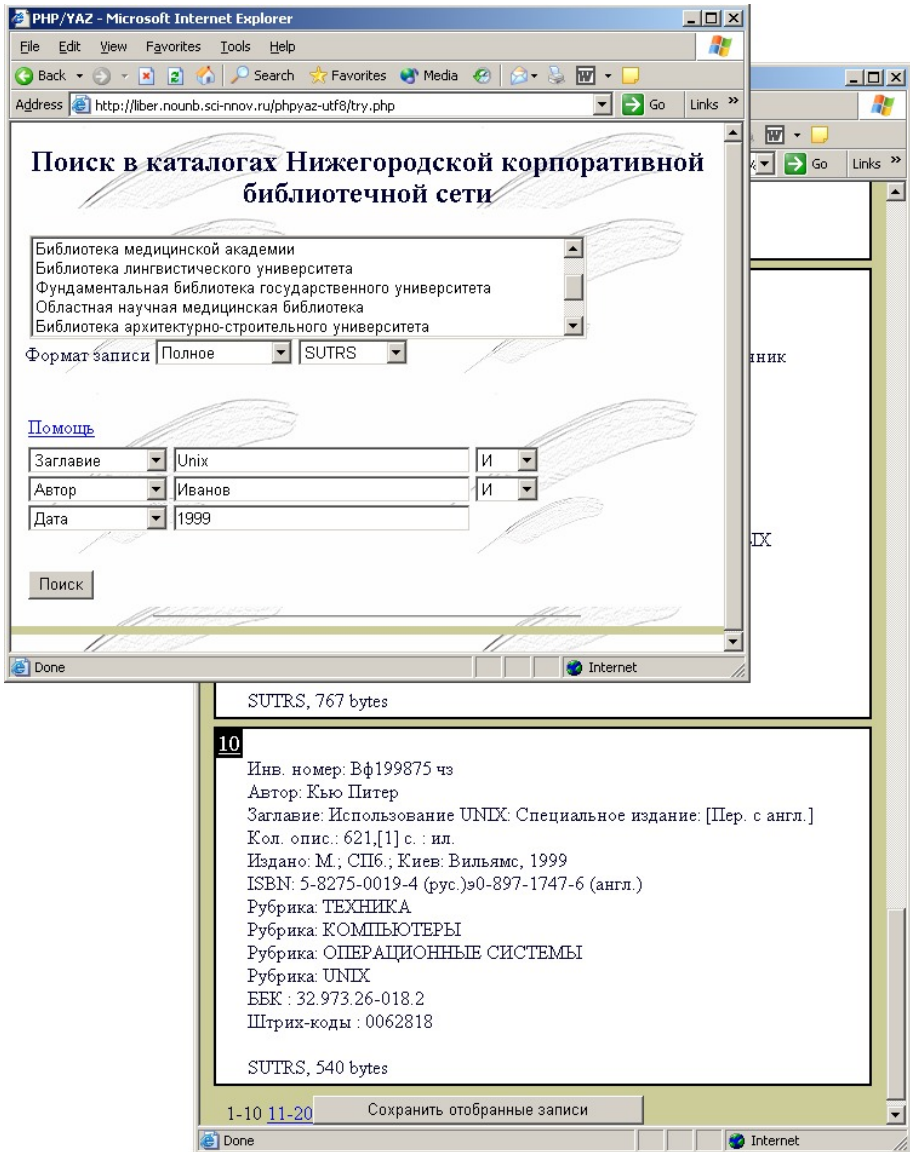


Рис. 7.7: Точка входа в Нижегородскую Корпоративную библиотечную сеть (плюс Z39.50-HTTP)

Адрес сервера	Организация	Тип
81.18.132.170:210	Нижегородская областная научная медицинская библиотека	<i>ZooPARK</i>
library.n-nov.mednet.com:2100	Нижегородская государственная медицинская академия	<i>D3ZServer</i>
liber.nounb.sci-nnov.ru:2100	Нижегородская областная универсальная научная библиотека	<i>D3ZServer</i>
z3950.nounb.sci-nnov.ru:210	Нижегородская областная универсальная научная библиотека	<i>ZooPARK</i>
lunnlib.lunn.sci-nnov.ru:210	Нижегородский лингвистический университет	<i>ZooPARK</i>
server-ibm.lib.unn.runnet.ru:210	Нижегородский государственный университет	<i>ZooPARK</i>
194.190.191.62:210	Нижегородский архитектурно-строительный университет	<i>ZooPARK</i>
liber.rsuh.ru:210	Российский государственный гуманитарный университет г. Москва	<i>D3ZServer</i>
www.bgunb.ru:5002	Белгородская государственная универсальная научная библиотека	<i>D3ZServer</i>

Таблица 7.3: Серверы *Z39.50* РИБС г. Нижний Новгород

- объединение информационных ресурсов библиотек и создание корпоративного информационного пространства библиотечных услуг Омска и Омской области;
- создание организационно-технических предпосылок для включения корпоративной системы в национальную информационно-библиотечную систему;
- обеспечение подключения основных библиотек Омска и Омской области к корпоративной библиотечной системе;
- поддержка современных стандартов организации и оформления потока информационного обмена между участниками корпоративной библиотечной системы;
- создание организационно-технических предпосылок для внедрения систем автоматизации библиотечного процесса, обеспечивающих "прозрачную" интеграцию библиотеки в корпоративную информационную среду;

- создание первоначальных узлов доступа к корпоративной библиотечной системе на базе основных образовательных библиотек города.

В результате проведенных работ была создана РИС, включающая серверы Z39.50 (см. Табл.7.4), реализующая модель распределенного каталога (см. раздел 4.6.3). Отличительной особенностью Омского корпоративного библиотечного консорциума является использование им только серверов ZooPARK и Z-IRBIS.

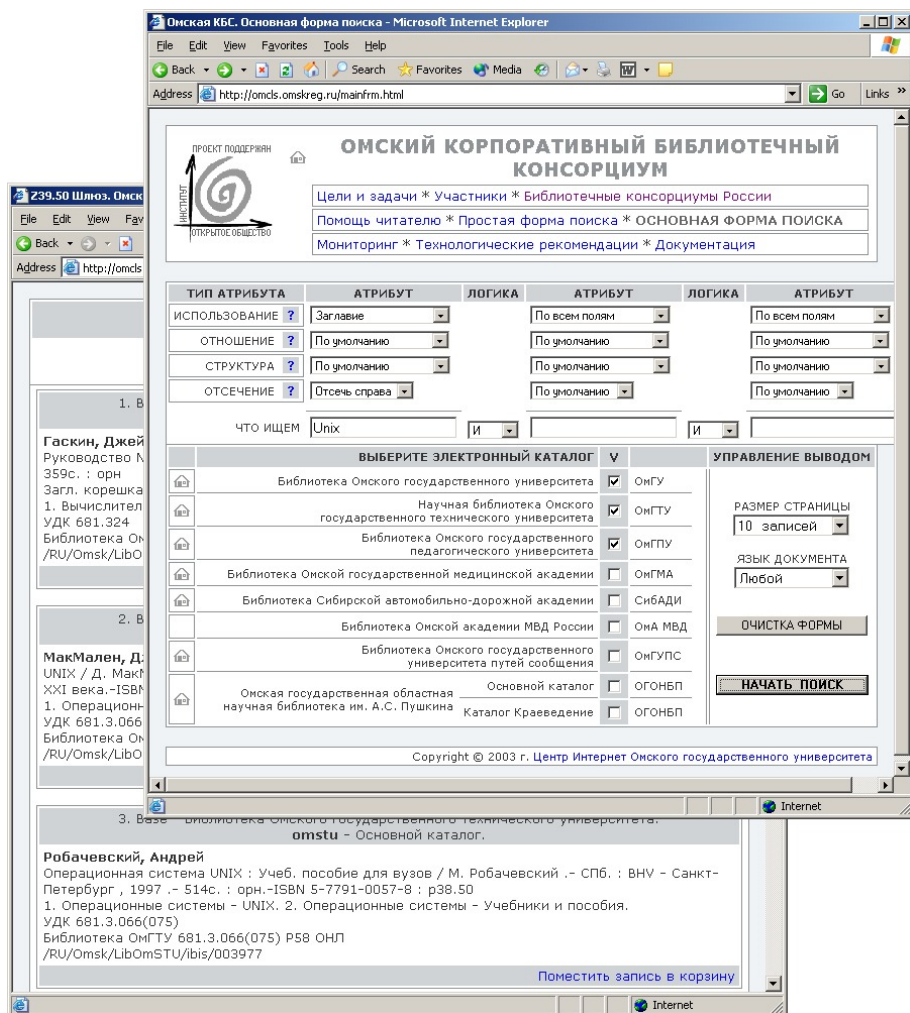


Рис. 7.8: Точка входа в Омскую Корпоративную библиотечную информационную систему (шлюз Z39.50-HTTP)

Адрес сервера	Организация	Тип
bibl.omamvd.ru:210	Омская академия МВД	Z-IRBIS
weblib.omsk-osma.ru:210	Омская государственная медицинская академия	Z-IRBIS
kbs.lib.omskreg.ru:210	Омская государственная областная научная библиотека им. А.С. Пушкина	ZooPARK
212.192.39.6:210	Омский государственный педагогический университет	Z-IRBIS
lib.omgtu.ru:210	Омский государственный технический университет	Z-IRBIS
omcls.omskreg.ru:210	Омский государственный университет	Z-IRBIS
lib.omgups.ru:210	Омский государственный университет путей сообщения	ZooPARK
195.162.33.166:210	Сибирская автомобильно-дорожная академия	Z-IRBIS

Таблица 7.4: Серверы Z39.50 РИБС г. Омска

7.5 Информационная система по газгидратам

7.5.1 Особенности проекта

Цель описываемого проекта¹ – создание специализированной информационной системы об уникальных объектах, которые являются предметом исследований ученых разных направлений, о газгидратах (см. [170–172]). Эта информационная система должна интегрировать разнородные базы данных со структурированной информацией о газгидратах и различные неструктурированные данные в виде текстов, схем, фотографий и т.п. Отбор информации, ее систематизация и аналитическая переработка составляли часть проекта, но за эту часть отвечали специалисты соответствующих областей знаний. Здесь обсуждается лишь та часть проекта, которая касается собственно информационной системы, технологии ее построения и интерфейсов доступа к ней.

На языке специалистов газгидраты представляют собой нечто среднее между твердым веществом и сильно сжатым и замороженным газом. Они реально существуют в природе. Уже открыто несколько месторождений. Ученые геологи находят газгидраты в совершенно разных местах планеты, например, в мировом океане на большой глубине. Некоторые газгидраты обладают интересными физико-химическими свойствами. Например, будучи извлеченными с большой глубины, они буквально тают в руках исследователей, испаряются, за несколько минут превращаясь в газ. Кроме

¹Проект *GASHYDAT (MAS3-CT98-0176)* финансировался Европейской Комиссией в рамках III программы *Marine Science and Technology*. Официальный сайт Проекта <http://www.gashydat.org>.

чисто научного интереса газгидраты привлекают и своими экономическими перспективами. Существуют оценки специалистов, что газгидраты могут служить основным сырьем для химической промышленности 21 века и быть эффективным источником энергии подобно нефти и природному газу. Несмотря на большой интерес к столь уникальным объектам, информация о них скудна, раздроблена по разным источникам и не систематизирована.

Настоящим проектом предполагалось улучшить ситуацию в области систематизации информации о газгидратах.

7.5.2 Структура информационной системы

Для реализации проекта, т.е. для построения распределенной информационной системы по газгидратам, были сформулированы требования, которым должна удовлетворять эта система, а именно:

- допускать хранение данных различных типов: тексты, карты, фотографии изображения, библиографические сведения о научных работах, ссылки на другие ресурсы;
- допускать хранение метаданных, т.е. данных об этих данных;
- допускать пополнение и изменение данных и метаданных;
- поддерживать целостность базы метаданных и их соответствие реальным данным;
- обеспечивать поиск по различным критериям и их комбинации в пределах информации описанной в базе метаданных;
- поддерживать разнообразное представление информации включая некоторые специальные формы визуализации, например, просмотр карт.
- быть доступной из Интернет через WWW;
- быть распределенной прозрачно для пользователя;
- удовлетворять существующим мировым и отраслевым стандартам;
- быть аппаратно независимой;
- быть легко управляемой, каждый сервер информационной системы должен допускать удаленное администрирование;
- быть достаточно живучей: отключение нескольких серверов не должно приводить к краху информационной системы.

Принимая во внимание, что удовлетворить описанным выше требованиями практически невозможно в рамках одной технологии, изначально был выбран подход, основанный на интеграции различных технологий, каждая из которых эффективна в своей части.

В качестве основы информационной системы был выбран протокол *Z39.50 (ISO-23950)*. Использование этого протокола позволяет не только остаться в рамках международных стандартов доступа к базам данных, но и достаточно эффективно построить аппаратно независимую распределенную информационную систему, как минимум на уровне метаданных, с единой системой запросов и стандартных форматов представления информации.

В отличие от уровня метаданных для уровня данных было выбрано несколько параллельных возможностей доступа: *Z39.50*, *WWW*, *FTP* и, вообще любые другие, определяемые *URL*. В качестве точки доступа к информационной системе со стороны *WWW* было решено использовать шлюз *Z39.50-HTTP*, выполняющий функции клиента *Z39.50*. Этот выбор обеспечивал достаточно простой и демократичный механизм не только для построения пользовательских интерфейсов в окружении привычного *WEB*-браузера, но и для создания существенно продвинутых решений на основе *Z39.50*.

Другой способ доступа к информационной системе – применение специализированных клиентов *Z39.50*. Сегодня в мире насчитывается несколько готовых решений, но все они ориентированы в первую очередь на доступ к библиографической информации, в то время как имеется необходимость отображения разнородной информации, в том числе и графической. Тем не менее информационная система допускает использование специализированных клиентов *Z39.50*.

Общая схема информационной системы представлена на Рис. 7.9.

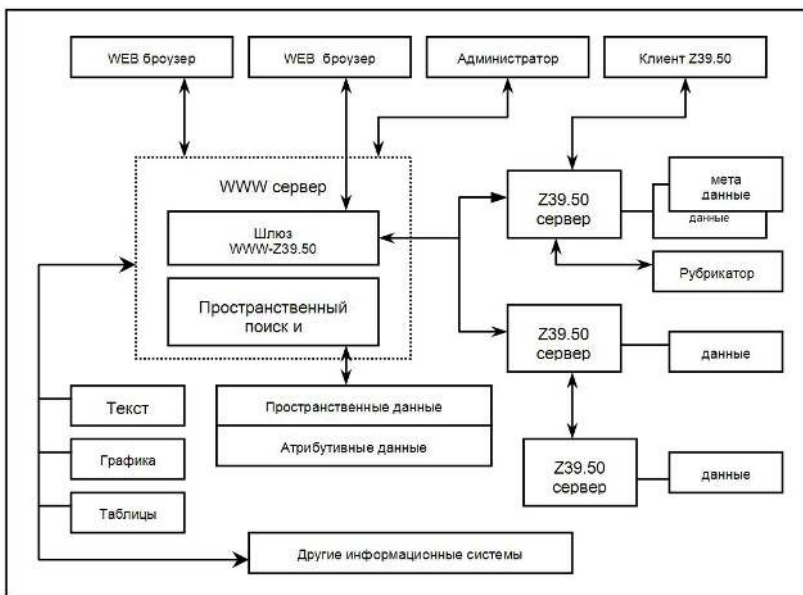


Рис. 7.9: Общая схема информационной системы по газгидратам

7.5.3 GILS как основа информационной системы

Поскольку в основу всей информационной системы было положено требование соответствия мировым стандартам, функциональность построенной системы никак не связана с конкретными способами хранения данных. На уровне метаданных, доступ к которым осуществляется по протоколу *Z39.50*, информация всегда отображается во внешнее представление в соответствии со схемой *GILS (Global Information Locator Service)* [47], поскольку *GILS* – это стандартная схема, отображение на которую данных из любой СУБД не представляет особой сложности. На первом этапе (опытной эксплуатации), для базы метаданных был выбран вариант хранения данных в виде обычных текстовых *XML*-файлов с теговой структурой *GILS*. Индексирование производилось при помощи свободно распространяемого программного обеспечения *Zebra* от компании *IndexData*, создающего совместимую по структуре базу данных с одним из провайдеров данных сервера *ZooPARK*. В дальнейшем, при увеличении объема хранимой информации, планируется задействовать реляционную СУБД с соответствующим провайдером данных для сервера *ZooPARK*. Ниже приведен пример записи базы метаданных в формате *XML*:

```
<gils>
  <controlIdentifier> mgashyd_1020 </controlIdentifier>
  <title> Gas composition and isotopic composition of
    gas hydrate from offshore Sakhalin Island,
    Okhotsk Sea </title>
  <originator><name> Ginsburg, G.D. </name></originator>
  <originator><name> Soloviev, V.A. </name></originator>
  <publicationDate> 1993 </publicationDate>
  <languageOfResource> eng </languageOfResource>
  <abstract> Gas hydrate recovered during R/V "Geolog Pyotr
    Antropov" cruise offshore Sakhalin Island in Okhotsk
    Sea in 1991. See table9 in url:Analysed gas hydrate
    recovered from Site 91-02-40 and Site 91-02-42.
    Parameters escribed: gas composition (methane,
    propane,... CO2, N2, He content) and isotopic
    composition (d(13)C-CH(4) PDB, dD-CH(4) SMOW,
    d(13)C-CO(2) PDB) </abstract>
  <controlledSubjectIndex>
    <subjectTermsControlled>
      <controlledTerm> 1.1.2.3 </controlledTerm>
    </subjectTermsControlled>
  </controlledSubjectIndex>
  <subjectTermsUncontrolled>
    <uncontrolledTerm> gas hydrate </uncontrolledTerm>
    <uncontrolledTerm> gas composition </uncontrolledTerm>
  </subjectTermsUncontrolled>
  <spatialDomain>
    <boundingCoordinates>
      <westBoundingCoordinate>144</westBoundingCoordinate>
      <eastBoundingCoordinate>145</eastBoundingCoordinate>
      <northBoundingCoordinate>55</northBoundingCoordinate>
      <southBoundingCoordinate>54</southBoundingCoordinate>
    </boundingCoordinates>
    <place><placeKeyword>Okhotsk Sea</placeKeyword></place>
  </spatialDomain>
  <sourcesOfData> Ginsburg G.D. and Soloviev V.A. Submarine
```

```

Gas Hydrates. VNIIOkeangeologia, St-Petersburg,
p.215, 1998.</sourcesOfData>
<originalControlIdentifier>1020</originalControlIdentifier>
<url> data/ginssolov98_t9.htm </url>
<recordSource>
<name> Poort J. </name>
<organization> RMCA (Tervuren, Belgium) </organization>
</recordSource>
<dateOfLastModification> 19991206 </dateOfLastModification>
<recordReviewDate> 19991230 </recordReviewDate>
</gils>

```

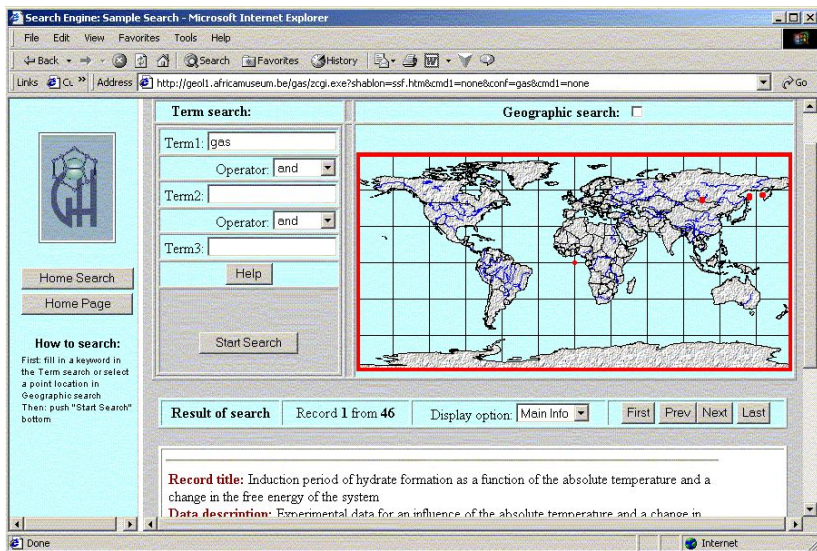


Рис. 7.10: Интерфейс пользователя при поиске

Следующий за метаданными информационный слой системы – собственно данные. Эти данные могут быть любыми: текст, графика, мультимедиа и даже отдельные базы данных с собственными интерфейсами доступа. В описываемой информационной системе доступ к этим данным осуществляется через *URL*, хранящийся в записи базы метаданных. Это позволяет задействовать механизм гипертекстовых ссылок при просмотре метаданных, предоставляя простой выход на сами данные.

На Рис.7.10 показан интерфейс пользователя при поиске информации в базе метаданных.

Следующий слой – доступ к накопленной ранее библиографической информации. Система позволяет подключать эту информацию в двух вариантах:

- осуществлять прямой поиск в любых библиографических базах данных, доступных по протоколу *Z39.50*;

- обращаться к специализированным библиографическим приложениям через URL, ссылка на который присутствует в GILS, для активизации расширенных возможностей этих приложений.

На Рис.7.11 показан интерфейс пользователя при выходе по URL в другую информационную систему.

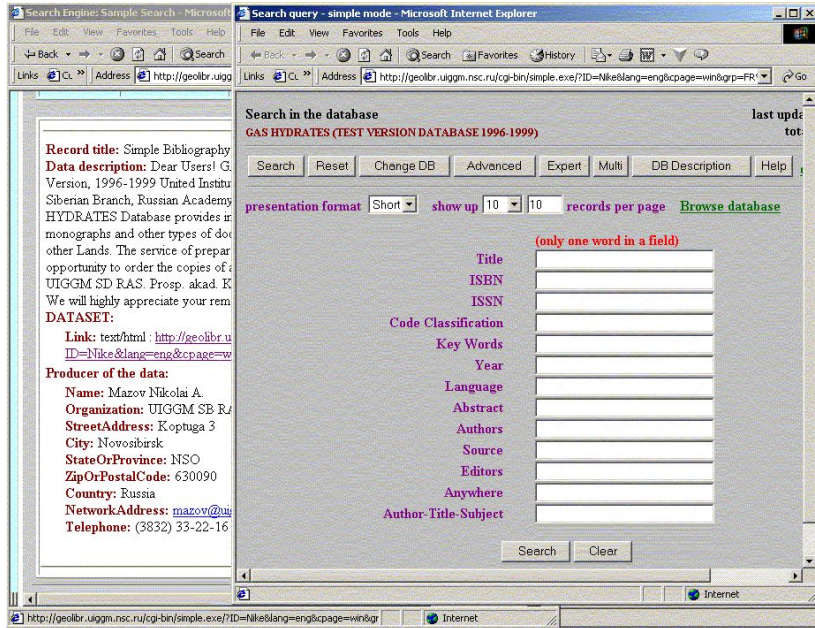


Рис. 7.11: Выход в другую библиографическую информационную систему

7.5.4 Пространственная привязка и интерфейсы пользователя

Другой информационный слой, который интегрирован с метаданными – географическая информация о положении и размерах описываемых месторождений газгидратов. Для четкой идентификации объекта и его визуализации на карте необходимо использовать данные ГИС (ГеоИнформационные Системы) 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. Специфика ГИС данных – большие объемы информации. Так, например, хранение только контура одного геологического объекта может занимать уже несколько мегабайт. Существуют специализированные системы и форматы для хранения такого рода информации. В описываемой системе был использован один из наиболее

популярных и стандартных форматов хранения – *SHP*-файлы, атрибутивная же информация была размещена в файлах форма *DBF*. *SHP/DBF* – очень распространенные форматы ГИС данных, любая ГИС система может их экспортировать и импортировать. В нашей системе доступ этой информации осуществляется *WEB*-сервером и шлюзом *Z39.50-HTTP* через специальный *CGI*-модуль.

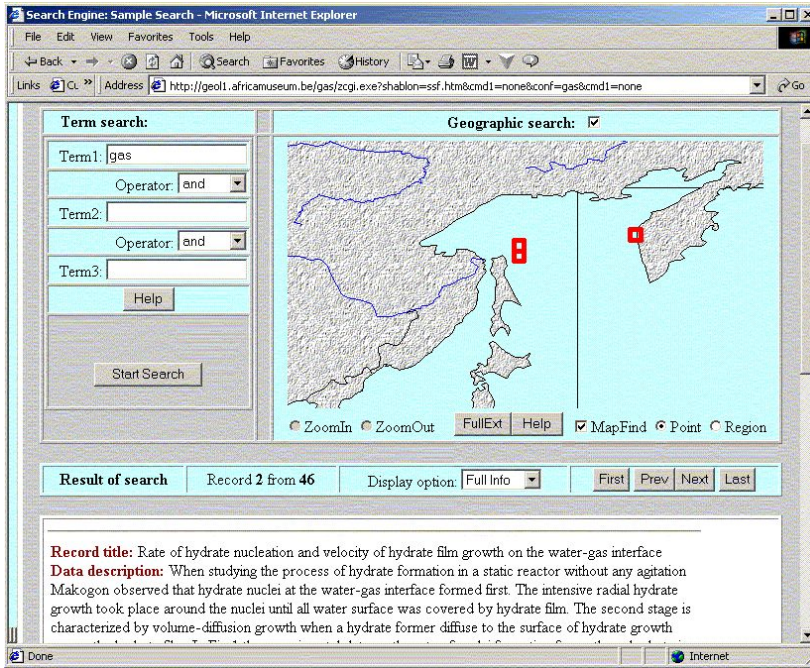


Рис. 7.12: Масштабирование и географический поиск

Интеграция поисковой подсистемы *Z39.50*, элементов *GILS* и ГИС-информации, позволила построить пользовательские интерфейсы с реализацией визуального поиска по масштабируемой карте и отображения на ней записей из баз метаданных. На Рис.7.12 показан интерфейс пользователя при масштабировании карты и географическом поиске в базе метаданных.

7.5.5 Рубрикатор

Наконец, необходимым элементом информационной системы является классификация информации. В этом случае в системе необходим доступ к тематическим рубрикаторам, возможность навигации по ним и проведение параллельного поиска. Это задача также была решена на основе протокола *Z39.50* и схемы данных *Zthes* [74], внедрение которой сейчас в мире только начинается. В схеме данных *Zthes* присутствуют все элементы, необходимые для построения рубрикаторов или тезаурусов: коды, названия, отношения, что позволяет создать классификационную базу дан-

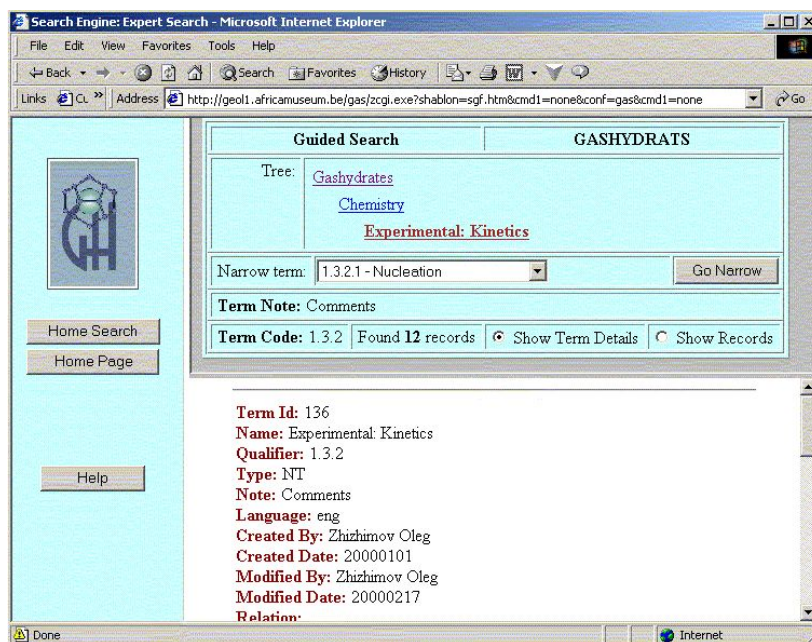


Рис. 7.13: Интерфейс навигации по рубриктору

ных, обеспечить спецификации *Z39.50* и предоставить через *Z39.50-HTTP* плюз интерфейс для перемещений по рубриктору с одновременным поиском в базах метаданных и выходом на найденные записи. Интерфейс навигации по рубриктору показан на Рис.7.13.

Глава 8

Проекты и решения

В предыдущих главах обсуждались распределенные информационные системы на основе протокола *Z39.50*. При этом основной акцент был направлен на внутреннее устройство РИС и на типы данных, которые могут быть доступны в этих РИС. В качестве примера были подробно рассмотрены РИС на основе серверов *ZooPARK*, приведены ссылки на реально существующие системы.

Однако было бы несправедливо ограничивать рассмотрение РИС на основе протокола *Z39.50* только системами *ZooPARK*, которые, во-первых, существуют только в России, и, во-вторых, даже в ней не исчерпывают все разнообразие программного обеспечения *Z39.50*.

В заключительной главе мы попытаемся сделать обзор существующих на сегодняшний день в мире проектов построения распределенных информационных систем на основе протокола *Z39.50*. Мы не претендуем на полноту этого обзора, но надеемся, что в нижеизложенном материале нашли свое отражение основные направления архитектурных изысков развития РИС.

8.1 Библиотечные проекты

Поскольку первоначально протокол *Z39.50* использовался преимущественно библиотечным сообществом, библиотечных проектов, основанных на нем, существует достаточно много. Многие ведущие научные, публичные и университетские библиотеки в развитых странах сегодня предоставляют доступ к своим электронным ресурсам по протоколу *Z39.50*.

Следует отметить, что мировое библиотечное сообщество внесло самый существенный вклад в самую идею и ее практические реализации построения информационных систем, в том числе и распределенных, на основе протокола *Z39.50*. И это представляется вполне закономерным, т.к. именно библиотечное сообщество имеет многовековые традиции работы с информационными ресурсами, причем как с первичными (книги, журналы, рукописи и т.д.), так и со вторичными (библиографические описания, карточки, каталоги, т.е. метаданные). Наличие в библиотечном сообществе устояв-

шихся жестких правил систематизации и каталогизации привело к быстрой его адаптации к новым возможностям обслуживания пользователей через глобальные сети и интеграции информационных ресурсов. Именно в библиотечном сообществе был создан протокол *Z39.50*, в котором воплотился опыт предоставления сетевого доступа к разнородным информационным ресурсам.

Естественно, факт предоставления ресурсов по протоколу *Z39.50* еще не означает построения распределенной информационной системы. Даже в минимальном требовании к РИС присутствует требование возможности проведения сквозного поиска во всех источниках прозрачным для пользователя образом. При отсутствии интеграции информационных ресурсов на уровне серверов (или на каком-либо промежуточном уровне) такой поиск можно осуществить в *Z39.50* только при интеграции ресурсов на уровне клиента *Z39.50*. Именно по последнему пути долгое время шло развитие идеи поиска в распределенных ресурсах. При этом в качестве клиентов *Z39.50* могут выступать:

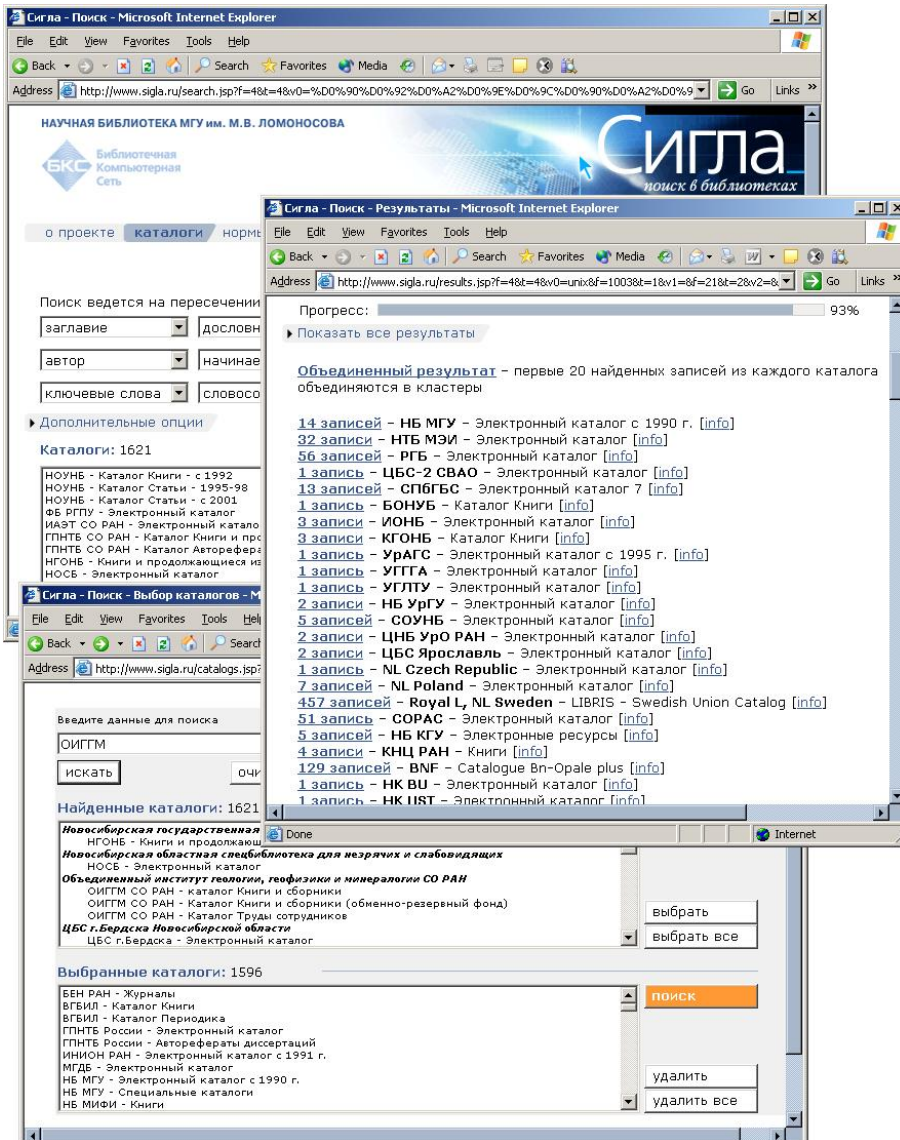
- Специализированное программное обеспечение, реализующее функции "толстого" клиента и предоставляющее пользователю полнофункциональные и громоздкие графические интерфейсы. При работе со множеством источников пользователь сам формирует список источников, с каждым из которых будет открыт сеанс связи для проведения поиска и извлечения информации. Очевидно, что такая модель предъявляет очень жесткие требования к ресурсам клиентского рабочего места и к каналам связи. Тем не менее, именно так работают многие известные клиенты *Z39.50*, например, *BookWhere (WebClarity Software Inc., [173])* или из отечественных *OLSC Client [174]* ("Открытые библиотечные системы", СПбГТУ).
- Заранее сконфигурированный шлюз *Z39.50-HTTP* как промежуточный слой между пользовательским *WEB*-браузером и ресурсами *Z39.50*. Несмотря на, как правило, ограниченную функциональность подобной модели, в ней переносится жесткость требований к рабочему месту пользователя на *WEB*-сервер, на котором исполняется программа шлюза. В простых шлюзах все сценарии работы пользователя жестко детерминированы и сводятся к поиску в выбранном списке статических источников и просмотру найденной информации. Однако наличие промежуточного слоя между серверами *Z39.50* и пользователем позволяет эмулировать для последнего примитивную распределенную информационную систему, исполняя запросы пользователя на указанном списке источников. Большинство шлюзов *Z39.50-HTTP* функционируют именно так. Наиболее известный пример – шлюз [175] Библиотеки Конгресса США, поддерживающим около 500 серверов *Z39.50* или российский шлюз *RUSLAN [176]* ("Открытые библиотечные системы", СПбГТУ). Однако на всех шлюзах существует ограничение на размер списка выбранных источников. Чаще всего, как в приведенных шлюзах, этот список ограничивается базами данных одного сервера *Z39.50*, что сводит на нет прин-

ципиальную возможность эмуляции распределенности в этой модели. Некоторым приятным исключением в семействе многочисленных шлюзов *Z39.50-HTTP* является недавно созданный шлюз российского проекта *Сигла* [177, 178] – совместный информационный проект Научной библиотеки МГУ (<http://www.lib.msu.su/>) и компании “Библиотечная компьютерная сеть” (<http://www.bks-mgu.ru/>). Шлюз *Сигла* включает более 1600 каталогов библиотек всего мира и допускает поиск в любой выбранной комбинации каталогов (см. Рис.8.1). Тем не менее, *Сигла* является всего лишь шлюзом, предоставляющего доступ только к библиографической информации и допускающего ее просмотр в predetermined библиографических форматах.

Если рассматривать интеграцию ресурсов на стороне серверов, то можно выделить несколько моделей:

- Модель централизованного сводного каталога (ЦСК) – данные в централизованную базу данных поступают из различных частных каталогов, сливаются с удалением дублируемых записей, информация об источниках дублируемых записей заносится в специальные поля. Информационную систему, реализующую эту модель нельзя назвать распределенной. Однако, эта модель была и остается очень популярной в библиотечном сообществе. В качестве примера ее реализации можно привести проект *Melvyl®*¹ по интеграции каталогов калифорнийских университетов (29 основных участников, более 20 млн. записей [13]) или российский проект “*ЛИБНЕТ*” [181] (800 тыс. записей сводного каталога, 42 участника на 13.04.2004).
- Модель централизованного хранения ресурсов (ЦХР) – каталоги участников хранятся централизованно, но в разных базах данных. Ответственность за каждую базу данных в этом случае лежит на конкретном участнике проекта. Интеграция ресурсов происходит в реальном времени в момент исполнения запросов. Дублирующие записи остаются. Подобная информационная система также распределенной не является. Модель ЦХР также достаточно популярна. В качестве примера можно привести Центр автоматизации библиотек Флориды (*Florida Center for Library Automation* [182]), обеспечивающий доступ к каталам десяти участников через одну точку входа [183]. В России по этой модели работают все системы, основанные на сервере “*Ruslan Z39.50 server*” (СПбГТУ, Центр “Открытые Библиотечные Системы”, ID=148), т.к. этот сервер обрабатывает запросы только к локальным базам данных.
- Модель распределенного виртуального каталога (РВК) – каталоги участников не перемещаются, а всегда хранятся в виде отдельных

¹Система *Melvyl®* [179] интегрирует ресурсы университетов Калифорнии в рамках системы, включающей ресурсы библиотек разных университетов (Университет Калифорнии в Лос-Анджелесе, Университет Калифорнии в Санта-Барбаре, Университет Калифорнии в Ирвине, и т.д.). *Melvyl®* – только один из ресурсов в Калифорнийской Электронной Библиотеки [180]. Доступ по *Z39.50* возможен по адресу: *melvyl.cdlib.org:210, dbname: CDL90*.

Рис. 8.1: Интерфейсы портала *Сигла*

баз данных на серверах участников. Объединение ресурсов происходит в реальном времени в момент обработки запроса пользователя. В отличие от модели ЦХР в момент обработки запроса здесь требуются дополнительные сетевые соединения и протоколы общения серверов. Таким протоколом чаще всего выступает *Z39.50*. Модель РВК широко обсуждается в литературе (см., например, [28,184,185]), существует множество проектов по ее реализации. В качестве примера можно привести Виртуальный каталог консорциума библиотек Бостона (*Virtual Catalog of the Boston Library Consortium*) [186]. В России модель виртуального каталога реализована в Новосибирской региональной корпоративной библиотечной системе [143–145].

Архитектура всех известных авторам библиотечных проектов по созданию РИС (чаще применяется название "виртуальный сводный каталог", т.к. термин "сводный каталог" ласкает слух библиотечного сообщества) имеет как минимум трехзвенную структуру, где промежуточный слой берет на себя функции переадресации запросов клиентов на соответствующие серверы *Z39.50* и группировки получаемых от последних ответов. Этот промежуточный слой представляет собой или шлюз *Z39.50-HTTP* (модель промежуточного шлюза – МПШ) или специально созданный сервер *Z39.50*, способный перенаправлять запросы клиента на другие серверы (модель промежуточного сервера – МПС). Единой терминологии для последний варианта не существует, промежуточный сервер *Z39.50* называют также шлюзом *Z39.50-Z39.50* и *Z39.50-PROXY* сервером. Модель МПС, очевидно, обладает намного большей функциональностью, чем модель МПШ, т.к., во-первых, она позволяет полностью скрыть от клиента детали обработки его запроса в информационной системе (имена баз данных, маршрут запроса, группировку результатов поиска), что полностью соответствует идеологии *Z39.50*, во-вторых, обеспечивает точку входа в РИС не только клиентам WEB, но и клиентам *Z39.50*, наконец, в-третьих, трехзвенная модель МПС легко расширяется в многозвенную модель с использованием того же самого программного обеспечения.

Характерным примером модели МПШ являются упоминавшийся выше шлюз [175] Библиотеки Конгресса США, но более развитый вариант этой модели реализован в проекте *vCuc* [187–189] по созданию шлюза, предоставляющего единую точку входа к ресурсам всех основных библиотек Канады (более 500). Этот шлюз позволяет выбирать из predetermined списка комбинаций источников для одновременного поиска, что отличает его от шлюза Библиотеки Конгресса США. Другой характерной особенностью шлюза *vCuc* является возможность просматривать индексированные термины (эмуляция *Z39.50* сервиса SCAN). Несмотря на некоторые странности шлюза, связанные с удержанием сессии *Z39.50* для клиента WEB, шлюз предоставляет весь заявленный сервис.

Следует заметить, что многозвенная (трех- и более) модель информационных систем оказалась очень продуктивной. При надлежащей организации обработки информации на промежуточном уровне можно интегрировать ресурсы не только из источников *Z39.50*, но и из других источников структурированной информации.

8.1.1 ZLOT

Проект *ZLOT* [190] (*LOT – The Library of Texas*) инициирован *TSLAC* (*Texas State Library and Archives Commission*) и *TIF* (*Telecommunications Infrastructure Fund Board*) и выполняется с декабря 2001 года. Проект с первоначальным бюджетом 500 тыс. USD призван в конечном счете вывести сервис доступа к информационным ресурсам более 600 публичных и академических библиотек штата на новый уровень². Он охватывает многие компоненты, включая:

- диалоговые базы данных, лицензированные Государственной Библиотекой для широкого доступа;
- сервис *GILS* и поддержку электронной программы сохранения ресурсов;
- виртуальный каталог, основанный на *ANSI/NISO Z39.50* протоколе;
- интерфейсы, интегрирующие поиск и извлечение информации для различных компонент;
- службу доставки документов, основанной на стандартных сервисах и других механизмах.

Для доказательства выполнения концепции на первом этапе (10 месяцев) планировалось создание как минимум виртуального каталога, интегрирующего ресурсы 20-30 библиотек с полнофункциональным сервисом поиска и извлечения данных. Архитектура виртуального каталога *ZLOT* представлена на Рис. 8.2.

Проект *ZLOT* имеет достаточно полное описание в периодических изданиях и на сайте проекта [191–196].

8.1.2 UNInverse

Проект *UNInverse* [197] финансировался Европейской комиссией в 4-ой рамочной программе *Telematics for Libraries* (DG XIII) для разработки и демонстрации масштабируемых интегрированных решений по предоставлению расширенных библиотечных сервисов. Проект был рассчитан на 33 месяца – с октября 1996 года по июнь 1999 года. Стоимость проекта – около 4 млн. экю [198]. Проект основывался на предыдущих проектах, финансируемых ЕС, и был направлен на решение многих проблем распределенного обслуживания и глобализации, с которыми столкнулись библиотеки. Проект ориентировался на открытые стандарты включая *ANSI Z39.50* для поиска и извлечения информации из баз данных, *ILL* [199, 200] для обслуживания заказов и *GEDI* рекомендации для электронной доставки документов.

²“The goal of this project is to demonstrate the viability and effectiveness of a standards-based approach for distributed access to Texas libraries resources.” (William E. Moen, Principal Investigator)

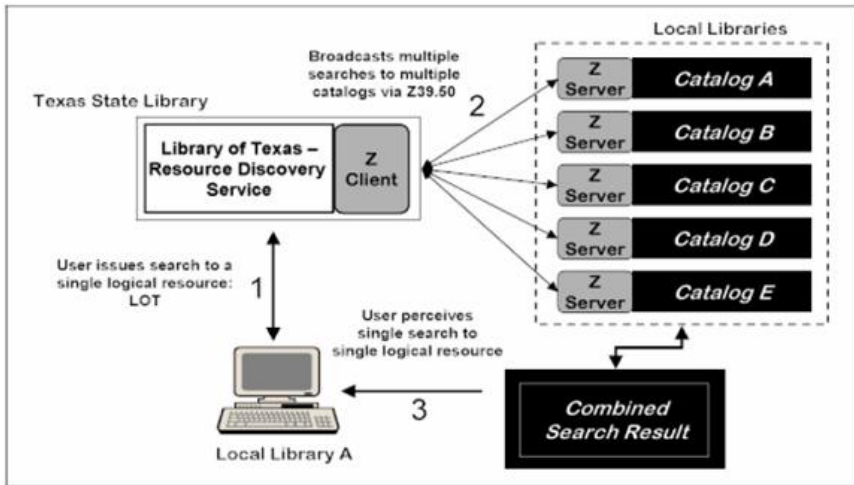


Рис. 8.2: Архитектура виртуального каталога ZLOT

Консорциум *UNiverse* во главе *Fretwell-Downing* (Великобритания) включал организации из 6 государств, что позволило объединить некоторых из ведущих европейских разработчиков сетевых информационных систем с влиятельными практиками от мира академических и национальных библиотек, включая британскую Библиотеку, греческую национальную Библиотеку и ирландский Совет Библиотек.

На первом этапе *UNiverse*-консорциум включал 17 партнеров:

- An Chomhairle Leabharlanna / Irish Library Council (<http://www.iol.ie/~libcounc/>);
- Delft University of Technology Library (<http://www.library.tudelft.nl/>);
- Ex Libris (<http://www.exlibris-usa.com/>);
- Forbairt;
- Fretwell-Downing Data Systems Ltd (<http://www.fdgrouop.co.uk/>);
- Harper Adams Agricultural College;
- Index Data (<http://www.indexdata.dk/>);
- Kyros;
- National Library of Greece;
- Q-Ray;
- Southampton Oceanographic Centre;

- Technical Knowledge Centre & Library of Denmark (<http://www.dtv.dk/>);
- Technical University Library of Norway;
- The British Library (<http://www.bl.uk/>);
- The Freshwater Biological Association (<http://www.ife.ac.uk/fba/>);
- University College Dublin (<http://dallas.ucd.ie/>);
- University of Sheffield (<http://www.shef.ac.uk/>).

UNVerse – логический преемник некоторых проектов, поддерживаемых ЕС, в частности:

- *IRIS* [201, 202] – один из первых основанных на *Z39.50* распределенных каталогов с функциями доставки документов;
- *EUROPAGATE* [203] – *Z39.50* и *ISO SR* шлюз с дополнительной конвертацией записей *MARC*;
- *DALI* [204] – *IRIS* с добавлением функций доставки документов мультимедиа.

Большое количество географически распределенных библиотечных каталогов, которые в рамках *UNVerse* должны вести себя как единый виртуальный каталог, комбинируется с множеством высококачественных библиотечных услуг и конечному пользователю и библиотекаря:

- Поиск и извлечение информации – масштабируемые, прозрачные для пользователя в многобазовом окружении;
- Доставка документов различного типа, интегрированная с механизмами поиска и извлечения;
- Поддержка *ILL* [199, 200], интегрированная с поиском и извлечением;
- Корпоративная каталогизация.

Ядром системы *UNVerse* является распределенный виртуальный каталог, с возможностями которого связаны основные особенности проекта:

- способность выполнять параллельный поиск во множестве физически различных базах данных, которые имеют различные методы доступа, оперируют различным синтаксисом, с различными языками и наборами символов, эмулируя многобазовость в одной логической базе данных;
- способность скрывать от пользователя разнообразие источников данных;

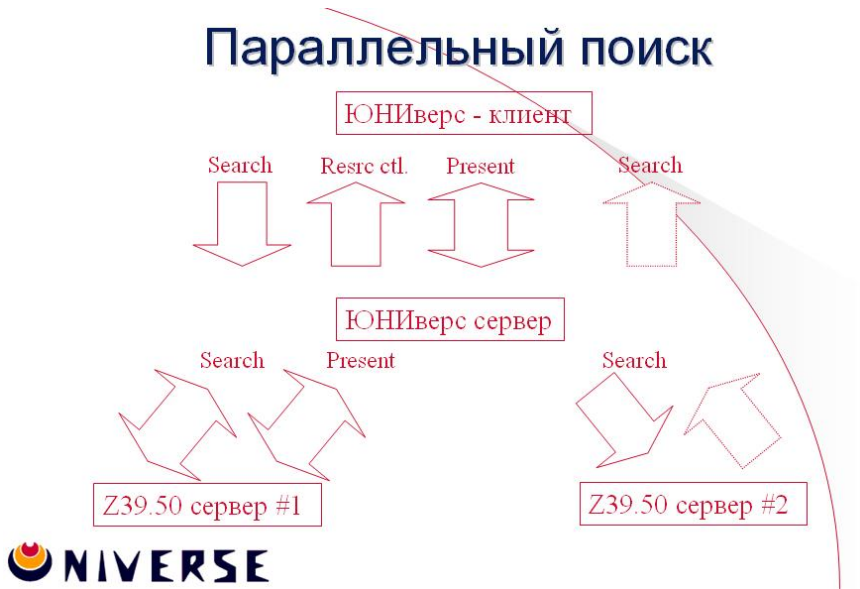


Рис. 8.3: Схема реализации параллельного поиска в *UNIVERSE*

- высокое качество обслуживания в том числе и при объединении данных с устранением дублирования в записях;
- практически неограниченная масштабируемость архитектуры системы, благодаря технологии *ODP* (*Open Distributed Processing*).

Результатом проекта можно считать:

- разработка сервера единого виртуального каталога;
- профиль статуса записи информационного поиска, обеспечивающий пользователя текущим статусом записи его поиска;
- разработка базы данных *Explain*, куда входят средства *Explain* для различных объектов в распределенном каталоге;
- обеспечение стержневой (kernel) операции, которая полностью конфигурируется из базы данных *Explain*, давая насыщенные "меж-областные" взаимодействия;
- проект профиля для доступа к многоязычному тезаурусу через *Z39.50*;
- предоставлению пользователю возможности группировки дублируемых записей;
- полную реализацию протокола *ISO ILL*, включая систему управления рабочим потоком для межбиблиотечного абонемента (МБА) и доставки документов;

- объединение электронной доставки документов на стандартной основе.

Одна из характерных особенностей проекта *UNiverse* состоит в том, что в нем существенно уточнены, правда только для нужд проекта, многие детали сервисов *Z39.50*, касающиеся функционирования распределенного виртуального каталога [205]- [206]. Следует также отметить, что именно в этом проекте, наверное, впервые была озвучена необходимость поддержки доступа к тезаурусам по протоколу *Z39.50* в соответствии со стандартной схемой.

Проект *UNiverse* явился стартовой площадкой для многих более поздних проектов. В частности, можно отметить проект *UNEX (UNiverse Expansion)* [207], в котором в качестве партнера участвовала Библиотека Академии Наук России (БАН). Проект исполнялся с ноября 1998 года по июнь 2000 года. Несмотря на то, что с системной и технологической точки зрения проект не представлял из себя ничего нового и являлся лишь демонстрацией возможностей программного обеспечения *UNiverse* по расширению списка партнеров (см., например, [208]), это был первый международный проект по интеграции ресурсов на основе *Z39.50*, в котором участвовала Россия.

8.1.3 АРБИКОН

Ассоциация региональных библиотечных консорциумов (*АРБИКОН*) [209] является некоммерческим партнерством, объединяющим российские библиотеки и информационные центры, ставшие его членами на добровольной основе. *АРБИКОН* зарегистрирован 18 мая 2002 года. Учредителями *АРБИКОНа* являются следующие организации:

- Российская книжная палата (РКП);
- Некоммерческий фонд «Пушкинская библиотека»;
- Санкт-Петербургский государственный политехнический университет (СПбГПУ).

Ассоциация *АРБИКОН* является результатом выполнения серии проектов *Института «Открытое общество» (Фонд Сороса)* в рамках программы «Автоматизация библиотек» по созданию корпоративных библиотечных систем, в которых участвовало около 250 библиотек и информационных центров. Общая сумма вложенных средств составила 6 миллионов долларов США, включая финансирование *ИОО (Институт «Открытое Общество»)* и софинансирование участников.

К настоящему моменту *АРБИКОН* интегрирует распределенные библиотечно-информационные ресурсы (см. [210]) множества ведущих библиотек России. Точками доступа к этим ресурсам является как центральный портал *АРБИКОН* (см. Рис. 8.4), так и периферийные шлюзы корпоративных библиотечных систем.

The screenshot displays the ARBICON portal interface. The top navigation bar includes links for 'Новости', 'Документы', 'О проекте', 'Обучение', 'Сервисы', 'Гостевая книга', and 'Авторизация'. The main content area is titled 'Сервисы' and provides information about expanded search capabilities and MARC record exports. Below this is a search interface with filters for 'Материалы' (All, Monographs, Serials, Articles) and 'Базы данных' (RUSLANet, Consensus Omnium, OKBC, Yaroslavl KBC). Search criteria include 'Заглавие' (unix), 'И' (I), and 'Тематика'. The number of records to retrieve is set to 20.

The 'Статистика' section features a bar chart titled 'Использование ресурсов портала (май 2004)'. The chart shows the number of requests per day from May 1st to 24th. Below the chart is a table with columns for 'Число Запросов', 'Доля', and 'Ср. часовая'.

A 'Классификатор' dialog box is open, showing a list of classification codes under the heading '38 ГЕОЛОГИЯ'. The selected code is '38.01.05 Материалы общего характера'. Buttons for 'Назад', 'Принять', and 'Отмена' are visible at the bottom of the dialog.

День	Число Запросов
1	596
2	872
3	550
4	548
5	1127
6	844
7	986
8	461
9	284
10	652
11	1019
12	780
13	1140
14	878
15	327
16	226
17	1318
18	1036
19	814
20	1141
21	
22	
23	272
24	163

Рис. 8.4: Интерфейсы портала Арбикон

По доступу к информационным ресурсам портал *АРБИКОН* является обычным шлюзом *Z39.50-HTTP* в реализации *РУСЛАН* [176] с достаточно примитивным пользовательским интерфейсом с предопределенным списком источников информации. Однако, существует несколько моментов, которые, на наш взгляд, выгодно отличают этот шлюз от аналогичных проектов.

- На шлюзе ведется статистика использования информационных ресурсов с наглядным представлением информации (см. Рис. 8.4). Это несомненно повышает уровень интеграции распределенной системы *АРБИКОН* как единого целого.
- На шлюзе реализована, пусть примитивно, работа с рубрикаторами и тезаурусами (*ГРНТИ*, *УДК*, *MeSH*), что выводит поисковую систему на качественно более высокий уровень.
- Система *АРБИКОН* и ее шлюз развиваются. В частности, ведутся работы по централизованной авторизации пользователей распределенной системы по технологии *LDAP* (см. [211]).

Наконец, следует отметить, что сегодня *АРБИКОН* является наиболее масштабным проектом по созданию РИС на основе протокола *Z39.50* в России, интегрирующим библиотечные ресурсы различных регионов страны. Мы надеемся, что эта работа будет продолжаться.

8.2 *GILS*

Вместе с развитием технологий создания электронных ресурсов постоянно растет их ассортимент и количество. Традиционные библиотечные технологии, ориентированные на каталогизацию печатной продукции и других материальных объектов, не могут полностью обеспечить информационное обслуживание общества. Для каталогизации электронных ресурсов, создания для них баз метаданных, адекватно описывающих эти ресурсы, для разработки более универсальных информационных систем нужны другие подходы, выходящие за рамки традиционных *MARC*-ориентированных систем.

Одной из первых попыток решения проблемы каталогизации электронных ресурсов было создание профиля *GILS* (*Government Information Locator Service*) для государственных учреждений США в 90-х годах прошлого века (см, например, [47–49]). *GILS* создавался вместе со стандартом *Z39.50* и как прикладной профиль является его частью. Этот профиль содержит описание схемы данных (метаданных) *GILS*, включающей наборы поисковых атрибутов, набор меток, наборы элементов и структуру записи *GILS* [1, 47].

На основе *GILS* в США была принята правительственная программа создания федеральной информационной системы для доступа общественности к правительственной информации. Программой предусматривалось создание ресурсов в соответствии с *GILS* в различных правительственных

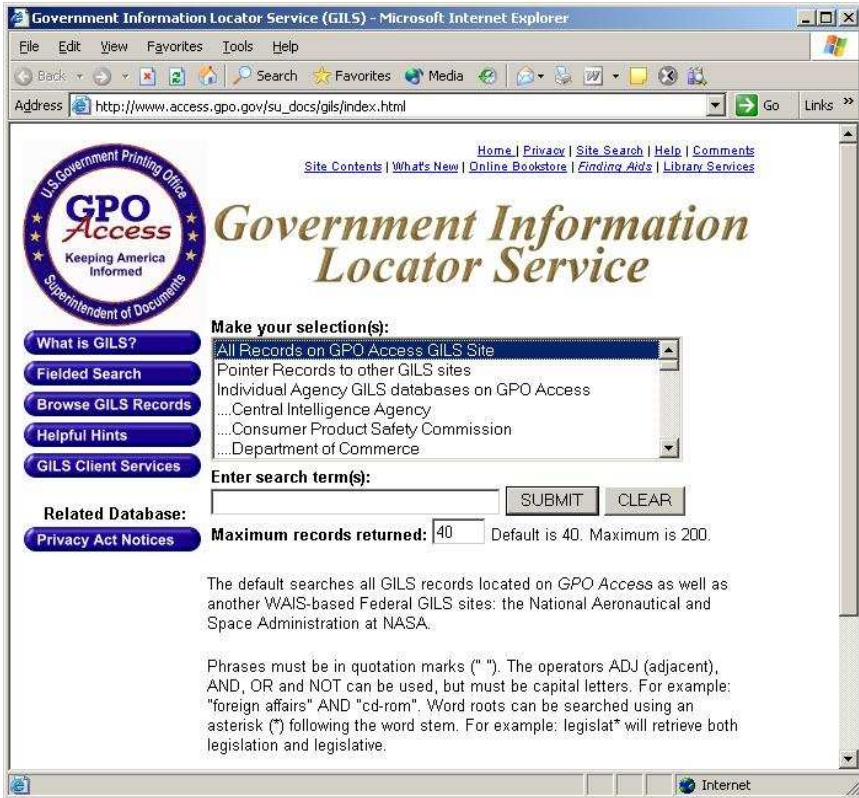


Рис. 8.5: Шлюз GILS U.S. Government Printing Office

учреждениях и обеспечение доступа к этим ресурсам по протоколу *Z39.50*, т.е. создания распределенной информационной системы на основе *Z39.50* и *GILS*. Следует заметить, что эта система была создана и сегодня существует [50, 212]. Она является составной частью, например, портала правительственной информации США (*FirstGov*) [213]. В качестве примера точки доступа можно указать *GPO (U.S. Government Printing Office)* [214] и многие другие (см., например, [215, 216]).

Информация о *GILS* может быть найдена на специализированном сайте [51]. Примечательно, что сегодня вне федеральной информационной системы США аббревиатуру *GILS* принято толковать как *Global Information Locator Service*, подчеркивая тем самым универсальность подхода *GILS* к описанию информационных ресурсов.

8.3 Электронные библиотеки

8.3.1 LAURIN

LAURIN [217] (*Libraries and Archives Collecting Newspaper Clippings Unified for their Integration into Networks*) – проект по созданию информационной системы по газетным публикациям. Проект финансировался ЕС с мая 1998 года по август 2000 года в 4-ой рамочной программе (*Telematics for Libraries*) с привлечением семнадцати участников от нескольких стран, включая две компании по разработке программного обеспечения и большую группу библиотек.

Главные цели проекта *LAURIN* (см. [218]– [232]):

1. Создание прототипа распределенной информационной системы оцифрованных газет и архива вырезок, предоставляющей доступ через Internet централизованным способом, для того, чтобы искать и отыскивать отсечения.
2. Создание универсальной модели, которая использовалась бы индивидуальными библиотеками для сканирования, оцифровки, сохранения, и индексирования газетных вырезок, делая их доступными через *LAURIN* сеть.

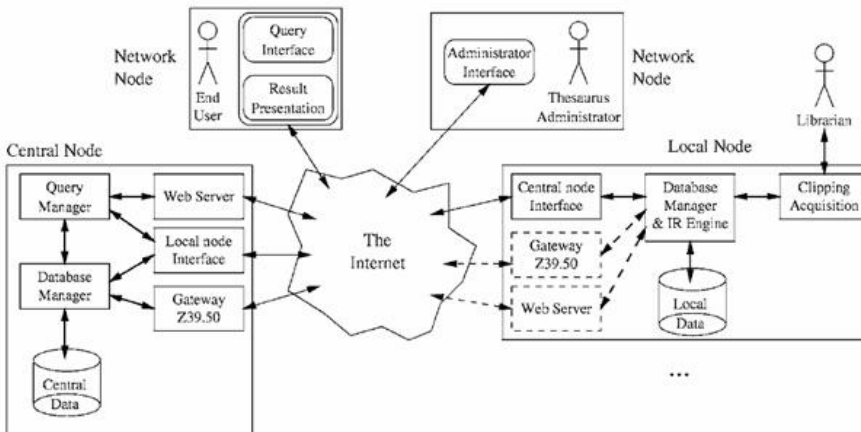


Рис. 8.6: Общая архитектура *LAURIN*

Система *LAURIN* включает три (четыре) основные части:

- *LAURIN* база данных – набор скриптов, запускаемых в СУБД Oracle, устанавливающей таблицы, синонимов, процедур, триггеров и т.п. *LAURIN* системы и заполняющий ее некоторыми основными данными. *LAURIN* база данных доступна свободно (Version: 2.0.2 от 2002-07-29 – http://laurin.uibk.ac.at/download_l_1.html).

- *LIS* – набор программ интерфейса *LAURIN*, т.е. несколько сетевых интерфейсов, написанных в PERL для индексации, поиска и административных целей. *LIS* свободно доступны (Ver. 2.0.6 от 2004-01-23 – http://laurin.uibk.ac.at/download_l_1.html)
- Тезаурус *LAURIN* – набор управляемого словаря для индексации. Он предназначен для охвата большинства подчиненных заголовков, которые могут быть найдены в газетах [233].
- *libClip* – инструмент для сканирования и распознавания текста. *LibClip* – коммерческий продукт, но он никогда не был доступен на рынке.

Следует отметить, что основные разработки в рамках проекта *LAURIN* распространяются бесплатно и могут быть загружены и использованы всеми заинтересованными сторонами.

8.3.2 DNER

DNER (*Distributed National Electronic Resource*) [234] – британская программа создания сети национальных распределенных информационных ресурсов. В рамках или под эгидой *DNER* исполнялось и исполняется несколько подпрограмм и проектов. Некоторые из них так или иначе связаны с применением протокола *Z39.50*.

Программа *JOIN-UP* [235] комитета *JISC* [236] (*Joint Information Systems Committee*) содержит четыре проекта (*Xgrain*, *ZBLSA*, *ZETOC*, *Docusend*), три из которых основаны на *Z39.50*.

Xgrain

*EDINA*³ *Xgrain* [239] проект (произносится как "cross grain") является частью программы *JOIN-UP* [235]. Проект предназначен для разработки *Z39.50* – совместимого брокера и прототипа портала для обучения и преподавания. В рамках проекта предусматривается

- Предоставление брокера для использования порталами *DNER* и местными организациями для взаимного поиска в разнородных *Z39.50* – совместимых источниках (аннотациях, оглавлениях, индексированных данных и др.).
- Предоставление порталам возможности для конечного пользователя выбора как грубого взаимного поиска, так и более детального, посредством перехода по ссылке на простые и интуитивные интерфейсы.

³"Edina" – древнее и поэтическое название Эдинбурга (Шотландия). Роберт Бернс написал "Адрес Эдинбургу", который начинается, "Эдина! Дорогое место Скотии!.. (Edina! Scotia's darling seat!)" (см. <http://edina.ed.ac.uk/about/poem.shtml>). Организация *EDINA* [237, 238] основана при Эдинбургском университете, финансируется *JISC* (*Joint Information Systems Committee* [236]) и предоставляет сетевой доступ к информационным ресурсам для образовательных и академических учреждений. Для британских пользователей ресурсы, как правило, бесплатны, но для большинства услуг требуется регистрация.

ZBLSA

ZBLSA (*Z39.50 Broker to Locate Serials and Articles*, см. [240]) – двухлетний проект *EDINA* [238] начат 1 января 2002 года, окончен 31 декабря 2003 года. Цель проекта *ZBLSA* состоит в том, чтобы расширить функциональность *DNER*, обеспечивая связь с брокером, который обрабатывает соединения по ссылкам на статьи из журналов, указывая их местоположение и условия доступа к полным текстам в напечатанной или электронной форме. Результаты *ZBLSA* планируется использовать на предметных порталах *RDN* (*Resource Discovery Network*). Задачи проекта *ZBLSA*:

- Разработка прототипа *ZBLSA* сервера, обслуживающего один или несколько порталов центров данных и предметных порталов.
- Расширение доступа к информации в *OPAC*-системах и объединение каталогов через *Z39.50*
- Разработка *HTTP*-интерфейсов, демонстрирующих для конечного пользователя прямой доступ к *ZBLSA*-сервисам.
- Идентификация подходящих электронных ресурсов, включая ссылки на полные тексты на сайтах издателей.
- Создание технической документации для проектировщиков порталов для включения сервисов *ZBLSA* в операции порталов.

ZETOC

Проект *ZETOC* [241]- [242] заключался в создании сервисов для *MIMAS* (*Manchester Information & Associated Services*) [243] по предоставлению *Z39.50* доступа к электронным ресурсам (оглавлениям журналов и материалов конференций) Британской библиотеки (<http://www.bl.uk/>). *ZETOC* является ключевым сервисом для трех тематических порталов.

Сегодня *ZETOC* обслуживает базы данных с объемом более 21 млн. записей с ежедневным пополнением 8000 статей. Количество журналов – 20 тыс., материалов конференций – 16 тыс. Доступ к ресурсам по *Z39.50* (требуется авторизация):

Address: *zetoc.mimas.ac.uk*
 Port: 2121
 Database name: *zetoc*
 Record syntax: *SUTRS, GRS-1* or *XML*

Специфической особенностью *ZETOC* является его ориентация на *Dublin Core* [244]. Библиографические записи не предоставляются в *MARC*-форматах, доступными структурированными форматами являются *GRS-1* и *XML*. Примеры записи *ZETOC* в форматах *SUTRS* и *XML* приведены ниже

Article Title: Dublin Core Metadata for Electronic Journals
 Author(s): Apps, A; MacIntyre, R
 Journal Title: LECTURE NOTES IN COMPUTER SCIENCE
 ISSN: 0302-9743
 Volume: 1923
 Year: 2000
 Jnl Issue Title: Research and Advanced Technology for Digital Libraries
 Page(s): 93-102
 Editor(s): Borbinha, J; Baker, T
 Publisher: Germany : Springer-Verlag
 Language: English
 Dewey Class: 004
 LC Class: TP372.5
 BLDSC shelfmark: 5180.185000
 ZETOC ID: RN085008791

```
<zetocrec>
  <dc:title>Dublin Core Metadata for ElectronicJournals</dc:title>
  <dc:creator scheme="zetoc">
    <zetoc:snm>Apps</zetoc:snm>
    <zetoc:inits>A</zetoc:inits>
  </dc:creator>
  <dc:creator scheme="zetoc">
    <zetoc:snm>MacIntyre</zetoc:snm>
    <zetoc:inits>R</zetoc:inits>
  </dc:creator>
  <!--Library of Congress-->
  <dc:subject scheme="LCC">TP372.5</dc:subject>
  <!--Dewey -->
  <dc:subject scheme="DDC">004</dc:subject>
  <dc:publisher>
    <zetoc:pnm>Springer-Verlag</zetoc:pnm>
    <zetoc:country>Germany</zetoc:country>
  </dc:publisher>
  <dc:contributor scheme="zetoc" role="editor">
    <zetoc:snm>Borbinha</zetoc:snm>
    <zetoc:inits>J</zetoc:inits>
  </dc:contributor>
  <dc:contributor scheme="zetoc" role="editor">
    <zetoc:snm>Baker</zetoc:snm>
    <zetoc:inits>T</zetoc:inits>
  </dc:contributor>
  <dc:date refine="issued" scheme="W3CDTF">2000</dc:date>
  <!--zetoc unique identifier-->
  <dc:identifier refine="zetoc">RN085008791</dc:identifier>
  <dc:identifier refine="shelfMark">5180.185000</dc:identifier>
  <dc:identifier refine="citation">
    <dccite:journalTitleFull>Lecture Notes in Computer
      Science</dccite:journalTitleFull>
    <dccite:journalVolume>1923</dccite:journalVolume>
    <dccite:journalPages>
      <zetoc:ppf>93</zetoc:ppf>
      <zetoc:ppl>102</zetoc:ppl>
    </dc:identifier>
</zetocrec>
```

```

</dctype:journalPages>
<zetoc:journalIssueTitle>Research and Advanced Technology
  for Digital Libraries </zetoc:journalIssueTitle>
</dc:identifier>
<dc:language scheme="RFC1766">en</dc:language>
<dc:relation refine="isPartOf" scheme="ISSN">0302-9743
  </dc:relation>
</zetocrec>

```

Кроме перечисленных *DNER* поддерживает несколько проектов под общим именем *Z*-проекты.

Z Projects

Gate-Z [245] – разработка шлюза *Z39.50-Z39.50* для взаимодействия *Z39.50* клиентов *Bath*-профиля [246] с другими серверами *Z39.50*.

Z39.50: Authentication and Development [247] – разработка метаданных для доступа к данным и сервисам *MIMAS* [243].

The JAFER Toolkit – см. ниже.

JAFER

JAFER (Java Access For Electronic Resources) [248–251] – стартовал в октябре 2000 года. Основная задача проекта – создание простого инструментария для доступа к информационным ресурсам *Z39.50*, необходимого для разработки информационных порталов. Основанием для создания такого инструментария явилась техническая сложность протокола *Z39.50* для создания простых интерфейсов пользователя.

JAFER Toolkit состоит из трех компонент:

1. Серверная компонента обеспечивает доступ к базам данных, например по *ODBC*, через *DNER* по протоколу *Z39.50*. Серверная компонента включает визуальный инструментарий по установке и настройке.
2. Компонента шлюза обеспечивает *Z39.50*-доступ к другим, не *Z39.50*-источникам информации, что увеличивает функциональность системы.
3. Компонента поддержки разработчиков порталов позволяет динамически формировать *WEB*-страницы для доступа по протоколу *Z39.50* к различной информации (библиотеки, музеи, архивы), используя единый инструментарий и дизайн.

JAFER написан на *Java*, поддерживает *Open Source* и тесно связан с другими проектами *DNER*. *JAFER* призван упростить применение протокола *Z39.50*.

В настоящее время *JISC* поддерживает четыре направления, связанные с доступом к распределенной информации, в том числе и по протоколу *Z39.50*.

SERVER:

Can be used in a variety of ways to provide different functionality, the simplest case being where a user needs to provide a Z39.50 interface to an existing non-Z39.50 database.

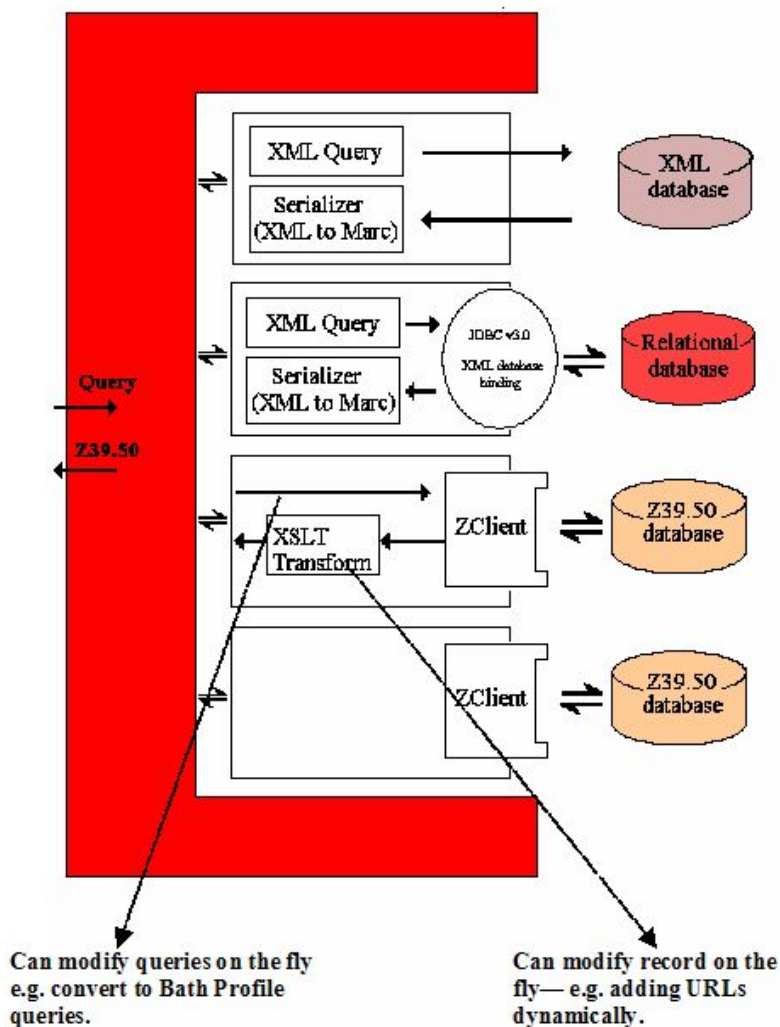


Рис. 8.7: Архитектура сервера JAFER

Рис. 8.8: Интерфейс *Cheshire II*

Interoperability Focus [252] – ориентирован на интеграцию различных ресурсов, в том числе на уровне метаданных, и обеспечения сквозного поиска в таких источниках как библиотеки, музеи, архивы и т.п. (см. также *AHDS* – раздел 8.4.3)

HILT (*High-Level Thesaurus*) [253] – проект по выработке рекомендаций для создания тезауруса, способного обеспечить потребности перекрестного поиска и навигации по ресурсам различного типа.

Renardus [254] – европейский проект доступа к разнородным ресурсам на основе классификации Дьюи (*DDC – Dewey Decimal Classification*) через единый интерфейс.

Cheshire [255, 256] является другим проектом, стремящимся облегчить доступ к разнородным ресурсам, в числе которых собрания ”оригинальных материалов” (старопечатные книги, отчеты, архивы, средневековые и литературные рукописи, и экспонаты музеев), статистические базы данных, полные тексты, геопространственные и мультимедийные данные. В этом проекте *JISC* и Университет Ливерпуля

сотрудничает с *NSF*. Проект ориентирован на архитектуру "клиент-сервер" с поддержкой *Z39.50* и *SGML*.

8.3.3 *ELISE II*

Проект *ELISE II* [257], поддержанный DGXIII ЕС, является хорошим примером совместной работы библиотек, музеев и архивов, для улучшения доступа к цифровым ресурсам. Это вторая стадия успешного проекта *ELISE* от Третьей Рамочной Программы (*Third Framework Telematics Programme, Libraries area*).

Полный бюджет Проекта составил 3072 тыс. эю, вклад ЕС составил 1606 тыс. эю. Проект начал осуществляться 1 октября 1996 года и первоначально планировался завершиться 30 сентября 1999 года.

В проекте *ELISE* приняло участие 9 партнеров (6 основных и 3 ассоциированных) из четырех государств (Великобритания, Нидерланды, Ирландия и Бельгия). Эти партнеры являлись представителями библиотек, музеев, академических учреждений, коммерческих и полукommerческих поставщиков изображений:

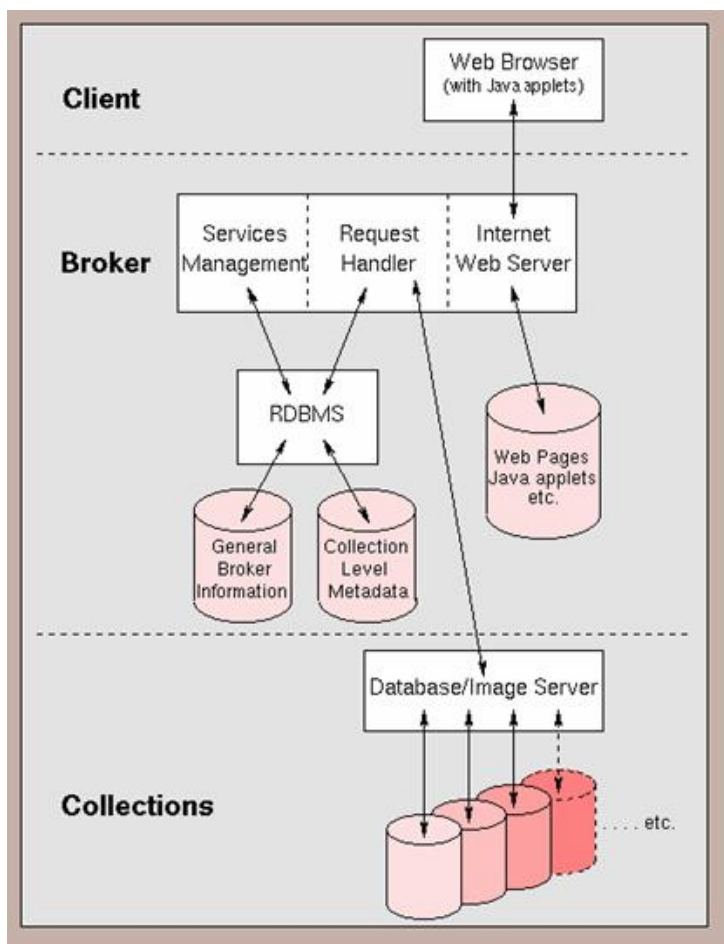
1. De Montfort University (UK);
2. Laboratium voor Biomedische Informatika (Netherlands)⁴;
3. Tilburg University (Netherlands);
4. IBM UK Scientific Centre (UK);
5. Victoria and Albert Museum (UK);
6. University of Limerick (Ireland);
7. Radio Telefis Eireann (Ireland)⁴;
8. Hunt Museum (Ireland)⁴;
9. Universite Libre de Bruxelles (Belgium).

Проект *ELISE II* позиционировал *ELISE* от ранней стадии прототипа до стадии опытно-промышленной эксплуатации банков изображений библиотек Европы.

Команда исполнителей Проекта должна была создать полнофункциональный демонстрационный вариант системы. Развитие этой демонстрационной системы было непосредственно связано с рыночным исследованием и анализом, также сделанным в проекте. Это исследование давалось не очень легко, поскольку потенциальные пользователи часто не сознавали технологию доступа к изображениям.

Шесть партнеров *ELISE II* производили банки данных с суммарным объемом около 40000 переведенных в цифровую форму изображений и их

⁴ассоциативные партнеры

Рис. 8.9: Архитектура *ELISE II*

метаописаний, а один партнер производил банк видеоизображений. Партнеры справлялись сами с такими проблемами как вербовка соответствующего штата, выбор изображений, индексация, решение технических вопросов, авторское право и т.д.

Краткое описание банков изображения *ELISE*, приведенное ниже, дает хороший пример того, как новые технологии могут улучшить доступ к разнородным ресурсам изображений. При этом важность совместной работы участников не может быть недооценена.

DMU: банк изображений (около 10 000 объектов). Представляет собой смесь диапозитивов, стеклянных пластинок, отпечатков и 35 мм слайдов. Содержание смешанное: от изображений типа "минувших дней", до изображений Gimson-мебели и других объектов искусства и музейных экспонатов.

HM: банк изображений (2000 объектов) собрания музеев и картин. Изображения обеспечивают визуальный охват полного содержания музея.

LBMI: банк изображений (3000 объектов) представляющий собой оригинальный материал, включающий: микроскопические препараты, фотографии, слайды, цифровые изображения и другие. Эти изображения разделены на различные категории согласно биомедицинским дисциплинам и должны включать: гематологию, дерматологию, пульмалогию, патологию, таксономию с патологией, которую считают особенно важной.

RTE: банк изображений (20000 объектов). Оригинальный материал различен, включая фотографии, слайды и стеклянные пластинки. Содержание также различно, охватывая ирландскую историю, культуру, искусства/ремесла, архитектуру, места, пейзажи, социальные действия, спортивные состязания и т.д. Выбор материала для обработки базировался на рейтинге интереса в радиовещании. Отобранный материал включает следующие разделы: Персоналии, Рукописи, Религия, Люди, Деррай Массэйкр, Демонстрации, Сепаратизм (Fenians), Встречи, Бунты – Северная Ирландия, Восстания, Выборы, Эмиграция, Правительства, Политические Партии, Забастовки, Земельная реформа, Бедность, Военнопленные, Гражданские права, Тюрьмы, Суды, Полиция, Выселения/Эмиграция (голод).

ULB: банк изображений (1000 объектов). Собрание *Bertelson* состоит из интересного собрания газет, которые сохранены в ULB библиотеках. Эти газеты датированы главным образом с конца 19-ого столетия и в начале 20-ого столетия и иллюстрируют широкое разнообразие тогдашней бельгийской журналистики. Банк изображений, показывая интересный графический аспект заголовков газет, обеспечивает своего рода витрину собрания Bertelson и более широко отдела редкой книги библиотеки.

V&A: банк изображений (3000 объектов) состоит из нескольких областей: коллекция "Западная Мебель: от 1350 года до настоящего времени"; коллекция картин Индуса Акбарнамы; экспонаты, связанные с Большой Выставкой 1850; экспонаты выставки "Лезвие" (мода 20-го столетия).

TU: video bank. Используется видео материал архива Брабанта (Северный Брабантский Архив Кино), банк будет состоять из 80 цифровых фильмов.

ELISE II использует в своих интересах последние достижения в цифровой технологии, связывая базы данных и банки изображений в различных странах с каждым партнером. Предоставляя доступ ко всем ресурсам, *ELISE II*, таким образом, сокращает потребность хранения копии информации в каждом центре. Проектом предусмотрены работы по отбору ресурсов для оцифровки и индексации, по авторскому праву, по управлению.

Проект *ELISE* стремился демонстрировать сервисы, которые обеспечивают доступ к сложным коллекциям и банкам изображений, особенно поощрялось использование открытых стандартов для передачи данных, например, *Z39.50*, и для представления структуры и содержания метаданных, например, *Dublin Core*.

Сотрудничество с другими проектами и странами, как и взаимодействие между партнерами, явилось существенной особенностью *ELISE II*. Работа проводилась в виде консультаций с другими группами, решающих аналогичные проблемы, из Европы, Северной Америки и других стран, например:

- Работа вместе с группой *ZIG* разработчиков *Z39.50* и Консорциумом *CIMI* для продвижения стандартов типа *CIMI* и *Dublin Core*.
- *CIMI*, членами которого являются *DMU* и *V&A*, работал в течение более чем 18 месяцев над интероперабельностью сетевого обмена данными между музеями, используя *WEB*, *HTML*, *SGML* и *Z39.50* как транспортное средство.
- Проект *COPEARMS* – *Co-ordinating Project for Electronic Authors Right Management Systems*
- Проект *Aquarelle* – The Information Network on Cultural Heritage
- *HELIX* – *Higher Education Library Image eXchange* (UK)
- *JIDI* – *JISC Image Digitisation Initiative* (UK)
- *TASI: The Technical Advisory Service for Images* (UK)

Работы по тематике *ELISE II* продолжаются, при этом приветствуются все новые группы участников.

8.3.4 *EULER*

Проект *EULER* (European Libraries and Electronic Resources in Mathematical Sciences) [258] финансировался с апреля 1998 года по ноябрь 2002 года Европейской Комиссией в рамках программы *Telematics Applications* подпрограммы *Telematics for Libraries*. Главная цель проекта *EULER* – интеграция всех доступных в электронном виде ресурсов в области математики в рамках электронной библиотеки гетерогенных источников [258, 259].

Отправной точкой для идеи проекта послужил тот факт, что быстрое увеличение объемов сетевых ресурсов в области математики заставляет пользователей постоянно переключаться между различными системами и различными пользовательскими интерфейсами. В качестве источников информации выступают:

- Научные библиографические базы данных;
- Библиотеки *OPAC* с услугами доставки документов;

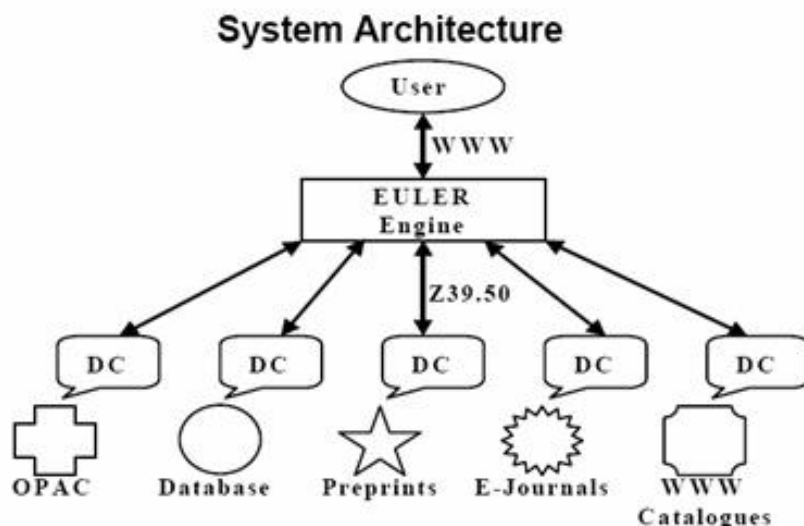


Рис. 8.10: Архитектура информационной системы *EULER*

- Электронные журналы от академических издателей;
- Архивы предпечатных изданий и серой литературы;
- Предметно-ориентированные системы Интернет (порталы);
- Различные поисковые систем Интернет.

Предполагается, что эти типы ресурсов являются наиболее часто используемыми при проведении научных исследований.

Цель проекта *EULER* – создание единственной точки входа в информационную систему, интегрирующую вышеупомянутые ресурсы, и предоставление пользователю единого интерфейса доступа. Интерфейс разрабатывался в тесном сотрудничестве с математическим пользовательским сообществом.

Для интеграции информационных ресурсов использовались общие описания (метаданные), основанные на *Dublin Core*, доступ к которым предоставляется по протоколу *Z39.50*. Информационная система имеет архитектуру, изображенную на Рис.8.10. При этом базы метаданных являются распределенными, каждая из них поддерживается своим поставщиком информации, интеграция ресурсов происходит в реальном времени. Фактически в проекте реализована модель виртуального каталога.

Основными результатами проекта *EULER* можно считать [260]:

- Разработка системы для гомогенного доступа к гетерогенным распределенным ресурсам по математике.

- Создание общего профиля метаданных, основанных на *Dublin Core*, но доработанных *EULER* для определенных потребностей.
- Участие важных поставщиков данных.
- Признание в пользовательском сообществе Европейского Математического Общества (*EMS*).

На сегодняшний день информационная система *EULER* имеет более 2,4 млн. записей, которые постоянно пополняются. Поставщиками являются:

Zentralblatt MATH – охватывает полный спектр математики, включая применения в информатике, механике, физике, и т.д., отредактированные Heidelberg Akademie der Wissenschaften, European Mathematical Society и Fachinformationszentrum Karlsruhe. Многие записи базы метаданных MATH содержат указания о заказе и электронной доставке полных текстов статей.

CWI Amsterdam, Library of the Dutch national research centre for mathematics and computer science. База данных CWI содержит ссылки на литературу по математике и компьютерным наукам из сводного каталога *CWI Library Union Catalogue* (OPAC). В настоящий момент имеются описания печатных и мультимедиа ресурсов.

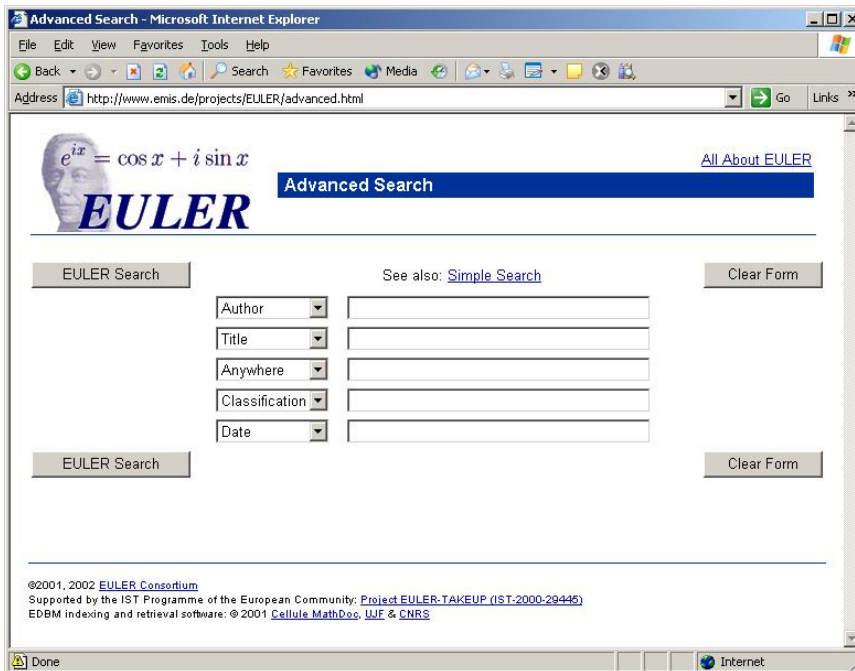


Рис. 8.11: Поисковый интерфейс системы *EULER*

Universita degli Studi di Firenze , mathematical library holdings of Florence University. Содержит ссылки на литературу по математике и компьютерным наукам из сводного каталога библиотек Флоренции.

NS und UG – Niedersächsische Staats- und Universitätsbibliothek Göttingen. Холдинг включает более 4,5 млн. томов, около 15 тыс. текущих журналов, 1,1 млн. микрофильмов.

Jahrbuch über die Fortschritte der Mathematik, Electronic Research Archive for Mathematics (ERAM). ERAM поддерживает электронный архив самых важных математических публикаций периода 1868-1942 и базы данных, основанный на *Jahrbuch*. Содержит ссылки к отсканированным факсимиле оригинальных работ. Метаданные для *Jahrbuch* могут быть использованы для некоторых услуг электронной доставки документов.

Проект *EULER* подготовил хорошую основу для создания всемирной электронной библиотеки по математике (см. [260]). Поисковый интерфейс пользователя представлен на Рис.8.11 (см. [261]).

8.4 Пректы по культурному наследию

8.4.1 *Aquarelle*

Aquarelle – трехлетний проект, поддержанный *Telematics Applications Programme* Европейского союза, ориентирован на проектирование системы доставки ресурсов через Интернет, относящихся к культурному наследию [262]. Система основана на протоколе *Z39.50* и должна поддерживать доступ к гетерогенным базам данных, включая репозитории *SGML*-документов [263].

Aquarelle – информационная система по культурному наследию, которая предлагает доступ к огромным информационным массивам, которые созданы общественными организациями и, до меньшей степени, некоторыми частными лицами, и которые вместе документируют культурное наследие четырех стран – Греции, Италии, Франции и Великобритании (см. [264]).

Полная архитектура *Aquarelle* разработана для освобождения пользователей от тяжелой ручной работы поддержки перекрестных ссылок. *Aquarelle* обеспечивает службу информационного поиска в гетерогенных базах данных.

Метаданные в системе *Aquarelle* основаны на *CIMI* [66]. Более того, сам профиль *CIMI* во многом обязан именно этому проекту [67, 265].

8.4.2 *Zavier*

Австралийский проект *Zavier* (*Z39.50 Arts Victoria Interoperability Project*) [266] – экспериментальный проект создания точки доступа через Интернет

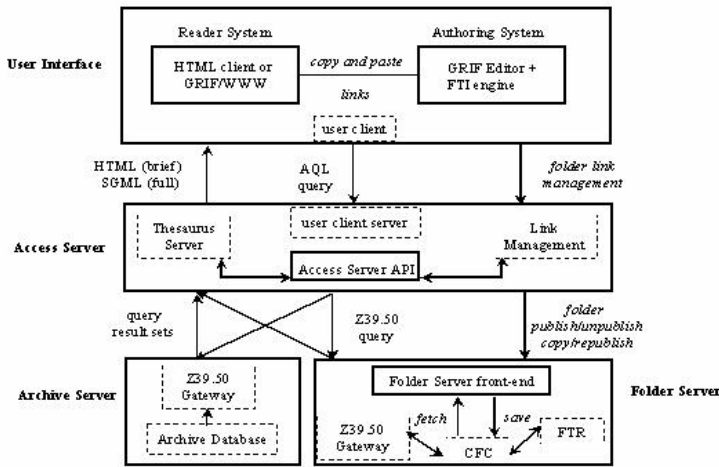


Рис. 8.12: Архитектура Aquarelle

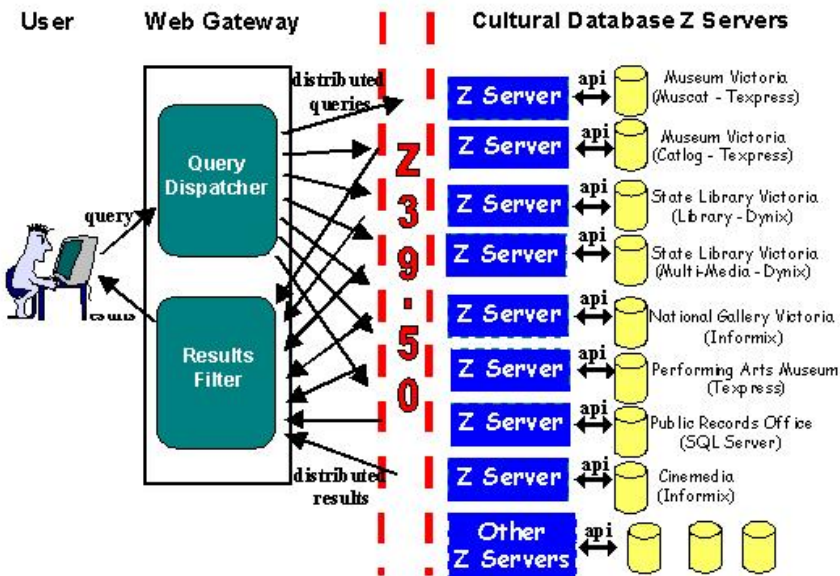


Рис. 8.13: Архитектура Zavier

к распределенным гетерогенным Викторианским культурным базам данных.

Проект *Zavier* был инициализирован *COMDIG* (*Victorian Cultural*

Organisations Metadata and Database Interoperability Group), подкомитетом *ARTVICMM (Arts Victoria Multimedia Group)*. Главная цель проекта – определение через экспериментальную систему возможности *Z39.50* и профиля *CIMI (Consortium for the Interchange of Museum Information)* [66] для сквозного поиска и извлечения информации из баз данных по культурному наследию.

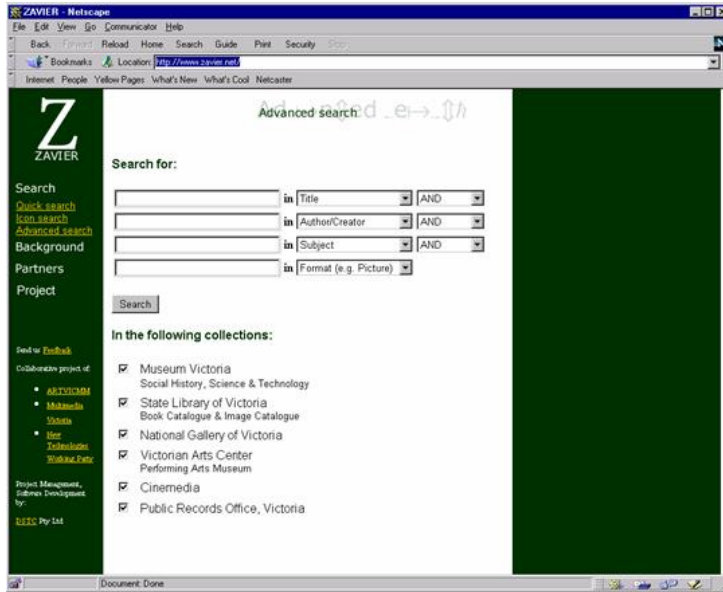


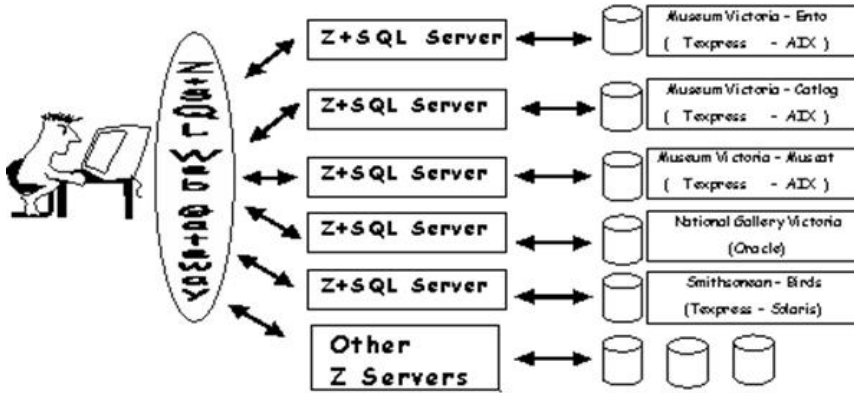
Рис. 8.14: Шлюз *Zavier*. Интерфейс для поиска

В пилотном проекте были интегрированы ресурсы шести Викторианских агентств: the Museum of Victoria, the National Gallery of Victoria, the State Library of Victoria, the Public Record Office, Cinemedia, and the Victorian Arts Centre Trust through its Performing Arts Museum (см. [267, 268]). В каждом из них был установлен специально разработанный сервер *Z39.50*, поддерживающий профиль *CIMI* и шлюз, поисковый интерфейс которого представлен на Рис.8.14.

Разработка программного обеспечения проекта *Zavier* выполнялась *DSTC Pty Ltd (Distributed Systems Technology Centre)*, вернее его подразделениями *DDU (Distributed Database Unit)* и *RDU (Resource Discovery Unit)*.

В разработку программного обеспечения были также вовлечены: *Crossnet Systems Ltd (Z39.50 toolkit providers)*, *KE Software Pty Ltd (object-oriented database vendors)* and *Vernon Systems Ltd (text-based database vendors)*.

Также следует заметить, что одновременно с проектом *ZAVIER* в Австралии исполнялся другой интересный пилотный проект *ZedMoV* [269] по "скрещиванию" *Z39.50* и *SQL (Z+SQL – [82])* применительно к профилю

Рис. 8.15: Распределенная архитектура *ZedMoV*

CIMI (см. [89,92,270]).

8.4.3 AHDS

Продолжающийся в настоящее время британский проект *AHDS* (*Arts & Humanities Data Service* – <http://ahds.ac.uk>) основан на интеграции различных технологических решений для создания информационных ресурсов и предоставления к ним публичного доступа. Проект *AHDS* финансируется *JISC* (см. раздел 8.3.2) и *AHRD* (*Arts and Humanities Research Board*) [271]. В обрабатываются технологии и стандарты, создаются методические рекомендации в разделах: археология [272–277], история [278,279], прикладные искусства [280,281], языкознание [282] и изобразительные искусства [283,284].

Если на начальных этапах проекта шлюз *AHDS*, физически находившийся в Лондоне, использовал *Z39.50* для запросов к пяти полностью различным базам данных, содержащих информацию относительно археологии (Йорк), истории (Колчестер), прикладных (Глазго) и изобразительные искусства (Ньюкасл), текстовые и лингвистические данные (Оксфорд), то в настоящее время число баз данных существенно возросло, и речь идет не о пяти базах данных, а о пяти координирующих центрах. Создаваемые в проекте базы данных описывают различные типы информации в соответствии с различными стандартами каталогизации. Доступ к данным обеспечивает различное программное обеспечение СУБД на различных аппаратных платформах. Однако, несмотря на эти различия, комбинация *Z39.50* и *Dublin Core* элементов достаточна, чтобы обеспечить сквозной многобазовый поиск.

Проект *AHDS* сегодня имеет несколько дочерних проектов, таких как *ADS* (*Archaeology Data Service*) [285], *OTA* (*The Oxford Text Archive*) [286], *VADS* (*Visual Arts Data Service*) [287] и другие.

AHDS Gateway: Advanced Search

Use this form to conduct more advanced searches of Service Provider catalogues and other resources. To issue a simple keyword search of Service Provider catalogues use the [Simple Search page](#).

Select AHDS Catalogues: Click on the boxes below to include/exclude AHDS catalogues in your search. A tick in the box indicates a catalogue is included.

Archaeology Data Service

History Data Service

Oxford Text Archive

Performing Arts Data Service

Visual Arts Data Service (including ADAM)

Other resources: Additional resources will be available in the future.

Choose a field name to search (e.g. all, subject), and a keyword or expression to search for (e.g. "city", "impressionism", "Dickens"). Use the operators (And, Or, and And Not) to construct more complex queries. See [help](#) for further details.

Field	Keyword	Operator	
All	<input type="text"/>	And	<input type="button" value="Search"/>
All	<input type="text"/>	And	
All	<input type="text"/>	And	

Use this drop-down box to determine how many search results are presented on each results page.

Records to display per page

Рис. 8.16: Поисковый интерфейс AHDS

8.4.4 ARTISTE

В настоящее время музеи и галереи зачастую имеют свои цифровые коллекции, тип которых простирается от обычных изображений, предназначенных для публичного доступа, до специализированных образов, созданных для целей сохранения. Одновременный перекрестный доступ к этим ресурсам признается сегодня очень важным, в том числе и, например, для задач контроля за состоянием картин, которые формируют основную часть культурного наследия Европы. Проект *ARTISTE* (*An integrated Art Analysis and Navigation Environment*) [288] с бюджетом 2,7 млн. евро, частично финансируемый Европейским Союзом через 5-ю рамочную программу, развил систему для автоматической индексации, перекрестного поиска и извлечения художественных изображений с высоким разрешением. В проект вовлечены четыре главных галереи Европы: Галерея Уффици во Флоренции, Национальная Галерея в Лондоне, Виктория и Альберт музей в Лондоне, Лувр во Франции. Суммарно информационный ресурс проекта *ARTISTE* превышает 160 тыс. изображений и 5 млн. текстовых метаописаний.

Проект *ARTISTE* имеет несколько характерных особенностей. В-первых, многоуровневая архитектура информационной системы (см. Рис.8.17 [289]) основана (см. [290]) на новом протоколе *SRW* [17, 18], который является развитием протокола *Z39.50* (упрощенное представление *Z39.50* в терминах *XML* для *SOAP*). Это один из немногих сегодняшних

проектов, основанный на развивающейся технологии *WEB*-сервисов. Более того, в процессе исполнения проекта *ARTISTE* были отмечены недостатки протокола *SRW*, в частности, применяемая в *SRW* система запросов *CQL* [109] (следует напомнить, что в *Z39.50* основной является система запросов *RPN*, а *SQL* и *CQL* являются факультативными) потребовала расширения, в результате чего в *ARTISTE* создана специальная система запросов *CAQL* [291, 292], допускающая работу с операндами, относящимся к изображениям. Во-вторых, на уровне метаданных активно задействованы технологии *XML*, *RDF* и *SOAP*. Наконец, в-третьих, *WEB*-сайт проекта оформлен с применением мультимедиа технологий и выглядит очень привлекательно.

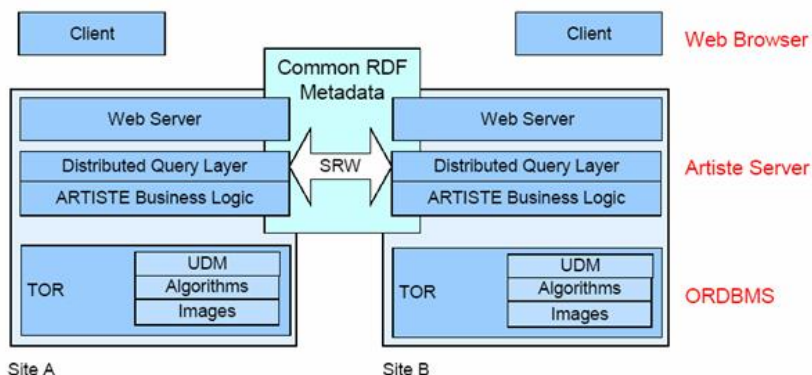


Рис. 8.17: Архитектура системы *ARTISTE*

Многие разработки проекта *ARTISTE* использованы в другом европейском проекте *SCULPTEUR* [293, 294], начавшемся в мае 2002 года и имеющим бюджет 2 млн. евро. Этот проект ориентирован на 3-х мерное представление объектов культурного наследия и интегрируется с многими другими проектами: *DADDI* (*Digital Archive through Direct Digital Imaging*) [295], *ARCO* (*Augmented Representation of Cultural Objects*) [296], *ORPHEUS* [297], *SCHEMA* (*Network of Excellence in Content-Based Semantic Scene Analysis and Information Retrieval*) [298], *VITRA* (*Veridical Imaging of Transmissive and Reflective Artefacts*) [299].

8.5 Проекты по естественным наукам

В последнее время идея создания глобальных распределенных информационных систем по биоразнообразию воистину завладела массами. Существует множество проектов, основанных на различных технологиях. В качестве примеров можно привести *GBIF* (*Global Biodiversity Information Facility*) [300], *CBIF* (*Canadian Biodiversity Information Facility*) [301],

BiOSC (Biological Observations, Specimens and Collections) [302] – доступ через шлюз к 2.4 млн. записей, *REMIB (The World Information Network on Biodiversity)* [303], *DiGIR (Distributed Generic Information Retrieval)* [304], *ITIS (Integrated Taxonomic Information System)* [305], *Species 2000* [306] с дочерними проектами *SPICE (Investigating a federated architecture for creating a catalogue of life)* [307] и *LITCHI* [308], *OBIS (Ocean Biogeographic Information System)* [309, 310], *ABRS (Australian Biological Resources Study)* [311, 312], *SABONET (Southern African Botanical Diversity Network)* [313], европейские проекты *BioCISE (Biological Collection Information Service in Europe)* [314] и *ENHSIN (European Natural History Specimen Information Network)* [315] и многие другие. Даже простое перечисление существующих проектов и информационных систем может составить весьма внушительный список.

Часть перечисленных проектов в своих распределенных информационных системах использует протокол *Z39.50*. Следует, наверное, предполагать, что развитие и внедрение протокола *SOAP/SRW* сблизит используемые и ориентированные на *WEB*-технологии с лучшими качествами *Z39.50*.

В качестве примера ниже рассмотрен проект *Species Analyst* и его дочерние проекты.

8.5.1 *Species Analyst*

Проект *Species Analyst* [316] является исследовательским проектом по разработке стандартов и программного обеспечения в области коллекций по естествознанию и баз данных наблюдений. Головной организацией проекта является Канзаский Университет (*University of Kansas – <http://www.ku.edu>*), Музей Естествознания и Центр исследований по биоразнообразию (*Natural History Museum and Biodiversity Research Center – <http://www.nhm.ku.edu>*).

В настоящее время задействовано более чем 120 коллекций по естествознанию, расположенных по всему миру и обслуживаемых через сеть *Species Analyst*. Во всех этих коллекциях может быть произведен одновременный поиск с объединением возвращаемых результатов в одном из поддерживаемых форматов. Сеть *Species Analyst* в настоящее время поддерживается по протоколу *Z39.50*.

Методы, используемые в *Species Analyst* для работы с распределенными базами данных, основаны на стандарте *ANSI/NISO Z39.50 (ISO-23950)* и *XML*. При этом *Z39.50*, как полагают разработчики проекта, обеспечивает превосходные механизмы для распределенного поиска информации, хотя его использование затруднено сложностью реализации. Все инструментальные средства, используемые *Species Analyst* преобразовывают наборы результатов запросов *Z39.50* в формат *XML* для дальнейшей обработки или просмотра. Подход, основанный на применении общепринятых стандартов, является осознанной необходимостью проекта.

В проекте развит клиент *Z39.50* (известный как *ZX* [317]), который является блоком управления *Z39.50* для *Internet Explorer* версии 5.0 и выше.

Этот клиент обеспечивает как поиск по *Z39.50*, так и поиск, определенный по отдельному *URL*.

В проекте разработаны и другие продукты:

ZASP – набор страниц *ASP*, которые предлагают альтернативу *Z39.50* серверу для того, чтобы обеспечить доступ к дынным по естествознанию и результатам наблюдения. Страницы генерируют вывод *XML* в формате, ожидаемом клиентами *Species Analyst*.

ZPortal – набор серверных *ASP* для универсального *HTTP-Z39.50* шлюза, который обеспечивает доступ к *Z39.50* ресурсам и формирует результаты в *XML* документе в том же формате, что и *ZX*.

ZBigServer является *Z39.50* сервером, предназначенным для обеспечения доступа к коллекциям естествознания. Программное обеспечение выполняется как служба *Windows NT/2000*.

Также в проекте *Species Analyst* был создан набор метаданных *Darwin Core – DwC* [318] (ср. *Dublin Core, DC*), определяющий структуру записей метаданных и точки доступа при поиске.

Одна примечательная особенность программного обеспечения *Species Analyst* – способность возвращать данные из множественных баз данных в виде отдельной электронной таблица *Microsoft Excel*.

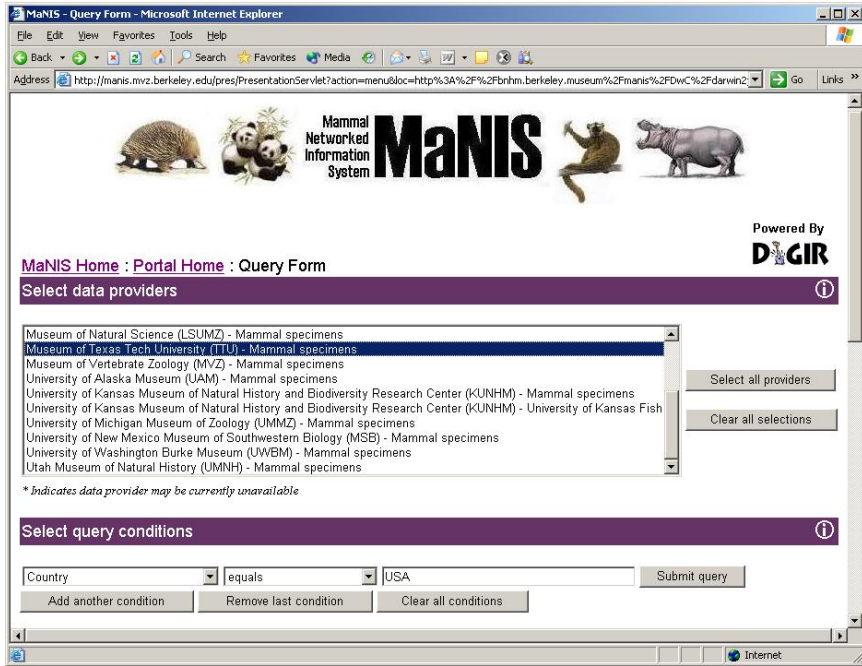
Проект *Species Analyst* явился основой для других проектов, которые рассмотрены ниже.

8.5.2 *FishNet*

Проект *FishNet* [319] является распределенной информационной системой, которая использует *Z39.50* и *XML* протоколы для "беспроводной" связи записей баз данных музеев и других учреждений в единой информационно-поисковой системе. Используя *FishNet* и *Species Analyst (TSA)*, любой пользователь с доступом к Internet может получить доступ к коллекциям партнеров *FishNet* и получить информацию о рыбах, описания которых находятся в базах данных.

FishNet отличается от более централизованных систем тем, что информация в *FishNet* распределена по множеству партнеров (см. [320]). Это означает, что базы данных расположены у партнеров проекта и сохраняют полное управление "на местах". Пользователи могут загрузить части баз данных в разнообразных форматах и анализировать эту информацию в соответствии с их собственными потребностями.

FishNet – является расширением системы *Species Analyst* в части информации о мире рыб. Количество партнеров проекта *FishNet* сегодня составляет 24. Двенадцать из них в настоящее время связаны интерактивно, и обеспечивают доступ к 20 миллионам записей. Другие партнеры будут добавляться. Дата окончания первого этапа проекта была намечена на декабрь 2002 года.

Рис. 8.18: Поисковый интерфейс *MaNIS*

8.5.3 *MaNIS*

MaNIS [321] (*Mammal Networked Information System*) – проект, поддерживаемый *NSF* (*National Science Foundation*). В этом проекте семнадцать североамериканских организаций участвуют в разработке распределенной информационной системы по млекопитающим. Цели этого проекта:

1. облегчить открытый доступ к комбинированным данным из WEB-браузера;
2. расширить предметные коллекции;
3. сохранить созданные информационные ресурсы;
4. использовать парадигму дизайна, которая могла бы быть применена в смежных областях.

Развитие этой сетевой информационной системы *MaNIS* связано с музеями естествознания в целях их объединения, формирования и совместной поддержки инфраструктуры открытым способом. Предполагается, что объединенный банк данных по биологической вариативности позволит использовать его для прогноза эволюционных процессов и экологических явлений, которые не были очевидны прежде. При этом доступ к детальной информации о биологической вариативности Земли является актуальным,

поскольку человечество сталкивается с проблемами изменения окружающей среды 21-ого столетия.

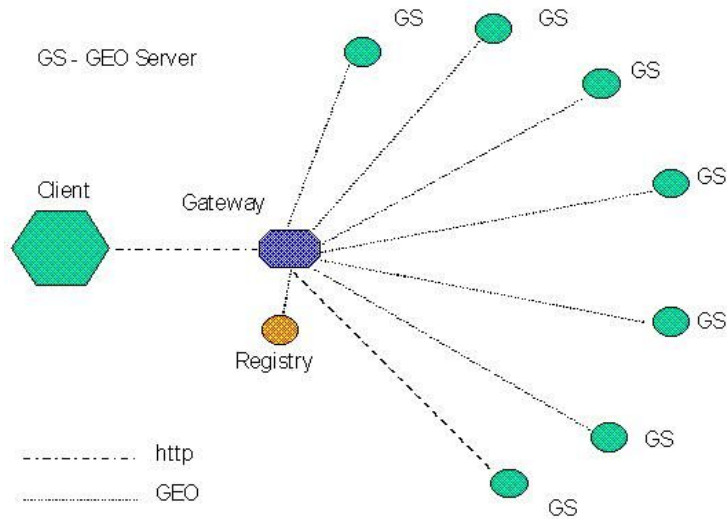
8.5.4 *HerpNET*

HerpNET [322] является совместным проектом организаций США, Канады и Мексики по интеграции ресурсов по пресмыкающимся. Проект поддерживается *NSF* (*NSF* No. 0132303) и объединяет 36 организаций.. Масштаб вызовов биологической вариативности требует интегральный подход, но индивидуальные совокупности обычно неадекватны для таких исследований. Миссия *HerpNET* состоит в том, чтобы интегрировать накопленные знания в музеях, сделать их доступными для науки и общества, создавая распределенную базу данных с доступом из различных порталов. *HerpNET* объединит большие массивы данных с малым количеством коллекций, которые имеют региональную специфику. В этом смысле *HerpNET* дает свой вклад в информационную систему по биоразнообразию, также как и другие аналогичные проекты: *MaNIS* [321], *FishNet* [319], *MaPSTeDi* [323].

8.6 Проекты по ГИС

Несмотря на достаточно большой успех технологий, основанных на протоколе *Z39.50*, в области построения распределенных информационных систем, интегрирующих метаданные библиотек, архивов, музеев и других, на наш взгляд, истинным бенефисом технологий *Z39.50* следует считать внедрение его как основы в геоинформационные системы. При этом речь идет о создании глобальных распределенных информационных систем по данным спутниковых наблюдений, архивам аэрокосмических снимков, цифровым картам и другой информации, имеющей пространственную привязку. Данные, интегрируемые в подобные системы, имеют свою специфику, которая кроме всего прочего заключается, во-первых, в очень больших объемах первичной информации, во-вторых, в необходимости составления очень подробных и очень структурированных метаописаний для даже грубого представления первичного объекта хранения, наконец, в-третьих, для доступа к подобным информационным системам требуются специфические пользовательские интерфейсы, включающие графические элементы пространственной навигации при поиске и визуального отображения результатов поиска на картах. Если дополнить все это потребностью, как правило, в жесткой абсолютной временной привязке (иногда до миллисекунд), мы получим грубое представление о тех рамках, в которых создаются подобные информационные системы.

Для информационных систем, которые интегрируют данные с пространственной привязкой, в рамках *Z39.50* был создан специальный прикладной профиль *GEO* [60], включающий набор поисковых атрибутов *GEO*, набор меток *GEO*, схему данных *GEO* и т.д. Профиль *GEO* привязан к стандарту метаданных *CSDGM* (*Content Standard for Digital Geospatial Metadata*) [59] Федерального Комитета по Географическим Данным США

Рис. 8.19: Схема *Clearinghouse*

(*FGDC*) [55]. На основании профиля *GEO* и протокола *Z39.50* была создана распределенная информационная система по пространственным метаданным *Clearinghouse* [324], включающая сотни серверов *Z39.50* (см. [325]).

Несмотря на достаточно простую архитектуру, реализующую на основе шлюзов *Z39.50-HTTP* трехзвенную модель, и большое количество участников, система *Clearinghouse* впечатляет объемами информации и простотой, с которой к этой информации можно получить доступ. Отличительной особенностью системы является постоянный мониторинг всех ассоциированных серверов с предоставлением пользователю информации об их доступности в различные промежутки времени [326]. Система оказалась настолько привлекательной, что по технологии *FGDC Clearinghouse* сегодня функционируют подобные системы для других предметных областей:

SEDAC (*Socioeconomic Data and Applications Center*) [327] – шлюз, обслуживаемый ***DAAC*** (*Distributed Active Archive Center*) ***NASA EOSDIS*** (*Earth Observing System and Data and Information System*), включает более 40 серверов *Z39.50* международных, федеральных и региональных организаций, предоставляющих доступ к базам метаданных в стандарте *FGDC CSDGM* [167].

GCMD (*Global Change Master Directory*) [328] – в отличие *Clearinghouse* предоставляет доступ к централизованному каталогу, интегрирующий ресурсы более чем 1500 различных организаций, поставляющих метаданные. Предназначен в основном для научного сообщества в области наук о Земле.

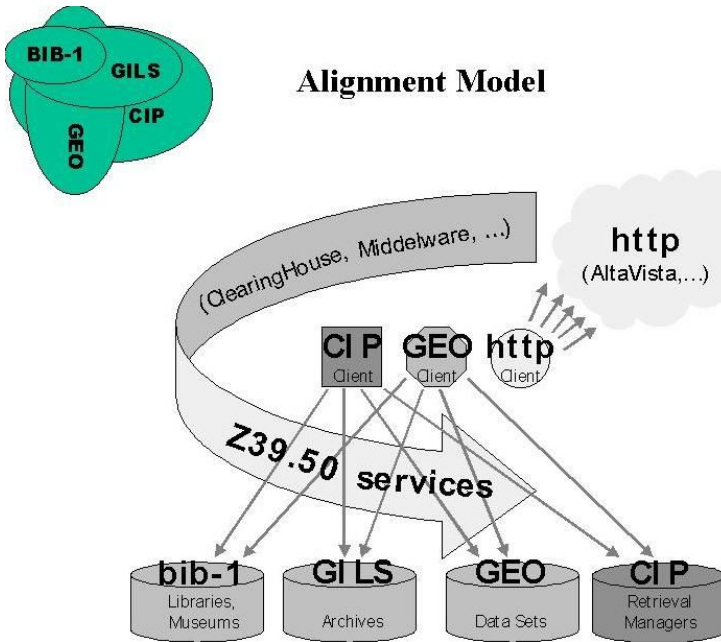


Рис. 8.20: Взаимодействие CIP и Z39.50 систем

NBII Metadata Clearinghouse [329] (*NBII – National Biological Information Infrastructure* [330]) является инициативой *USGS (U.S. Geological Survey)* [331] по созданию распределенной информационной системы, которая содержит метаданные, описывающие биологические данные и информационные продукты, и основана на подмножестве *CSDGM – NBII Biological Profile* [58, 332].

IAI-DIS (Inter-American Institute for Global Change Research – Data and Information System) [333] – американская распределенная информационная система по Глобальным Изменениям. Состоит из множества узлов, доступных по протоколу *Z39.50*. Существуют узлы и в других странах, например, Бразилии и Уругвае.

Сознавая привлекательность технологий *Z39.50* для построения РИС, с одной стороны, и отсутствие в *Z39.50* необходимых компонент для функционирования РИС, с другой стороны, *CEOS (Committee on Earth Observation Satellites)* [334] разработал специальный протокол для построения своей распределенной информационной системы – *CIP (Catalogue Interoperability Protocol)* [25, 26] как модернизированный протокол *Z39.50*. Система предназначена для доступа к данным по дистанционному зондированию Земли. *CIP* наследует все свойства *Z39.50*, оба протокола могут использоваться совместно.

В отличие от базового протокола *Z39.50*, протокол *CIP* ориентирован на обработку распределенной информации, хранящейся на различных сер-

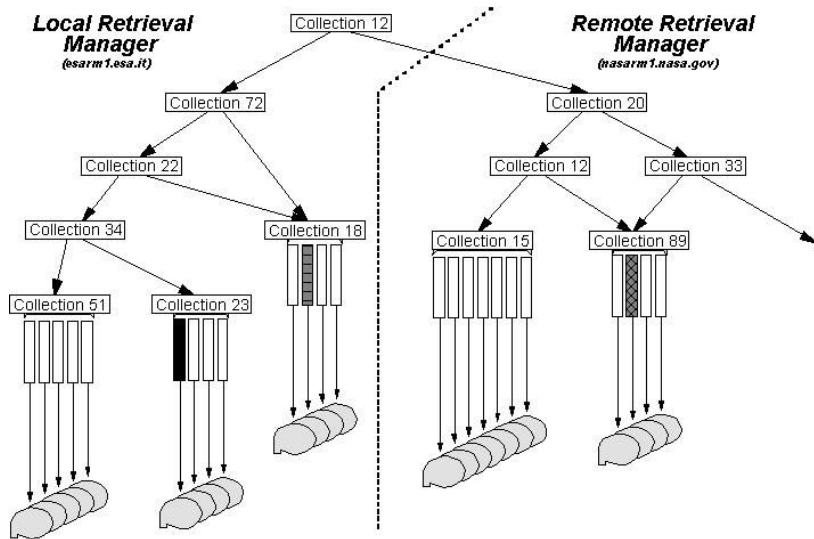


Рис. 8.21: Иерархические коллекции CIP

верах и организованной в иерархические коллекции и базы данных. При этом коллекция может быть распределенной и находится на разных серверах CIP. Доступ к коллекциям обеспечивают соответствующие серверы, именуемые *Retrieval Manager*.

Протокол CIP используется многими информационными системами, связанными космическим мониторингом. В частности, можно указать системы *INFEO* (*Information on Earth Observation*) [335] (см. также *Earth Observation Portal* [336]), *DIMS* (*Data Information and Management System*) [337] и др. (см. [338]). Следует отметить, что соответствующие организации в России также пытаются включиться в процесс создания глобальных библиотек данных по космическому мониторингу. В качестве примера можно указать Институт космических исследований РАН (ИКИ РАН) [339].

Заключение

Если Вы дочитали этот труд до этого места, то мы надеемся, что основные выводы Вы уже сделали сами. Однако, хотелось бы подчеркнуть несколько моментов, ради которых, собственно, эта книга и создавалась.

Во-первых, технологии построения распределенных информационных систем не исчерпываются технологиями *Z39.50*. Существует множество других технологий, которые не исключают, а дополняют те, которые обсуждались выше. Более того, в последнее время в мировом сообществе существует и продолжает появляться множество проектов, в которых тесно переплетаются различные технологические подходы и идеологии. И это вполне закономерно, т.к. развитие информационных технологий не стоит на месте, человеческая мысль ищет оптимальные решения для организации доступа к разнородным информационным ресурсам и упрощения пользовательских интерфейсов. Поэтому ставка на какую-то единственную технологию, с помощью которой можно было бы решить все проблемы, заведомо обречена на провал. Технологии приходят и уходят, а информация остается, и остаются задачи по ее обработке, которые требуют решения.

Во-вторых, ни одна из существующих технологий построения распределенных информационных систем не стоит на месте. Развиваются все, даже те, которые зафиксированы жесткими стандартами. Так обсуждаемый выше стандарт *Z39.50 (ISO-23950)* постоянно дополняется новыми возможностями, готовится выход новой версии стандарта.

Наконец, появляются новые технологии. Чаще всего факт появления технологии с новым именем является маркетинговым шагом той или иной компании. Далеко не всегда заявляемые новые технологии являются действительно новыми. Чаще всего новое – это, или хорошо забытое старое, или старое, которое подверглось хорошей переработке в свете новых подходов. Во всяком случае, новое всегда опирается на старое. Яркий пример этому – *ZING (Z39.50 International: Next Generation)* [17], интегрирующий технологии *Z39.50*, *WEB*-сервисов (*SOAP*) и *XML*.

Поэтому мы думаем, что изложенный в этой книге подход к решению определенных задач будет актуальным не только в рамках сегодняшнего *Z39.50*, но и в будущих технологиях, которые придут ему на смену.

Мы надеемся, что чтение этой книги для Вас не явилось пустой тратой времени.

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. Moen William E. The Development of ANSI/NISO Z39.50: A Case Study in Standards Evolution // A Dissertation, School of Information Studies Syracuse University, August 1998.
(<http://www.unt.edu/wmoen/dissertation/DissertationIndex.htm>).
3. Abstract Syntax Notation One (ASN.1): Specification of Basic Notation // International Organization for Standardization, Information Technology (ISO/IEC) 8824-1. - 1995.
4. Larmouth J. ASN.1 Complete, Morgan Kaufmann Publishers. - 1999. - 400 p. - ISBN: 0-12-233435-3.
5. International Organization for Standardization, Information Technology: ASN.1 Encoding Rules: Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER), ISO/IEC 8825-1. - 1995.
6. Z39.50 Target Statistics
(<http://bagel.indexdata.dk/targettest/stat.php>).
7. Жижимов О.Л. Введение в Z39.50: 4-е изд. доп. и перераб. - Новосибирск: Изд-во НГОНБ. - 2003. - 263 с.
8. Жижимов О.Л., Мазов Н.А. Модель распределенной информационной системы Сибирского Отделения РАН на базе протокола Z39.50 // Электронные библиотеки: научный электронный журнал. - М., Институт развития информационного общества. - 1999. - 2(2).
9. Когаловский М.Р. Энциклопедия технологий баз данных. - М.: Финансы и статистика. - 2002. - 800 с.
10. Цимбал А.А., Аншина М.Л. Технологии создания распределенных систем. Для профессионалов. - СПб.: Питер. - 2003. - 576 с.

11. Laverdure, L., Collonna-Romano, J., Srite, P. Network Application Support Architecture Reference Manual. Digital Press. - 1993.
12. Microsoft Corporation. Справочник по Microsoft OLE DB 1.1 / Пер. с англ. - М.: Издательский отдел "Русская Редакция" ТОО "Channel Trading Ltd.". - 1997. - 624 с.
13. Галатенко В., Таранов А. Компонентная объектная модель JavaBeans // СУБД. - № 4. - 1997. - С. 42-60.
14. Пуха Ю. CORBA/ПООР и Java RMI. Основные возможности в сравнении // СУБД. - № 4. - 1997. - С. 24-36.
15. Coyle Frank P. XML, Web Services, and the Data Revolution // Addison Wesley Professional. - 2002. - 400 p.
16. Михайлов В. Web-сервисы, WSDL и SOAP // Технология Клиент-Сервер. - 2002. - № 2.
17. ZING - Z39.50-International: Next Generation.
(<http://www.loc.gov/z3950/agency/zing>).
18. SRW: Search/Retrieve Web Service (<http://srw.o-r-g.org/docs>).
19. Approach Paper for GILS support in MSFindit // Submitted to USGS (United States Geological Survey). - Satyam Computer Services Limited. - December 2001. - (<http://www.gils.net/ms-findit.doc>).
20. Colmer M., Brittain J.M. Access to Information: New Directions // International Journal of Information Management. - 1998. - Vol. 18. - № 2. - P. 81-90.
21. Poo D.C.C., Toh T.-K., Khoo C.S.G. Design and implementation of the E-referencer // Data & Knowledge Engineering. - 2000. - Vol.32. - № 2. - P. 199-218.
22. Hands J., Bessonov M., Blinov M., Patel A., Smith R. An inclusive and extensible architecture for electronic brokerage // Decision Support Systems. - 2000. - Vol. 29. - № 4. - P. 305-321.
23. Chang Y.-S., Ho M.-H., Yuan S-M. A unified interface for integrating information retrieval // Computer Standards & Interfaces. - 2001. - № 23. - P. 325-340.
24. Lo W., Chang Y.S., Chou C.L., Sheu R.K., Yuan S.M. An Information Store and Retrieval Facility on CORBA // Lecture Notes in Computer Science LNCS of Springer-Verlag. - Vol. 1846. - Heidelberg:Springer-Verlag. - Germany. - 2000. - P. 374-379.
25. Catalogue Interoperability Protocol (CIP) Specification - Release B 2.4 - CEOS/WGISS/PTT
(http://www.dfd.dlr.de/ftp/pub/CIP_documents/cip2.4/S_cover.pdf).

26. Best C., Hayes L., Nebert D., Reich L., Percival G., Hylton J., Smith S. Alignment of CIP and GEO Profiles // Proceedings of the Earth Observation & Geo-Spatial Web and Internet Workshop '98 (EOGEO'98). - Salzburger Geographische Materialien. - Vol. 27. - Instituts fur Geographie der Universitdt Salzburg. - 1998.
27. Lynch C.A. Building the infrastructure of resource sharing: Union catalogs, distributed search, and cross-database linkage // Library Trends. - Vol. 3. - № 45. - 1997. - P. 448-461.
28. Coyle K. The Virtual Union Catalog: A Comparative Study // D-Lib Magazine. - 2000. - Vol. 6. - № 3. (<http://www.dlib.org/dlib/march00/coyle/03coyle.html>).
29. Son J. H., Kim M. H. An analysis of the optimal number of servers in distributed client/server environments // Decision Support Systems, 2004. - № 36. - P. 297-312.
30. Жижимов О.Л., Скибин С.В. О времени обработки запросов в распределенных информационных системах на основе протокола Z39.50 // Пятая Всероссийская науч. конф. "Электронные библиотеки: перспективные методы и технологии, электронные коллекции - "RCDL-2003", г. Санкт-Петербург, 29-31 окт., 2003 г.: Труды конф. - СПб., НИИ Химии СПбГУ. - 2003. - С. 73-77.
31. Российские Z39.50 серверы: статистика. (<http://z3950.ruslan.ru/stat/stat.shtml>).
32. Lukas C., Roszkowski M. The Isaac Network: LDAP and Distributed Meta-data for Resource Discovery // Proceedings of the Third IEEE META-DATA Conference, April 6-7, 1999, Bethesda, Maryland. (<http://scout.wisc.edu/Projects/PastProjects/Isaac/ldap.html>).
33. Roszkowski M., Lukas C. A Distributed Architecture for Resource Discovery Using Metadata // D-Lib Magazine, June 1998. - ISSN 1082-9873. (<http://www.dlib.org/dlib/june98/scout/06roszkowski.html>).
34. The PRIDE project team / Oiling the Works: the PRIDE Project Develops an Information Brokerage Service // Exploit Interactive. - Issue 1. - 10 April 1999 (<http://www.exploit-lib.org/issue1/pride>).
35. PRIDE Architecture and Interfaces / PRIDE Logical Components and Integration Designs // Work Package 2 of Telematics for Libraries project PRIDE (LB 5624). - 1999. (<http://www.ukoln.ac.uk/metadata/pride/wp2/d241/>).
36. Jones S., Cunningham Sally J., McNab R., Boddie S. A transaction log analysis of a digital library // Int. J. Digit. Libr. - 2000. - № 3. - P. 152-169.

37. Sfakakis M., Kapidakis S. User Behavior Tendencies on Data Collections in a Digital Library // Lecture Notes in Computer Science. - 2002. - № 2458. - P. 550-559.
38. ZooPARK. Модульный сервер Z39.50. Версия 4.06 // Новосибирск, 2004. (<http://z3950.uiggm.nsc.ru:210/ZooPARK>).
39. Patel A. Access control mechanisms in digital library services // Computer Standards & Interfaces. - 2001. - Vol. 23. - № 1. - P. 19-28.
40. International Organization for Standardization. Documentation: format for bibliographic information interchange on magnetic tape. [2nd ed.] Geneva, ISO, 1981 (ISO 2709-1981). The first edition was published in 1973.
41. Формат USMARC. Краткое описание: В 3-х ч. М.: ГПНТБ России. - 1996.
42. Руководство по UNIMARC: Руководство по применению международного коммуникативного формата UNIMARC. - М.: ГПНТБ России. - 1992. - 320 с.
43. AACR - "Anglo-American Cataloguing Rules, Second Edition", Revision, 1988.
44. Российский коммуникативный формат представления библиографических записей в машиночитаемой форме: (Рос. вариант UNIMARC). СПб.: Изд-во РНБ. - 1998.
45. ISBD, International Federation of Library Associations and Institutions, "ISBD(G): General International Standard Bibliographic Description: Annotated Text", London, IFLA International Office for UBC, 1. 1997.
46. ГОСТ-7.1-84 "Библиографическое описание документа". - 1986.
47. Application Profile For The Government Information Locator Service (GILS), Version 2, November 24, 1997. (http://www.gils.net/prof_v2.html).
48. Hutto Dena H. Old Solutions in a New Age: Cataloging and the Future of Access to Government Information // Journal of Government Information. - 1996. - Vol. 23. - № 3. - P. 335-344.
49. Adams M.O., Thibodeau S.G. The Government Information Locator Service: Origins and Potential // Journal of Government Information. - 1996. - Vol. 23. - № 5-6. - P. 747-748.
50. Mullen A. GILS metadata initiatives at the state level // Govt. Inform. Quart., - 2001. - Vol. 3. - № 18. - P. 167-180.
51. GILS (<http://www.gils.net>).

52. Stuart L. Weibel, Lagoze C. An element set to support resource discovery. The state of the Dublin Core: January 1997 // International Journal on Digital Libraries. - 1997. - Vol. 1. - № 2. - P. 176-186.
53. The Z39.50 Cross-Domain Attribute Set. Version 1.4 (04.08.1999) (<http://www.oclc.org/levan/docs/crossdomainattributeset.html>).
54. Z39.50 Attribute Architecture. Version 1.1 (09.08.1999) (<http://lcweb.loc.gov/z3950/agency/attrarch/arch.html>).
55. FGDC - Federal Geographic Data Committee (<http://www.fgdc.gov>).
56. David L. Tulloch, Milo Robinson. A progress report on a U.S. National Survey of Geospatial Framework Data // Journal of Government Information. - 2000. - № 27. - P. 285-298.
57. Communication Protocols for a Distributed Geospatial Metadata Service. 3rd Report // National Geospatial Data Framework, 1998. (www.gigateway.org.uk/moreinformation/pdf/archive/-NGDF%20Progress%20Report.pdf).
58. Content Standard for Digital Geospatial Metadata. Part 1: Biological Data Profile // Biological Data Working Group. FGDC and USGS Biological Resources Division, October 1999. (<http://www.fgdc.gov/standards/documents/standards/biodata/biodatap.html>).
59. Content Standard for Digital Geospatial Metadata (<http://www.fgdc.gov/metadata/contstan.html>).
60. Douglas D. Nebert. Z39.50 Application Profile for Geospatial Metadata or 'GEO' Version 2.2. U.S. Federal Geographic Data Committee. (<http://www.blueangeltech.com/Standards/GeoProfile/geo22.htm>).
61. Crosswalk: FGDC Content Standards for Digital Geospatial Metadata to USMARC // Geography and Map Division, Library of Congress. - 1997. (<http://alexandria.sdc.ucsb.edu/public-documents/metadata/fgdc2marc.html>).
62. Crosswalk: USMARC to FGDC Content Standards for Digital Geospatial Metadata // Geography and Map Division, Library of Congress. - 1997. (<http://alexandria.sdc.ucsb.edu/public-documents/metadata/marc2fgdc.html>).
63. Z39.50 Profile for Access to Digital Collections. (Final Draft). May 3, 1996. - (<http://lcweb.loc.gov/z3950/agency/profiles/collections.html>).
64. Bull R. A. Analysis of the Z39.50 Profile for Access to Digital Collections and the Z39.50 Explain Service for UKOLN // Crossnet Job No. F17. - 29 October 1997. - CSL017-TN-001. - Issue 2. (<http://www.ukoln.ac.uk/metadata/cld/study/crossnet/zpadc.pdf>).

65. Moen William E. Resource Discovery Using Z39.50: Promise and Reality // Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium. - 2001. - 574 p.
(http://lcweb.loc.gov/catdir/bibcontrol/moen_paper.html).
66. The CIMI Profile. Release 1.0H. A Z39.50 Profile for Cultural Heritage Information (http://www.cimi.org/public_docs/HarmonizedProfile/-HarmonProfile1.htm).
67. Moen William E. Accessing Distributed Cultural Heritage Information // Comm. of ACM, - April 1998. - Vol. 41. - № 4. - P. 45-48.
68. Moen William E. CIMI's Z39.50 Interoperability Testbed: Search and Retrieval of Distributed Cultural Heritage Information. 1998.
(http://www.unt.edu/wmoen/publications/CIMI_ACMrev.htm).
69. Velegrakis Y., Christophides V., Constantopoulos P. On Z39.50 wrapping and description logics // Int. J. Digit. Libr. - 2000. - № 3. - P. 208-220.
70. Жижимов О.Л., Мазов Н.А. О доступе к информационным ресурсам по культурному наследию по протоколу Z39.50 // Электронные библиотеки: научный электронный журнал, М., Институт развития информационного общества. - 2001. - Т. 4. - № 4.
(<http://www.iis.ru/el-bib/2001/200104/ZM/ZM.ru.html>).
71. Bechhofer S., Goble C. Thesaurus construction through knowledge representation // Data & Knowledge Engineering, - 2001. - Vol. 37. - № 1. - (April 2001). - P. 25-45.
72. Binding C., Tudhope D. KOS at your Service: Programmatic Access to Knowledge Organisation Systems // Journal of Digital Information, Vol. 4. - № 4, Article No. 265, 2004-02-05.
(<http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Binding/>).
73. Amann B., Fundulaki I., Scholl M. Integrating ontologies and thesauri for RDF schema creation and metadata querying // Int. J. Digit. Libr., - 2000. - № 3. P. 221-236.
74. Taylor M. Zthes: a Z39.50 Profile for Thesaurus Navigation. 15th November 2000.
<http://lcweb.loc.gov/z3950/agency/profiles/zthes-04.html>.
75. International Organization for Standardization. ISO 2788: Guidelines for the establishment and development of monolingual thesauri, 2nd ed. Geneva: ISO. - 1986.
76. Жижимов О.Л., Мазов Н.А. Тезаурусы и классификационные схемы в распределенных информационных системах: проблемы и решения // Междунар. конф. "EVA-2000". Электронная конвергенция: новые технологии в музеях, галереях, библиотеках и

- архивах, Москва, 30 октября - 3 ноября 2000 г.: Матер. конф. - М., Центр ПИК Минкультуры РФ, 2000. - С. 08-9-1 - 08-9-3
(<http://www.artinfo.ru/eva/EVA2000M/eva-papers/200008/Zhizhimov2-R.htm>).
77. Жижимов О.Л., Мазов Н.А. Доступ к тезаурусам и классификационным схемам из Интернет по протоколу Z39.50 // Вычислительные технологии. - 2000. - Спец. вып. - С. 23-28.
 78. Мазов Н.А., Жижимов О.Л. О доступе к рубрикам по протоколу Z39.50 при поиске в базах данных (на примере рубрикатов ВИНТИ и ГРНТИ) // 6-я междунар. конф. Информационное общество. Интеллектуальная обработка информации. Информационные технологии, Москва, 16-18 окт., 2002 г.: Материалы конф. - М., ВИНТИ. - 2002. - С. 229-232.
 79. Мазов Н.А., Жижимов О.Л. Применение протокола Z39.50 для работы с тезаурусами и классификационными схемами // 7-я междунар. конф. "Крым 2000". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 3-11 июня 2000 г.: Труды конф. - Симферополь, Таврида. - 2000. - Т. 1. - С. 402-405.
 80. Мазов Н.А., Жижимов О.Л. Унификация построения и организации доступа к тезаурусам и классификационным схемам в распределенных информационных системах по протоколу Z39.50 // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Вторая Всерос. науч. конф., Протвино, 26-28 сентября 2000 г.: Сб. докл. - Протвино, ГНЦ ИФВЗ, 2000. - С. 230-233.
 81. Z39.50 Utility Attribute Set / Modified March 2003.
(<http://lcweb.loc.gov/z3950/agency/attrarch/util.html>).
 82. Z+SQL Profile. February 2000.
http://archive.dstc.edu.au/DDU/projects/Z3950/Z+SQL/-Z+SQL_profile.html.
 83. ISO/IEC 9075-1:1999, Information technology - Database language - SQL - Part 1: Framework (SQL/Framework), 1999. - 67 p.
 84. ISO/IEC 9075-2:1999, Information technology - Database language - SQL - Part 2: Foundation (SQL/Foundation), 1999. - 1121 p.
 85. ISO/IEC 9075-3:1999, Information technology - Database language - SQL - Part 3: Call-Level Interface (SQL/CLI), 1999. - 401 p.
 86. ISO/IEC 9075-4:1999, Information technology - Database language - SQL - Part 4: Persistent Stored Modules (SQL/PSM), 1999. - 152 p.
 87. ISO/IEC 9075-5:1999, Information technology - Database language - SQL - Part 5: Host Language Bindings (SQL/Bindings), 1999. - 243 p.

88. Li H., Orłowska M.E. Liu C. A Normal Form on Functional Dependencies in Object-Oriented Data Models // Proceedings for 10th Australian Database Conference (ADC 99), Auckland, 18-22 January 1999.
89. Finnigan S., Bird L. Harnessing cultural resources: the Web, Z39.50 and the CIMI profile // Proceedings for CIDOC 98 at International Council of Museums Conference (ICOM98), Melbourne, 10-16 October 1998 (<http://www.willpowerinfo.myby.co.uk/cidoc/cidoc98/finnigan.doc>).
90. Wing H., Mineau G., Colomb R.M. Using CG Formal Contexts to Support Business System Interoperations // 6th International Conference on Conceptual Structures (ICCS'98), Montpellier, 10-13 August 1998.
91. Finnigan S. and Ward N. Tasmanian Cultural Information Access Framework // Consultancy Report for the Tasmanian Department of Education, Community and Cultural Development, April 1998.
92. Finnigan S., Bird L. and Bull R. Z+SQL for Advanced Interoperable Search Services for Geo-Spatial Collections // Proceedings for EOGeo'98 and IDRISI GIS'97 (CD-ROM), Salzburg, 17-19 January 1998.
93. Жижимов О.Л. Z+SQL: достоинства и недостатки на примере реализации // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Вторая Всероссийская научная конференция, Протвино, 26-28 сентября 2000 г.: Сб. докл. - Протвино, ГНЦ ИФВЗ, 2000. - С. 211-213.
94. Жижимов О.Л., Скибин С.В. О доступе к реляционным СУБД по протоколу Z39.50 // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Третья Всероссийская конференция по Электронным Библиотекам - "RCDL-2001", г. Петрозаводск, 11-13 сентября 2001 г.: Сб. тр. - Петрозаводск, Карельский научный центр. - 2001. - С. 123-128.
95. Егоров В.К., Вислый А.И. Многообразие и конкретность в создании информационно-библиотечной компьютерной сети страны // Библиотечные компьютерные сети: Россия и Запад. Современные тенденции корпоративной работы библиотек в сетях передачи данных. - М., 1998. - С. 107-109.
96. Каленов Н.Е. Библиографические ресурсы библиотек РАН и типовые средства для их представления в электронном виде // Инф. рес. России. - 1998. - № 6. - С. 26-27.
97. Шрайберг Я.Л. Современное состояние развития библиотечных технологий и место CDS/ISIS // Научные и технические библиотеки. -1995. - Спец. выпуск. - С. 5-8.

98. Шрайберг Я.Л., Воройский Ф.С. Автоматизированные библиотечно-информационные системы. - М.: Либерия, ГПНТБ России, 1996. - 273 с.
99. Эпштейн М.Я. Использование современных информационно-поисковых систем // Инф. рес. России. - 1998. - № 1. - С. 28-33.
100. Воройский Ф.С., Грибов В.Т., Каленов Н.Е. Рынок программного обеспечения для автоматизированных библиотечно-информационных систем // Инф. рес. России. - 1998. - № 5. - С. 18-22.
101. Племнек А.И., Усманов Р.Т. Z39.50: Открытый доступ к библиографической информации // Научные и технические библиотеки. - 1998. - № 8. - С. 24-28.
(http://www.unilib.neva.ru/rus/olsc/publications/z39_04.html).
102. Племнек А.И., Усманов Р.Т., Сова Д.Н. Использование протоколов Z39.50 и HTTP в современных библиотечных информационных системах // Информационный бюллетень РБА. - 1998. - № 12. - С. 240-257.
(http://www.unilib.neva.ru/rus/olsc/publications/z39_01.html).
103. Baranov V., Plemnek A., Sokolova N., et al. Review of Z39.50 servers and Z39.50 environment in Russia // Library Hi Tech. - 2000. - Vol. 18. - № 4. - P. 304-314.
104. Племнек А.И., Соколова Н.В. Центр "Открытые Библиотечные Системы": опыт разработки корпоративной библиотечной системы регионального уровня в системах // Информационный бюллетень РБА. - 1999. - № 15.
(http://www.unilib.neva.ru/rus/olsc/publications/corp_exp.html).
105. Попов В. В., Грибов В. Т., Левова Л. В. Еще раз про АИБС "МАРК- SQL" // 10-я Международная конференция "Крым 2003". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 7-15 июня, 2003 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2003. - С. 129-131.
106. Жижимов О.Л., Мазов Н.А., Поморцев А.Ю. Доступ к базам данных CDS/ISIS/M из INTERNET // Автоматизированные библиотечно-информационные системы: VI Сиб. науч. семинар с междунар. участием, г. Новосибирск, 1-7 июля 1996 г.: Тез. докл. и сообщ. - Новосибирск. - ГПНТБ СО РАН. - 1996. - С. 32-36.
107. Мазов Н.А., Жижимов О.Л., Поморцев А.Ю. Организация доступа к базам данных CDS/ISIS/M из INTERNET // 3-я Междунар. конф. "Крым 96". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, п. Форос - г.

- Ялта, 1-9 июня 1996 г.: Материалы конференции. - М. - 1996. - С. 177-179.
108. Жижимов О.Л., Мазов Н.А., Фролов А.С. Доступ к базам данных ISIS из Internet и построение распределенной информационной системы // Вычислительные технологии. - 1997. - 2. - № 3. - С. 45-50.
109. CQL - Common Query Language (<http://www.loc.gov/z3950/agency/zing/cql/>).
110. D3ZServer (<http://d3zsoft.newmail.ru>).
111. БКС – Библиотечная компьютерная сеть (<http://www.bks-mgu.ru/products>).
112. A-Elite - ЗАО 'GIPER' (<http://www.giper.ru/abis.htm>).
113. АС Библиотека-3 – ГИВЦ МК РФ (<http://www.givc.ru/info/work/1.html>).
114. Hammer S., Dickmeiss F., Levanto H., Taylor M. Zebra - User's Guide and Reference / Index Data, 1995-2004 (<http://www.indexdata.dk/zebra/doc>).
115. Елепов Б.С. Основные принципы создания автоматизированной системы научно-технической информации региона на базе Сибирского отделения АН СССР: Автореф. дисс. ... докт. техн. наук. - Новосибирск, 1990. - 67 с.
116. Кириллова О.В., Мазов Н.А. Современные технологии в библиотеке академического института // 2-я Международная конференция "Крым 95". Библиотеки и ассоциации в меняющемся мире: Новые технологии и новые формы сотрудничества, г. Евпатория, 10-18 июня, 1995 г.: Материалы конференции. - М., 1995. - Т. 1. - С. 218-220.
117. Кириллова О.В., Мазов Н.А. Организация и функционирование информационно-библиотечного центра по наукам о Земле // Автоматизированные библиотечно-информационные системы: VI Сиб. науч. семинар с междунар. участием, 1-7 июля 1996 г., г. Новосибирск: Тез. докл. и сообщ. - Новосибирск, ГПНТБ СО РАН. - 1996. - С. 115-117.
118. Кириллова О.В., Мазов Н.А. Новые формы информационного и библиотечного обслуживания в научно-исследовательском институте // Информационные ресурсы. Интеграция. Технологии: "НТИ-97", 26-28 ноября 1997 г., г. Москва: Материалы конф. - М., ВИНТИ. - 1997. - С. 103-104.
119. Шокин Ю.И., Федотов А.М., Никульцев В.С. Вопросы развития сети Internet новосибирского научного центра // Вычислительные Технологии. - 1997. - Т. 2. - № 3. - С. 97-103.

120. Шокин Ю.И., Федотов А.М., Чубаров Л.Б., Жижимов О.Л. Информационно-телекоммуникационные ресурсы Сибирского отделения РАН // Выездное заседание координационного научного совета СО РАН по целевой программе: Информационно-телекоммуникационные ресурсы СО РАН, г. Иркутск, 29-30 июля, 2002 г.: Материалы заседания. - Иркутск, СО РАН. Иркутский научный центр. - 2003. - С. 150-163.
121. Шокин Ю.И., Федотов А.М., Белов С.Д., Бредихин С.В., Жижимов О.Л., Жуков А.Г., Никульцев С.В. Сеть Интернет Новосибирского научного центра СО РАН // Выездное заседание координационного научного совета СО РАН по целевой программе: Информационно-телекоммуникационные ресурсы СО РАН, г. Иркутск, 29-30 июля, 2002 г.: Материалы заседания. - Иркутск, СО РАН. Иркутский научный центр. - 2003. - С. 126-138.
122. Жижимов О.Л., Мазов Н.А., Фролов А.С. Доступ к базам CDS/ISIS данных через Internet // II международное рабочее совещание по электронным публикациям - "EL-PUB-97", Новосибирск, Академгородок, ИВТ СО РАН, 24-26 апреля 1997 г.: Тез. докл. - Новосибирск. - 1997. (<http://www.ict.nsc.ru/win/mathpub/comp-tech/content/t2n3/zhiz-htm/zhizhim.htm>).
123. BIREME/PAHO/WHO. WWWISIS: Ver. 3.0. Sao Paulo, Brazilian, Oct. 1997. (<http://www.bireme.br/wwwisis.htm>).
124. 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 // Научные и технические библиотеки. - 2000. - №2. - С. 52-57.
125. Жижимов О.Л., Мазов Н.А. Построение распределенной информационной системы на базе протокола Z39.50 (ISO 23950) // Новые информационные технологии в университетском образовании: Материалы междунар. науч.-метод. конф. - Новосибирск: НГУ, 1999. - С. 165-166.
126. Мазов Н.А., Жижимов О.Л. Интеграция Z39.50 и CDS/ISIS: состояние и перспективы развития // Научные и технические библиотеки. - 2000. - № 5. - С. 76-79.
127. Мазов Н.А. Разработка и построение распределенной информационно-библиотечной системы на основе СУБД CDS/ISIS: Автореф. дисс... канд. техн. наук. - Новосибирск, 2000. - 16 с.
128. Жижимов О.Л., Мазов Н.А. Состояние и перспективы использования протокола Z39.50 в информационном сообществе России // Информационное общество. - 2000. - № 2. - С. 39-43.
129. Жижимов О.Л., Мазов Н.А. Применение протокола Z39.50 в распределенной информационной системе Сибирского отделения

- РАН // 1-я Международная конференция "Centarl Asia - 1999".
Библиотечно-информационные ресурсы в науке, образовании,
культуре и бизнесе, Самарканд, Узбекистан, 10-16 октября 1999 г.:
Материалы конф. - Самарканд. - 1999. - С. 118-125.
130. Программа "Электронная библиотека Сибирского отделения РАН".
Целевая программа
(http://www.ict.nsc.ru/win/elbib/data/show_page.dhtml?2+148).
131. Электронная библиотека Сибирского отделения РАН: Проекты
(http://www.ict.nsc.ru/win/elbib/data/show_page.dhtml?2+147).
132. Жижимов О.Л., Мазов Н.А. Распределенная информационная
система СО РАН // V рабочее совещание по электронным
публикациям - "EL-PUB-2000", Новосибирск, Академгородок, ИВТ
СО РАН, 21-23 июня 2000 г.: Тез. докл., Новосибирск. - 2000.
(http://www.ict.nsc.ru/ws/show_abstract.dhtml?ru+1+16).
133. Шокин Ю.И., Барахнин В.Б., Пищик Б.Н., Ламин В.А., Жижимов
О.Л., Федотов А.М., Мазов Н.А., Покровский Н.Н. Распределенная
информационная система "Виртуальный музей науки и техники СО
РАН" // Пятая Всероссийская науч. конф. Электронные
библиотеки: перспективные методы и технологии, электронные
коллекции - "RCDL-2003", г. Санкт-Петербург, 29-31 окт., 2003 г.:
Труды конф. - СПб., НИИ Химии СПбГУ. - 2003. - С. 112-116.
134. Шокин Ю.И., Ламин В.А., Федотов А.М., Барахнин В.Б.,
Жижимов О.Л., Мазов Н.А., Пищик Б.Н., Покровский Н.Н.,
Рычкова Е.В. Виртуальный музей науки и техники СО РАН //
Выездное заседание координационного научного совета СО РАН по
целевой программе: Информационно-телекоммуникационные
ресурсы СО РАН, г. Иркутск, 29-30 июля, 2002 г.: Материалы
заседания. - Иркутск, СО РАН. Иркутский научный центр. - 2003. -
С. 118-125.
135. Хохлов Ю.С. Проект LIBWEB сетевой интеграции
информационных ресурсов ведущих библиотек России // 3-я
Междунар. конф. "Крым 96". Б-ки и ассоц. в меняющ. мире: нов.
технол. и нов. формы сотрудничества, п. Форос - г. Ялта, 1-9 июня,
1996 г.: Матер. конф., М. - 1996. - С. 18-22.
136. Хисамутдинов В.Р. Информационные ресурсы ИНИОН по
социальным и гуманитарным наукам в рамках проекта Libweb //
3-я Междунар. конф. "Крым 96". Б-ки и ассоц. в меняющ. мире:
нов. технол. и нов. формы сотрудничества, п. Форос - г. Ялта, 1-9
июня, 1996 г.: Матер. конф., М. - 1996. - С. 16-18.
137. Хохлов Ю.Е.; Воройский Ф.С. Проектирование распределенной
логической библиотечной сети LIBWEB // Науч. и техн. б-ки. -
1998. - № 2. - С. 24-30.

138. Куйбышев Л.А. Взаимодействие проекта LibWeb с другими сообществами и проектами // 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 1042.
139. Хохлов Ю.Е.; Аветисов М.А. Проект LibWeb - прошлое, настоящее и будущее // 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 1039-1040.
140. Жижимов О.Л. Об архитектуре распределенной информационной системы LibWeb на основе протокола Z39.50 // 8-я междунар. конф. "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня 2001 г.: Труды конф. - М., Изд-во ГПНТБ России. - 2001. - Т. 2. - С. 1043.
141. Елепов Б., Соболева Е., Федотова О., Шабанов А. Создание Российского центра, подобного OCLC: Первые шаги // Библиотека. - 2000. - № 6. - С. 27-29.
142. Баженов С.Р., Бредихина Н.А., Елепов Б.С., Жижимов О.Л., Мазов Н.А. Корпоративная библиотечно-информационная система Сибирского региона: проблемы и решения // Вторая Международная Конференция Интернет-Общество-Личность-2000 "ИОЛ-2000": Новые информационно-педагогические технологии, г. Санкт-Петербург, 28 февраля - 3 марта 2000 г.: Материалы Конференции. - СПб. - 2000. - С. 354.
143. Баженов С.Р., Елепов Б.С., Жижимов О.Л., Мазов Н.А., Бредихина Н.А., Кон Л.Ф. Принципы корпоративной распределенной библиотечной системы Сибирского региона // 7-я междунар. конф. "Крым 2000". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 3-11 июня 2000 г.: Труды конф. - Симферополь, Таврида. - 2000. - Т. 2. - С. 72-74.
144. Елепов Б.С., Баженов С.Р., Жижимов О.Л., Мазов Н.А. Формирование корпоративной распределенной информационно-библиотечной системы г. Новосибирска // 8-я междунар. конф. "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня 2001 г.: Труды конф. - М., Изд-во ГПНТБ России. - 2001. - Т. 2. - С. 688-692.
145. Елепов Б.С., Баженов С.Р., Жижимов О.Л., Мазов Н.А., Амельченко С.А. Новосибирская корпорация: итоги второго этапа

- // 9-я междунар. конф. "Крым 2002". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 8-16 июня 2002 г.: Труды конф. - М., Изд-во ГПНТБ России. - 2002. - Т. 2. - С. 570-574.
146. Мазов Н.А., Жижимов О.Л. Ресурсы НТИ библиотек и информационных центров НИУ в распределенной информационной системе СО РАН: текущее состояние и перспективы // Региональная научно-практ. конф. Библиотеки и духовно-культурное наследие, г. Новосибирск, 23-27 сентября 2002 г.: Материалы конф. - Новосибирск. - 2002. - С. 188-195.
147. Елепов Б.С., Баженов С.Р., Мазов Н.А., Жижимов О.Л. Корпоративная распределенная информационно-библиотечная система Сибирского региона на основе протокола Z39.50 // Библиотеки и духовно-культурное наследие: Материалы региональной научно-практ. конф., г. Новосибирск, 23-27 сентября 2002 г., Новосибирск. - 2002. - С. 195-202.
148. Амельченко С.А., Злотников М.А. Электронный каталог Новосибирской государственной областной научной библиотеки в едином информационном пространстве России // 9-я Международная конференция "Крым 2002". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 8-16 июня, 2002 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2002. - С. 103-106.
149. Баженов С.Р. Формирование корпоративной распределенной информационно-библиотечной системы г. Новосибирска // Научные библиотеки в новом тысячелетии: проблемы взаимодействия ресурсов. Новосибирск, Изд-во ГПНТБ СО РАН. - 2002. - С. 109-114.
150. Воройский Ф.С. Создание корпоративной сети публичных библиотек Москвы: проблемы и решения // Науч. и техн. б-ки. - 2000. - № 12. - С. 23-30.
151. Мазов Н.А., Жижимов О.Л., Бродовский А.И. Отображение электронного каталога ГПНТБ России в Z39.50 // 7-я междунар. конф. "Крым 2000". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 3-11 июня 2000 г.: Труды конф. - Симферополь, Таврида. - 2000. - Т. 1. - С. 221-224.
152. Воройский Ф. О корпоративных библиотечных системах и перспективах их развития в России // Инф. ресурсы России. - 2001. - № 3-4. - С. 7-10.
153. Батова Л.А., Замуруев Э.Н., Золоторева В.И. Создание корпоративной библиотечной сети вузов Москвы // Информационные технологии в учебном процессе университета, Москва. - 2000. - С. 131-136.

154. Рассина Э.Б. Корпоративные библиотечные проекты в области музыкального искусства // Межвузовское совещание "Проблемы развития вузовских сетевых информационно- библиотечных систем", Москва, 19-20 дек., 2000 г., М., Изд-во РХТУ. - 2000. – С. 53-55.
155. Древис Ю.Г. Опыт создания Московской корпоративной библиотечной вузовской сети // Межвузовское совещание "Проблемы развития вузовских сетевых информационно-библиотечных систем", Москва, 19-20 дек., 2000 г., М., Изд-во РХТУ. - 2000. – С. 22-23.
156. Дорошкова Е.Ф. Проблемы создания узла Корпоративной сети публичных библиотек в Северном округе г. Москвы // 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. – С. 730-733.
157. Круглова Н.В. Создание и использование местных корпоративных ресурсов в рамках реализации проекта "Создание корпоративной сети публичных библиотек г. Москвы" // 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 725-726.
158. Линдеман Е.В., Острая С.А. Использование корпоративных информационных ресурсов в публичных библиотеках Москвы: опыт работы ЦБС "Люблино" // 9-я Международная конференция "Крым 2002". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 8-16 июня, 2002 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2002. – С. 604-606.
159. Ярославская корпоративная библиотечная сеть (<http://lib.yar.ru/z3950.html>).
160. Бугрова Т.М. Совершенствование информационной поддержки научно-технической и инновационной деятельности на территории Ярославской области (Роль Центральной научно-технической библиотеки ЦНТИ в корпоративном проекте ярославских библиотек) // 7-я Международная конференция "Крым 2000". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 3-11 июня, 2000 г.: Труды конференции. - Симферополь. - 2000. – С. 85-88.
161. Лищук Т.А., Аккуратов Е.Г. Электронная медицинская библиотека как составная часть корпоративной библиотечной системы г. Ярославля: опыт создания, пути и перспективы развития // 8-я

- Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 790-793.
162. Смирнов В.Н., Палей Д.Э., Курчинский Д.Н. Корпоративная библиотечная сеть Ярославского региона // 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 683-687.
163. Палей Д.Э., Курчинский Д.Н., Смирнов В.Н. Принципы программно-технической организации Ярославской корпоративной библиотечной сети // 9-я Международная конференция "Крым 2002". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 8-16 июня, 2002 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2002. - С. 582-588.
164. Мершиева Л.В. Реализация проекта создания Нижегородской корпоративной библиотечной системы, 7-я Междунар. конф. "Крым 2000". Б-ки и ассоц. в меняющ. мире: нов. технол. и нов. формы сотрудничества, г. Судак, 3-11 июня, 2000 г.: Труды конференции. - Симферополь. - 2000. - С. 88-92.
165. Мершиева Л.В., Захарчук С.Н. Состояние и перспективы развития Нижегородской корпоративной библиотечной системы, 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 698-702.
166. Шлюз Z39.50 (<http://liber.nounb.sci-nnov.ru/phpyz/try.php>).
167. Шадрин Д.В. Корпоративная библиотечная система Омского региона // 8-я Международная конференция "Крым 2001". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 9-17 июня, 2001 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2001. - С. 693-695.
168. Шадрин Д.В. Развитие Корпоративной библиотечной системы Омского региона // 9-я Международная конференция "Крым 2002". Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества, г. Судак, 8-16 июня, 2002 г.: Труды конференции. - М., Изд-во ГПНТБ России. - 2002. - С. 588-591.
169. Тарелкин А.И. Состояние и перспективы развития информатизации в социальной сфере Омской области // Омск. науч. вестн. - 2001. - № 17. - С. 133-134.

170. Жижимов О.Л., Коджесян В.С., Мазов Н.А. Пример распределенной информационной системы на основе метаданных и международных стандартов // Вторая Всерос. науч. конф. Электронные библиотеки: перспективные методы и технологии, электронные коллекции, Протвино, 26-28 сентября 2000 г.: Сб. докл. - Протвино, ГНЦ ИФВЗ. - 2000. - С. 102-106.
171. Жижимов О.Л., Коджесян В.С., Мазов Н.А. Синтез технологий на примере распределенной информационной системы по газогидратам // Вычислительные технологии. - Новосибирск. - 2000. - №5. - С. 14-22.
172. Klerkx J., Poort J., De Batist M., Kuznetsov F., Dyadin Y., Belosludov V., Tohidi B., Ostergaard K., Soloviev V., Nesterov A., Chuvilin E., Yakushev V., Kwamme B., Bazhin N., Duchkov A., Dobretsov N., Zhizhimov O. A gas hydrate database online and for everyone // EOS, Transactions, American Geophysical Union. - 2000, 81. - № 22. - P. 63.
173. WebClarity Software Inc.
(<http://www.webclarity.info/products/bookwhere.html>).
174. Baranov V., Plemnek A., Sokolova N. et al. Review of Z39.50 servers and Z39.50 environment in Russia // Library Hi Tech. - 2000. - Vol. 18. - №. 4. - P. 304-314.
175. Library of Congress WWW-Z39.50 Gateway
(<http://lcweb.loc.gov/z3950/gateway.html>).
176. RUSLANet Z39.50 шлюз
(<http://www.ruslan.ru:8001/rus/z3950/gateway.php>).
177. Сигла — поиск в библиотеках (<http://www.sigla.ru>).
178. Хохлов А.Ю. Организация адаптивного распределенного поиска по библиотечным каталогам с использованием протокола Z39.50 // Электронные библиотеки. - 2003. - Т. 6. - № 2.
(<http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2003/-part2/xoxlov>).
179. Melvil® – The Catalog of the University of California Libraries
(<http://melvyl.cdlib.org>).
180. California Digital Library (<http://www.cdlib.org/>).
181. Национальный информационно-библиотечный центр "ЛИБНЕТ"
(<http://www.nilc.ru/>).
182. Florida Center for Library Automation (<http://www.fcla.edu/>).
183. WebLuis - Library User Information Service / State University System of Florida (<http://www.fcla.edu/z3950info.html>).

184. Miller P. Z39.50 for All // Ariadne, - 1999. - Issue 21.
(www.ariadne.ac.uk/issue21/z3950/).
185. Preece B., Thomson J. Union Catalogs and virtual Union Catalogs – repositioning Interlibrary Loan // 67th IFLA Council and General Conference, August 16-25, 2001
(<http://www.ifla.org/IV/ifla67/papers/091-108e.pdf>).
186. The Boston Library Consortium (<http://www.blc.org/index.html>).
187. Virtual Canadian Union Catalogue (vCuc)
(<http://www.collectionscanada.ca/8/6/index-e.html>).
188. Lunau Carrol D. The virtual Canadian union catalogue project (vCuc): Using Z39.50 to emulate a Centralized Union Catalogue // Resour. Sharing and Inf. Networks - 2000. - Vol. 14. - N 2. - P. 21-35.
189. Lunau Carrol D. The Virtual Canadian Union Catalogue Project (vCuc): Using Z39.50 to Emulate a Centralized Union Catalogue // 64th IFLA General Conference, August 16 - August 21, 1998.
(<http://www.ifla.org/IV/ifla64/081-160e.htm>).
190. ZLOT Project - Center for Digital Knowledge
(<http://www.unt.edu/zlot>).
191. Moen William E., Murray Kathleen R. ZLOT Project Phase 2: System Implementation of the LOT Resource Discovery Service. Z39.50 Server Implementation Issues and Recommendations. // Texas Center for Digital Knowledge, University of North Texas, Denton, Texas, June 23, 2003. (http://www.unt.edu/zlot/Deliverables/WA4_del_z-server_v2_krm_23jun2003.doc).
192. Moen William E., Murray Kathleen R., Lopatovska I. ZLOT: Technical Evaluation Methodology for RFP Demonstration Systems and RFP Written Responses // Texas Center for Digital Knowledge, University of North Texas, Denton, Texas, March 21, 2003. (http://www.unt.edu/zlot/project_docs/del_tech_eval_master_wem_21mar2003.doc).
193. Moen William E., Murray Kathleen R. A Service-Based Approach for Virtual Libraries // Texas Library Journal. - Vol. 78. - № 3. - 2002. - P. 5-8. (http://www.txla.org/pubs/tlj78/TLJ78_3.PDF).
194. Moen William E., Murray Kathleen R. Designing and Demonstrating a Resource Discovery Service for the Library of Texas // Texas Library Journal. - Vol. 78. - № 3. - 2002. - P. 9-14.
(http://www.txla.org/pubs/tlj78/TLJ78_3.PDF).
195. Moen William E., Murray Kathleen R. A Resource Discovery Service for the Library of Texas: Requirements, Architecture, and Interoperability Testing Texas Center for Digital Knowledge, School of Library and Information Sciences, University of North Texas, Denton
(<http://www.cni.org/tfms/2002b.fall/handouts/H-Resource-Moen.doc>).

196. Murray, K. R., Moen, W.E., Simon, S.J. (2002). ZLOT Project deliverable D: Architecture design. Retrieved May 6th, 2003. (<http://www.unt.edu/zlot/ProjectDocs.htm>).
197. UNiverse Project WWW Site (<http://www.fdggroup.co.uk/research/universe/>).
198. Ward S. The UNiverse project - a European Demonstration which adds value to the Virtual Union Catalogue. // 64th IFLA General Conference, 16-21 August 1998. (<http://www.ifla.org/ifla/IV/ifla64/132-160e.htm>).
199. Profile for the Use of Z39.50 Item Order Extended Service to Transport ILL Protocol APDUs (Z39.50/ILL Profile 1), 11 april 1996. (<http://www.nlc-bnc.ca/iso/ill/document/standard/z-ill-1a.pdf>).
200. Profile for the Use of Parameters from the ILL-Request APDU in Z39.50 Item Order (Z39.50/ILL Profile 2), 11 september 1996. (<http://www.nlc-bnc.ca/iso/ill/document/standard/z-ill-2a.pdf>).
201. IRIS: History and Background Information (<http://www.iris.ie/IRCO.htm>).
202. Peare J.D.T. IRIS - a Z39.50 national information system // The New Review of Information Networking. - 1995. - Vol. 1. - P. 17-32.
203. Ciardhuain S., Mogens S. The EUROPAGATE project // VINE. - 1994. - № 97 (Dec). - P. 18-23. (<http://www.dvt.dk/egate/egate.html>).
204. Murray R. IRIS and DALI: the dawn of a new era in distributed library systems // Library Technology News. - 1995. - № 19. - P. 1-3.
205. Murray, R; Smith, N; Pettman, I (1999). The UNiverse project: a review of progress up to the demonstration phase. // New Library World. - Vol. 100. - № 4. - P. 153-163.
206. Pettman I. Project UNiverse - an introduction // Proceedings of the Sixth Meeting of the European Association of Aquatic Sciences Libraries and Information Centres (EURASLIC) held at the Foundation for International Studies, Valetta, Malta, 25-26 April 1996. Plymouth: Plymouth Marine Laboratory, 1996. - P. 75-79.
207. UNEX – Home Page (<http://genie.ucd.ie/unex/index.html>).
208. NLUC'00 Материалы Международной конференции "Сеть и Логический Сводный Каталог", 17 апреля 2000, Санкт-Петербург, Россия. (http://genie.ucd.ie/unex/docs/NLUC_proceedings.rtf).
209. АРБИКОН - проект. (<http://www.arbicon.ru/project>).

210. Племнек А.И. Отчетный доклад за 2002 – 2003 гг. Общее собрание членов АРБИКОН // Второй Съезд Ассоциации Региональных Библиотечных Консорциумов АРБИКОН. - 18 декабря 2003 года (http://www.arbicon.ru/downloads/pdf/forum_final_report.pdf).
211. Племнек А.И., Усманов Р.Т. Подходы к реализации службы идентификации пользователей сервисов АРБИКОН // 11-я междунар. конф. "Крым 2004". Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса, г. Судак, 3-13 июня 2004 г.: Труды конф. - М., Изд-во ГПНТБ России. - 2004. - CD-ROM.
212. Eliot C. Experiences with Information Locator Services // Journal of Government Information. - 1999. - Vol. 26. - № 3. - P. 271-285.
213. FirstGov.gov: The U.S. Government's Official Web Portal (<http://www.firstgov.gov/index.shtml>).
214. GPO Access - GILS (http://www.access.gpo.gov/su_docs/gils/index.html).
215. GILS Utah Home Page (<http://gils.utah.gov>).
216. Government Links (<http://dlis.dos.state.fl.us/fgils/government.html>).
217. LAURIN (<http://laurin.uibk.ac.at/>).
218. Muhlberger G. FINAL REPORT: LAURIN. Version 01 // Telematics Applications Programme. Project LB-5629/A – LAURIN. - November, 2000. (<http://laurin.uibk.ac.at/old/reports/finalrep01.pdf>).
219. Design of the Local Node // LAURIN PROJECT. Libraries Project LB-5629/A, Report d50_004, August, 10th 1999, 103 p. (http://laurin.uibk.ac.at/old/reports/d50_004.pdf).
220. The overall Laurin Architecture // LAURIN PROJECT. Libraries Project LB-5629/A, Report d39_003, 15 p. (<http://laurin.uibk.ac.at/old/reports/d39003.pdf>).
221. Calvanese D., Catarci T., Santucci G. Building a Digital Library of Newspaper Clippings: The LAURIN Project // IEEE Advances in Digital Libraries 2000 (ADL 2000), May 22-24, Washington, 2000. - P. 15-26. (<http://csdl.computer.org/comp/proceedings/adl/2000/0659/-00/06590015abs.htm>).
222. Newspaper Clippings in a Digital World: The LAURIN Project. (<http://www.exploitlib.org/issue2/laurin/>).
223. Digitisation of Newspaper Clippings: The LAURIN Project. In: RLG DigiNews: Volume 3, Number 6, December 15, 1999. (<http://www.rlg.org/preserv/diginews/diginews3-6.html#feature>).

224. Retti G., Stehno B. The LAURIN-Thesaurus: A large, multilingual, electronic thesaurus for newspaper clipping archives. // *Journal of Documentation*. - 2004. - Vol. 60. - № 3. (accepted).
225. Retti G. The LAURIN Interface Suite: A software package for newspaper clipping archives. // *Journal of Digital Information Management*. - 2003. - Vol. 1. - № 4.
226. Calvanese D., Catarci T., Lenzerini T., Santucci G. The multilingual thesaurus of LAURIN. // *SEKE*. - 2002. - P. 83-90.
227. Calvanese D., Catarci T., Santucci G. LAURIN: A Distributed Digital Library of Newspaper Clippings. // *World Wide Web*. - 2001. - Vol. 4. - № 1-2. - P. 5-20.
228. Habitzel K., Muhlberger G. Offentliche Zeitungsausschnittsarchive in Deutschland, Osterreich und der Schweiz. Ein statistischer Ueberblick. // *Info. Information und Dokumentation in Archiven, Mediotheken, Datenbanken* Jg. 16. 2001, Heft 2, S. 92-97.
229. Muhlberger G., Habitzel K. Formale und inhaltliche Erschlie?ung von Zeitschriften mittels Digitalisierung. Stand der Technik, Probleme und Perspektiven // In: Wolfgang Hackl, Kurt Krolop (Hrsg.) *Wortverbunden - Zeitbedingt. Perspektiven der Zeitschriftenforschung* Innsbruck u.a.: StudienVerlag, 2001, S. 325-342.
230. Calvanese D., Catarci T., Santucci G. LAURIN: A Distributed Digital Library of Newspaper Clippings. // *Kyoto International Conference on Digital Libraries 2000*. - P. 187-194.
231. Muhlberger G. Electronic Clipping Of Articles From Printed Newspapers - A Report On The Outcomes Of The Eu-Funded Project Laurin. // *Library Computing*. - 2000. - Vol. 19. - P. 77-85.
232. Muhlberger G. Newspaper Clippings in a Digital World: The LAURIN Project. // *Exploit Interactive*, issue 2, 20 July 1999.
233. LAURIN – NUTS (<http://laurin.uibk.ac.at/nuts/>).
234. DNER - Distributed National Electronic Resource. (<http://www.jisc.ac.uk/dner/>).
235. The JOIN-UP Programme. (<http://edina.ed.ac.uk/projects/joinup/index.shtml>).
236. JISC - Joint Information Systems Committee. (<http://www.jisc.ac.uk>).
237. EDINA (<http://edina.ac.uk>).
238. EDINA ANNUAL REPORT for the Academic Year 2002 / 2003. (<http://edina.ac.uk/about/annrep/0203>).
239. Xgrain (<http://edina.ed.ac.uk/projects/joinup/xgrain/index.shtml>).

240. ZBLSA: Z39.50 Broker to Locate Serials and Articles
(<http://edina.ac.uk/projects/joinup/zblsa/>).
241. About zetoc (<http://zetoc.mimas.ac.uk/about.html>).
242. Stevenson J. What Everyone Wants: an evaluation of the zetoc service.
// Sconul Newsletter Spring 2003; 28, P. 50-53
(http://www.sconul.ac.uk/pubs_stats/newsletter/28/ART17.PDF).
243. MIMAS - Manchester Information & Associated Services
(<http://www.mimas.ac.uk/>).
244. Apps A., MacIntyre R. zetoc: a Dublin Core Based Current Awareness Service // Journal of Digital Information, Volume 2, Issue 2, Article No. 61, 2001-11-06 (<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Apps/>).
(<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Apps/apps-v2.pdf>).
245. Gate-Z A protocol gateway to support use of the Bath Profile / Enhancing JISC Activities for Learning and Teaching A: DNER enhancement // 24 January 2000
(<http://www.rdn.ac.uk/projects/gate-z/proposal/>).
246. The Bath Profile / An International Z39.50 Specification for Library Applications and Resource Discovery
(<http://www.ukoln.ac.uk/interop-focus/bath/>).
247. MIMAS Metadata for the DNER / JISC Services DNER: Z39.50/Authentication (<http://epub.mimas.ac.uk/dner.html>).
248. JAFER: Java Access For Electronic Resources toolkit project
(http://www.jisc.ac.uk/index.cfm?name=project_jafer).
249. Corfield A., Dovey M., Mawby R., Tatham C. JAFER ToolKit Project – Interfacing Z39.50 and XML // JCDL'02, July 13-17, 2002, Portland, Oregon, USA.
(<http://portal.acm.org/citation.cfm?doid=544220.544287>).
250. Corfield A., Dovey M., Mawby R., Tatham C. Z39.50 and XML - Bridging the old and the new // THE ELEVENTH INTERNATIONAL WORLD WIDE WEB CONFERENCE (WWW-2002), 7-11 May 2002, Honolulu, Hawaii, USA. (<http://www2002.org/presentations/dovey.pdf>, <http://www2002.org/CDROM/alternate/XS2/>).
251. Mawby R., Corfield A., Dovey M., Tatham C. A visual toolkit for information retrieval // Libraries Without Walls 4. The Delivery of Library Services to Distant Users: Distributed Resources - Distributed Learning The Fourth Conference, 14-18 September, 2001, Molyvos, Lesvos, Greece.
(<http://www.cerlim.ac.uk/conf/lww4/slides/mawby.ppt>).
252. UK Interoperability Focus
(<http://www.ukoln.ac.uk/interop-focus/about/>).

253. HILT (<http://hilt.cdlr.strath.ac.uk>).
254. Renardus EU-project (<http://www.renardus.org/index.html>).
255. Cheshire II Project (<http://sca.lib.liv.ac.uk/cheshire>,
<http://cheshire.berkeley.edu/>).
256. Larson Ray R., Moon R., McDonough J., Kuntz L., O'Leary P. Cheshire II: Design and Evaluation of a Next-Generation Online Catalog System (<http://sca.lib.liv.ac.uk/cheshire/papers.html>).
257. ELISE - The Introduction (<http://severn.dmu.ac.uk/elise>).
258. EULER: European Libraries and Electronic Resources in Mathematical Sciences. Telematics for Libraries, Project LB-5609. (<http://www.emis.de/projects/EULER/eulerp.pdf>).
259. EULER - Project Reports and Deliverables (<http://www.emis.de/projects/EULER/Reports/index.html>).
260. Jost M., Becker Hans J. EULER - A Real Virtual Library for Mathematics // High Energy Physics Libraries Webzine, Issue 9, February 2004 (<http://library.cern.ch/HEPLW/9/papers/5/>).
261. EULER - Advanced Search (<http://www.emis.de/projects/EULER/advanced.html>).
262. The AQUARELLE project / TELEMATICS Application Program of the European Commission, Project IE-2005. - 1996.
263. Vercoustre A.-M., Sutcliffe D., Michard A., Stapleton M., Scholl M., Christophides V. The Aquarelle resource discovery system // Computer Networks and ISDN Systems. - 1998. - Vol. 30 . - № 13. - P. 1185-1200.
264. Michard A (1998) Aquarelle: sharing cultural heritage through multimedia telematics: final report. (<http://aqua.inria.fr/aquarelle/EN/final-report.html>).
265. Signore O. The Aquarelle - CIMI Z39.50 Profile and its Mapping to Cultural Databases // ERCIM News. - 1998. - № 33. (http://www.ercim.org/publication/Ercim_News/enw33/signore.html).
266. The ZAVIER Project (<http://www.dstc.edu.au/Research/Projects/Z3950/ZAVIER>).
267. Bird L., Finnigan S. ZAVIER: Executive Report Summary // DSTC, Issue 1.0, 14 May 1999 (<http://www.dstc.edu.au/Research/Projects/Z3950/ZAVIER/ZAV-NewExecSum.doc>).
268. Beaumont A. ZAVIER - Wider than Libraries Deeper than the Web // 10th VALA Biennial Conference and Exhibition. Books and bytes: technologies for the hybrid library (VALA-2000), 6-18 February, Melbourne, 2000. (<http://vala.org.au/vala2000/2000pdf/Beaumont.PDF>).

269. ZedMoV - Z+SQL Internal Testbed
(<http://www.dstc.edu.au/Research/Projects/Z3950/ZedMoV>).
270. Finnigan S.M., Bird L.J., Colomb R.M. Distributed searching across cultural resources using Z39.50 and SQL: a powerful combination // Information online & on disk 99. Strategies for the next millennium. Proceedings of the Ninth Australasian Information Online & On Disc Conference and Exhibition. Sydney Convention and Exhibition Centre, Sydney Australia, 19–21 January 1999.
(<http://www.csu.edu.au/special/online99/proceedings99/104b.htm>).
271. The Arts and Humanities Research Board (<http://www.ahrb.ac.uk>).
272. AHDS A Guides to Good Practice / Archaeology Data Service / Geophysical Data in Archaeology: A Guide to Good Practice by Armin Schmidt with contributions from Tony Austin, Andrew David, Kate Fernie, William Kilbride, Paul Linford, Julian Richards and Damian Robinson (<http://ads.ahds.ac.uk/project/goodguides/geophys/>).
273. AHDS A Guides to Good Practice / Archiving Aerial Photography and Remote Sensing Data: A Guide to Good Practice Robert Bewley, Danny Donoghue, Vince Gaffney, Martijn van Leusen, and Alicia Wise (<http://ads.ahds.ac.uk/project/goodguides/apandrs/>).
274. AHDS A Guides to Good Practice / Archaeology Data Service / CAD: A Guide to Good Practice by Harrison Eiteljorg II, Kate Fernie, Jeremy Huggett and Damian Robinson with an additional contribution by Bernard Thomason (<http://ads.ahds.ac.uk/project/goodguides/cad/>).
275. AHDS A Guides to Good Practice / Digital Archives from Excavation and Fieldwork: Guide to Good Practice. Second Edition / edited by Julian Richards and Damian Robinson with additional contributions by Tony Austin, Adrian Brown, Duncan Brown, Geraint Coles, David Dawson, Anne Dodd, Kate Fernie, Julie Gardiner, William Kilbride, Gary Lock, Christine Longworth, Nick Merriman, Paul Miller, Diana Murray, Jez Reeve and Alicia Wise
(<http://ads.ahds.ac.uk/project/goodguides/excavation/>).
276. AHDS A Guides to Good Practice / GIS Guide to Good Practice edited by Mark Gillings and Alicia Wise with contributions by Mark Gillings, Peter Halls, Gary Lock, Paul Miller, Greg Phillips, Nick Ryan, David Wheatley, and Alicia Wise
(<http://ads.ahds.ac.uk/project/goodguides/gis/index.html>).
277. AHDS A Guides to Good Practice / Creating and Using Virtual Reality: a Guide for the Arts and Humanities / edited by Kate Fernie and Julian D. Richards with contributions by: Tony Austin, Rachael Beach, Aaron Bergstrom, Sally Exon, Marc Fabri, Michael Gerhard, Catherine Grout, Stuart Jeffrey, Mike Pringle, Damian Robinson, Nick Ryan, Melissa Terras and additional Case Studies by: Kate Allen, Clive

- Fencott, Learning Sites, Anthony McCall. AHDS 2002
(http://vads.ahds.ac.uk/guides/vr_guide/index.html).
278. AHDS A Guides to Good Practice / Digitising History A Guide to Creating Digital Resources from Historical Documents SEAN TOWNSEND, CRESSIDA CHAPPELL AND OSCAR STRUIJVE
Sean Townsend, Cressida Chappell, Oscar Struijve 1999
(http://hds.essex.ac.uk/g2gp/digitising_history/index.asp).
279. AHDS A Guides to Good Practice / A Place in History A Guide to Using GIS in Historical Research IAN GREGORY Ian Gregory 2002
(<http://hds.essex.ac.uk/g2gp/gis/index.asp>).
280. AHDS A Guides to Good Practice / Creating Digital Performance Resources A Guide to Good Practice Performing Arts Data Service Guide to Good Practice Creating Digital Performance Resources Series editors Catherine Owen, Stephen Arnold, Tony Pearson Edited by Barry Smith, 2003 with contributions by Mark Batty, Alan Beck, Christie Carson, Steve Dixon, David Hughes, Sophia Lycouris, Barry Russell and Barry Smith (<http://ahds.ac.uk/creating/guides/-performing-resources/contents.htm>).
281. AHDS A Guides to Good Practice / Creating Digital Audio Resources A Guide to Good Practice Performing Arts Data Service Guide to Good Practice Creating Digital Audio Resources, 2003 Nick Fells, Pauline Donachy and Catherine Owen Editors: Catherine Owen, Kate Iles (http://ahds.ac.uk/creating/guides/audio-resources/-GGP_Audio_Contents.htm).
282. AHDS A Guides to Good Practice / Creating and Documenting Electronic Texts: A Guide to Good Practice by Alan Morrison, Michael Popham, Karen Wikander
(<http://ota.ahds.ac.uk/documents/creating/>).
283. AHDS A Guides to Good Practice / Creating and Using Virtual Reality: a Guide for the Arts and Humanities / AHDS 2002 / edited by Kate Fernie and Julian D. Richards with contributions by: Tony Austin, Rachael Beach, Aaron Bergstrom, Sally Exon, Marc Fabri, Michael Gerhard, Catherine Grout, Stuart Jeffrey, Mike Pringle, Damian Robinson, Nick Ryan, Melissa Terras and additional Case Studies by: Kate Allen, Clive Fencott, Learning Sites, Anthony McCall
(http://vads.ahds.ac.uk/guides/vr_guide/index.html).
284. AHDS A Guides to Good Practice / Creating Digital Resources for the Visual Arts: Standards and Good Practice authors: Visual Arts Data Service: Catherine Grout, Phill Purdy, Janine Rymer Technical Advisory Service for Images: Karla Youngs, Jane Williams, Alan Lock, Dan Brickley, Oliver Moss University of Bristol & Catherine Grout, Phill Purdy & Janine Rymer (Visual Arts Data Service)
(http://vads.ahds.ac.uk/guides/creating_guide/contents.html).

285. Archaeology Data Service (<http://ads.ahrb.ac.uk>).
286. The Oxford Text Archive (<http://ota.ahds.ac.uk>).
287. The Visual Arts Data Service (<http://vads.ahds.ac.uk>).
288. ARTISTE Project (<http://www.artisteweb.org>).
289. Addis M., Boniface M., Goodall S., Grimwood P., Kim S., Lewis P., Martinez K., Stevenson A. Integrated Image Content and Metadata Search and Retrieval across Multiple Databases // Lecture Notes in Computer Science. - 2003. - № 2728. - P. 91-100.
290. Stevenson A. User Guide: Accessing ARTISTE using SRW // Technical Note ARTISTE T57, 16 August 2002 (http://www.it-innovation.soton.ac.uk/ftpdata/artiste/srw/-T57_UserGuide_ArtisteSRW_02.pdf).
291. D6.1 Distributed Query Layer and Metadata Report (<http://www.it-innovation.soton.ac.uk/artiste/documentation/D6.1final.pdf>).
292. D6.2 Impact on World-Wide Metadata Standards (<http://www.it-innovation.soton.ac.uk/artiste/documentation/D6.2.final.pdf>).
293. SCULPTEUR - Semantic and content-based multimedia exploitation for European benefit (<http://www.sculpteurweb.org>).
294. Addis, M., Boniface, M., Goodall, S., Grimwood, P., Kim, S., Lewis, P., Martinez, K. and Stevenson, A. SCULPTEUR: Towards a New Paradigm for Multimedia Museum Information Handling. In Proceedings of Semantic Web ISWC 2870. - 2003. - P. 582 -596.
295. DADDI - Digital Archive through Direct Digital Imaging (<http://www.uffizi.firenze.it/Dta/daddi-eng.html>).
296. ARCO - Augmented Representation of Cultural Objects (<http://www.arco-web.org>).
297. ORPHEUS (<http://www.orpheus-edu.org>).
298. SCHEMA - Network of Excellence in Content-Based Semantic Scene Analysis and Information Retrieval (<http://www.iti.gr/SCHEMA>).
299. VITRA - Veridical Imaging of Transmissive and Reflective Artefacts (<http://www.vitra.org>).
300. GBIF Portal (http://www.gbif.org/GBIF_org/what_is_gbif).
301. CBIF (http://www.cbif.gc.ca/home_e.php).
302. BiOSC (http://www.cbif.gc.ca/pls/bb/bcin_specs.portal?p_ifx=cbif).

303. REMIB - The World Information Network on Biodiversity
(http://www.conabio.gob.mx/remib_ingles/doctos/remib_ing.html).
304. DiGIR - Distributed Generic Information Retrieval (<http://digir.net/>).
305. ITIS - Integrated Taxonomic Information System
(<http://www.itis.usda.gov>).
306. Species 2000 Home (<http://www.sp2000.org>).
307. SPICE Species 2000 (<http://www.systematics.reading.ac.uk/spice>).
308. LITCHI (<http://litchi.biol.soton.ac.uk>).
309. OBIS - Ocean Biogeographic Information System
(<http://www.iobis.org>).
310. Zhang Y., Grassle J.F. A portal for the Ocean Biogeographic Information System // *Oceanologica Acta*. - 2003. - № 25. - P. 193–197.
311. ABRS - Australia's Virtual Herbarium.// *Biologue*, 2001, 24, 9.
312. ABRS - Australian Biological Resources Study
(<http://www.ea.gov.au/biodiversity/abrs/>).
313. SABONET - Southern African Botanical Diversity Network
(<http://www.sabonet.org>).
314. BioCISE (<http://www.bgbm.fu-berlin.de/biocise>).
315. ENHSIN (<http://www.nhm.ac.uk/science/rco/enhsin>).
316. Species Analyst (<http://speciesanalyst.net/>).
317. Z.X Home (<http://tsadev.speciesanalyst.net/z.x/>).
318. The Darwin Core./ Version: 1.3
(http://tsadev.speciesanalyst.net/DarwinCore/darwin_core.asp).
319. FishNet (<http://habanero.nhm.ku.edu/fishnet/>).
320. Vieglais D., Wiley E.O., Robins C.R., Peterson A.T. Harnessing Museum Resources for the Census of Marine Life: The FISHNET project // *Oceanography*. - 2000. - Vol. 13. - № 3. - P.10–13.
(<http://www.iobis.org/PDFs/2000Vieg.pdf>).
321. MaNIS Home (<http://elib.cs.berkeley.edu/manis>).
322. HerpNet (<http://herpnet.org/>).
323. MaPSTeDi (<http://mapstedi.colorado.edu/>).
324. Clearinghouse
(<http://www.fgdc.gov/clearinghouse/clearinghouse.html>).

325. Websites Associated with Clearinghouse
(http://clearinghouse4.fgdc.gov/registry/clearinghouse_sites.html).
326. Hourly Status of NSDI Clearinghouse Nodes
(<http://registry.gsdi.org/serverstatus>).
327. SEDAC - Socioeconomic Data and Applications Center
(<http://sedac.ciesin.columbia.edu>).
328. Global Change Master Directory Web Site (<http://gcmd.gsfc.nasa.gov>).
329. NBII - Metadata Clearinghouse: Clearinghouse Nodes
(<http://www.nbii.gov/search/clearinghouse/nodes.html>).
330. The National Biological Information Infrastructure (NBII)
(<http://www.nbii.gov>).
331. USGS (U.S. Geological Survey) Home Page (<http://www.usgs.gov>).
332. Colorized Content Standards for Digital Geospatial Metadata –
Biological Data Profile, FGC-STD-001.1-1999
(<http://badger.state.wi.us/agencies/wlib/sco/meta/colorstd/-colorstdebio99.html>).
333. IAI Data and Information System Project (<http://disbr1.iai.int>).
334. CEOS Home Page (<http://www.ceos.org>).
335. INFEO - Information on Earth Observation (<http://www.infeo.org>).
336. Earth Observation Portal (<http://www.eoportal.org>).
337. Mikusch E., Diedrich E., Gohmann M., Kiemle S., Reck C., Reisig R.,
Wildegger W., Wolfmuller M. Data Information and Management
System for the Production, Archival and Distribution of Earth
Observation Products.
(http://www.dfd.dlr.de/archive/dims/dims_www_paper.html).
338. EOPOLE: Useful Links
(<http://www.geog.ucl.ac.uk/eopole/useful-links.html>).
339. Кузнецов Е.Б., Балашов А.Д. Интеграция электронной библиотеки
спутниковых данных в международную систему космической
информации // Труды 5-ой Всероссийской научной конференции
"Электронные библиотеки: перспективные методы и технологии,
электронные коллекции" RC DL-2003, 29-31 окт. 2003 г. -
Санкт-Петербург. - С. 88-95.

Список иллюстраций

1.1	Структура запроса <i>RPN</i>	21
1.2	Модель извлечения записей	22
3.1	Типичная схема распределенной модульной информационной системы	36
3.2	Схема многослойной РИС	37
3.3	Уровни сетевого взаимодействия в <i>Z39.50</i>	38
3.4	Уровни сетевого взаимодействия <i>Z-data provider</i> с сетевой СУБД	39
3.5	Уровни сетевого взаимодействия <i>Z-data provider</i> с другим сервером <i>Z39.50</i>	39
3.6	Модель РИС на основе протокола <i>Z39.50</i>	40
3.7	Временная диаграмма режима 1 взаимодействия клиента и серверов в РИС	44
3.8	Временная диаграмма режима 2 взаимодействия клиента и серверов в РИС	45
3.9	Временная диаграмма режима 3 взаимодействия клиента и серверов в РИС	46
5.1	Основные блоки сервера <i>ZooPARK</i>	73
5.2	Доступ к данным в сервере <i>ZooPARK</i>	74
5.3	Доступ к статистической информации через шлюз <i>Z-GW</i> сервера <i>ZooPARK</i>	79
5.4	Доступ к базам данных через шлюз <i>Z-GW</i> сервера <i>ZooPARK</i>	82
5.5	Библиографическая запись в схеме <i>UIGGM</i>	83
6.1	Свойства сервиса <i>ZooPARK</i> при сбое.	90
6.2	Первая страница <i>WWW</i> сервера	94
6.3	Проверка сервера <i>SRW/SRU</i>	94
6.4	Проверка модуля поддержки <i>PHP</i>	95
6.5	Проверка модуля поддержки <i>PERL</i>	96
6.6	Шлюз <i>Z-GW</i> и <i>Explain</i> -запись <i>databaseInfo</i> для <i>tgils</i>	97
7.1	Точка входа в РИС СО РАН	158
7.2	Точка входа в распределенную информационную систему <i>LibWeb</i> (ЦНСХБ, шлюз <i>Z39.50-HTTP</i>)	160

7.3	Точка входа в распределенную информационную систему LibWeb (БЕН РАН, шлюз <i>Z39.50-HTTP</i>)	163
7.4	Точка входа в Новосибирскую распределенную библиотечно-информационную систему (шлюз <i>Z39.50-HTTP</i>)	165
7.5	Точка входа в Московскую распределенную библиотечно-информационную систему (шлюз <i>Z39.50-HTTP</i>)	167
7.6	Точка входа в Ярославскую корпоративную библиотечно-информационную систему (шлюз <i>Z39.50-HTTP</i>)	170
7.7	Точка входа в Нижегородскую Корпоративную библиотечную сеть (шлюз <i>Z39.50-HTTP</i>)	171
7.8	Точка входа в Омскую Корпоративную библиотечную информационную систему (шлюз <i>Z39.50-HTTP</i>)	173
7.9	Общая схема информационной системы по газгидратам	176
7.10	Интерфейс пользователя при поиске	178
7.11	Выход в другую библиографическую информационную систему	179
7.12	Масштабирование и географический поиск	180
7.13	Интерфейс навигации по рубриктору	181
8.1	Интерфейсы портала <i>Сигла</i>	186
8.2	Архитектура виртуального каталога <i>ZLOT</i>	189
8.3	Схема реализации параллельного поиска в <i>UNiverse</i>	191
8.4	Интерфейсы портала <i>Арбикон</i>	193
8.5	Шлюз <i>GILS U.S. Government Printing Office</i>	195
8.6	Общая архитектура <i>LAURIN</i>	196
8.7	Архитектура сервера <i>JAFER</i>	201
8.8	Интерфейс <i>Cheshire II</i>	202
8.9	Архитектура <i>ELISE II</i>	204
8.10	Архитектура информационной системы <i>EULER</i>	207
8.11	Поисковый интерфейс системы <i>EULER</i>	208
8.12	Архитектура <i>Aquarelle</i>	210
8.13	Архитектура <i>Zavier</i>	210
8.14	Шлюз <i>Zavier</i> . Интерфейс для поиска	211
8.15	Распределенная архитектура <i>ZedMoV</i>	212
8.16	Поисковый интерфейс <i>AHDS</i>	213
8.17	Архитектура системы <i>ARTISTE</i>	214
8.18	Поисковый интерфейс <i>MaNIS</i>	217
8.19	Схема <i>Clearinghouse</i>	219
8.20	Взаимодействие <i>CIP</i> и <i>Z39.50</i> систем	220
8.21	Иерархические коллекции <i>CIP</i>	221

Список таблиц

1.1	Классы объектов <i>Z39.50</i>	19
1.2	Стандартные наборы поисковых атрибутов	20
1.3	Распространенность сервисов <i>Z39.50</i>	22
2.1	Объекты <i>Z39.50</i> , переопределенные в <i>CIP</i>	33
3.1	Поддерживаемость категорий <i>Explain</i> зарегистрированными серверами	48
4.1	Структура маркера записи ISO-2709	53
4.2	Поисковые USE-атрибуты <i>Bib-1</i> для <i>Dublin Core</i>	56
4.3	Объекты <i>Z39.50</i> для <i>GEO</i>	57
4.4	Объекты <i>Z39.50</i> для <i>Digital Collections</i>	59
4.5	Объекты <i>Z39.50</i> для <i>CIMI</i>	60
4.6	Связь поисковых атрибутов и элементов схемы данных <i>Zthes</i>	63
5.1	Локальные объекты сервера <i>ZooPARK</i>	75
5.2	Схемы данных, поддерживаемые сервером <i>ZooPARK</i>	84
5.3	Распространенность серверов <i>ZooPARK</i> и <i>Z-IRBIS</i>	85
6.1	Состав дистрибутива <i>ZooPARK</i>	87
6.8	Параметры запуска сервера <i>ZooPARK</i>	123
6.9	Параметры конфигурационного файла сбора статистики <i>stat.cfg</i>	124
6.10	Команды шлюза <i>Z-GW</i> и переменные	140
6.11	Простые переменные шлюза <i>ZGW</i>	141
6.12	Переменные-шаблоны шлюза <i>ZGW</i>	144
7.1	Серверы <i>Z39.50</i> РИБС г. Новосибирска	166
7.2	Серверы <i>ZooPARK</i> и <i>Z-IRBIS</i> в РИС г. Москвы	168
7.3	Серверы <i>Z39.50</i> РИБС г. Нижний Новгород	172
7.4	Серверы <i>Z39.50</i> РИБС г. Омска	174

Предметный указатель

- .NET, 26, 28, 29, 65
АРБИКОН, 13, 192, 194
АС Библиотека-3, 86
БАН, 161, 192
БЕН РАН, 161, 163, 168
БКС, 85
Библиотека-2000, 86
Библиотека-4, 169
ЦНСХБ РАСХН, 60, 162
Центр ПИК МК РФ, 162
ГЦНМБ, 60
ГИВЦ МК РФ, 86, 169
ГПНТБ России, 67, 85, 161, 162, 168
ГПНТБ СО РАН, 150, 153, 156, 162, 164, 166
ГРНТИ, 60, 194, 337, 352
ИК СО РАН, 150
ИКИ РАН, 221
ИНИОН РАН, 162
ИОО, 85, 163, 168
ИРБИС, 68
ИРИО, 162
Информсистема, 75, 169
ЛИБНЕТ, 13, 185
МАРК-SQL, 68, 75, 87, 119, 169
МГУ, 85, 162
ОИГГМ СО РАН, 71, 84, 149–151, 153–155, 162, 163, 166
РФФИ, 159
РГБ, 162
РГНФ, 159
РКП, 192
СПбГПУ, 192
СПбГУ, 162
Сигла, 185
УДК, 194, 336, 351
ВЦ РАН, 161, 162
ВГБИЛ, 168
ВИНИТИ, 60, 149
A-Elite, 86
ABRS, 215
ADS, 212
AHDS, 202, 212
AHRD, 212
Apache, 133
Aquarelle, 206, 209
ARCO, 214
ARTISTE, 213, 214
ARTVICMM, 211
ASN.1, 18, 19, 21, 28, 31, 32, 53
ASP, 28, 216
BATH, 200
BioCISE, 215
BiOSC, 215
BNSC, 33
BookWhere, 184
CBIF, 214
CCRS, 33
CDS/ISIS, 71, 74, 84, 87, 109–111, 150–152, 154
CEO, 33
CEOS, 32, 33, 220
Cheshire, 202
CIMI, 55, 59, 72, 84, 206, 209, 211, 212, 346
CIP, 32, 33, 44, 49, 220, 221
cmd, 74
CNIDR Zserver, 86
COMDIG, 210
COPEARMS, 206
CORBA, 26, 28, 29, 65
Crossnet, 211

- CSDGM, 56
- D3Zserver, 85, 172
- DAAC, 219
- DADDI, 214
- DALI, 190
- Darwin Core, 216
- DCE, 27
- DCOM, 27
- DDC, 202
- DDU, 211
- DiGIR, 215
- DIMS, 221
- DLR, 33
- DNER, 197, 198, 200
- Docusend, 197
- DSTC, 211
- Dublin Core, 20, 55, 59, 206–208, 212, 216
- EDINA, 197, 198
- ELISE II, 203
- EMS, 208
- ENHSIN, 215
- ESA, 33
- EULER, 206–209
- EUROPAGATE, 190
- FGDC, 56
- FGDC Clearinghouse, 219
- FirstGOV, 195
- FishNet, 216, 218
- GASHYDAT, 174
- Gate-Z, 200
- GBIF, 214
- GCMD, 219
- GEO, 56, 72, 84, 218, 292, 352
- GILS, 20, 54–56, 64, 72, 84, 95, 159, 177, 179, 180, 188, 194, 339
- GIPER, 86
- GPO, 195
- HELIX, 206
- HerpNet, 218
- HILT, 202
- IAI-DIS, 220
- IBM DB2, 67
- IndexData, 21, 23, 47, 84–86, 95, 99, 101, 130, 177
- INFEO, 221
- Informix, 63, 67
- Interoperability Focus, 202
- IRBIS, 84
- IRIS, 190
- ISI CDE, 151
- ITIS, 215
- J2EE, 26, 28, 29, 65
- JAFER, 200
- JAVA, 27, 28, 153, 200
- JDBC, 25, 27–29
- JIDI, 206
- JISC, 197, 206, 212
- JOIN-UP, 197
- JSP, 28
- LAURIN, 196, 197
- LDAP, 18, 31, 47, 49, 100, 194
- libClip, 197
- Liber, 68, 85, 166, 169
- LiberMedia, 75, 85
- LibWeb, 13, 159–163
- LIS, 197
- LITCHI, 215
- MaNIS, 217
- MaPSTeDi, 218
- Melvyl®, 185
- MeSH, 60, 194
- MIMAS, 198, 200
- MS ADO, 27–29, 71, 74, 118
- MS IIS, 133
- MS SQL Server, 37, 67, 71, 74, 75, 87, 111, 112, 118, 119, 125, 155
- MySQL Server, 74, 87
- NASA, 33
- NASDA, 33
- NBII Clearinghouse, 220
- NSF, 203, 217, 218
- OBIS, 215
- ODBC, 25, 27, 29, 65, 74, 75, 119, 200

- OLE DB, 27, 28, 75
OLSC Client, 184
OMG, 28
ORACLE, 28, 37, 63, 67, 75, 196
ORPHEUS, 214
OTA, 212

PERL, 74, 75, 89, 133, 134, 197
PHP, 74, 75, 89, 133, 134

RDN, 198
RDU, 211
REMB, 215
Renardus, 202
RMI, 27
RMON, 18
RUSLAN, 47, 84, 85, 184, 194
RUSmarc, 14, 52, 54, 56, 66, 72,
84, 85, 103, 107, 108, 110,
113, 121, 127, 135, 138,
159, 164, 166, 169, 358

SABONET, 215
SCHEMA, 214
SCULPTEUR, 214
SEDAC, 219
sh, 74
SNMP, 18
SOAP, 28, 71, 72, 95, 133, 213, 215
Species Analyst, 215, 216
Species 2000, 215
SPICE, 215
SPIRS, 151
SRW, 28, 71, 72, 95, 133, 213, 215
Sybase, 75

TASI, 206
TCL, 89
tcl, 74

UNEX, 192
UNIMARC, 52, 72, 84, 85, 135,
166
UNIverse, 188–190, 192
USMARC, 52, 54, 55, 57, 60, 66,
72, 84, 85, 107, 108, 135,
159, 166, 169

VADS, 212

vCuc, 187
VITRA, 214

WWWISIS, 152

X.500, 18
Xgrain, 197

Z+SQL, 63, 65, 77, 81, 84, 211
Z-GW, 88, 89, 126, 134–136, 139,
144
Z-IRBIS, 71, 84, 85, 168, 173, 174
Z-ISIS, 87, 88, 109, 144
Z-MMARK, 87, 119
Z-MSADO, 118
Z-MSSQL, 87, 88
Z-MySQL, 87, 88
Z-ORIGIN, 88
Z-REMOTE, 76, 87, 104, 126, 145,
146, 148
Z-SHELL, 87
Z-TEST, 87, 88
ZASP, 216
Zavier, 209, 211
ZBigServer, 216
ZBLSA, 197, 198
Zebra, 88, 92, 96, 98, 99, 101, 103,
130, 133, 177
ZedMoV, 211
ZETOC, 197, 198
ZIG, 15, 206
ZING, 223
ZLOT, 188
ZPortal, 216
Zthes, 20, 61–63, 72, 84, 135, 159,
180, 342
ZX, 215, 216

Приложение А

Наборы атрибутов

А.1 Атрибуты Bib-1

А.1.1 Use attributes

Name	Val	Name1	Text	USmarc
Personal name	1	Name-personal	A person's real name, pseudonym, title of nobility nickname, or initials.	100, 400, 600, 700 800
Corporate name	2	Name-corporate	An organization or a group of persons that is identified by a particular name. (Subject name headings are included.)	110, 410, 610, 710 810
Conference name	3	Name-conference	A meeting of individuals or representatives of various bodies for the purpose of discussing topics of common interest. (Subject name headings are included.)	111, 411, 611, 711 811

Name	Val	Name1	Text	USmarc
Title	4	Title	A word, phrase, character, or group of characters, normally appearing in an item, that names the item or the work contained in it.	130, 21X-24X, 440, 490, 730, 740, 830 840, subfield \$t in the following: 400, 410, 410, 600
Title series	5	Title-series	Collective title applying to a group of separate, but related, items.	440, 490, 830, 840 subfield \$t in the following: 400,410
Title uniform	6	Title-uniform	The particular title by which a work is to be identified for cataloging purposes.	130, 240, 730, subfield \$t in the following: 700,710
ISBN	7	Identifier-ISBN	International Standard Book Number – internationally agreed upon number that identifies a book uniquely. Cf. ANSI/NISO Z39.21 and ISO 2108.	020
ISSN	8	Identifier-ISSN	International Standard Serial Number – internationally agreed upon number that identifies a serial uniquely. Cf. ANSI/NISO z39.9 and ISO 3297.	022, 4XX\$\$x, 7XX\$\$x
LC card number	9	Control number-LC	Character string that uniquely identifies a record in the Library of Congress database.	010, 011
BNB card number	10	Control number-BNB	Character string that uniquely identifies a record in the British National Bibliography.	015
BGF(sic) number	11	Control number-BNF	Character string that uniquely identifies a record in the Bibliotheque Nationale Francais.	015
Local number	12	Control number-local	Character string that uniquely identifies a record in a local system (i.e., any system that is not one of the four listed above).	001, 035
Dewey classification	13	Classification-Dewey	A classification number from the Dewey Decimal Classification, developed by Melvyl Dewey.	082

Name	Val	Name1	Text	USmarc
UDC classification	14	Classification-UDC	A classification number from Universal Decimal Classification, a system based on the Dewey Decimal Classification.	080
Bliss classification	15	Classification-Bliss	A classification number from the Bliss Classification, developed by Henry Evelyn Bliss.	
LC call number	16	Classification-LC	A classification number from the US Library of Congress Classification.	050
NLM call number	17	Classification-NLM	A classification number from the US National Library of Medicine Classification.	060
NAL call number	18	Classification-NAL	A classification number from the US National Agriculture Library Classification.	070
MOS call number	19	Classification-MOS	A classification number from Mathematics Subject Classification, compiled in the Editorial Offices of Mathematical Reviews and Zentralblatt für Mathematik.	
Local classification	20	Classification-local	A local classification number from a system not specified elsewhere in this list of attributes.	
Subject heading	21	Subject	The primary topic on which a work is focused.	600, 610, 611, 630 650, 651, 653, 654
Subject Rameau	22	Subject-RAMEAU	Subject headings from Repertoire d'autorité de matières encyclopedique unifie – maintained by the Bibliothèque Nationale (France).	
BDI index subject	23	Subject-BDI	Subject headings from Bibliotek Dokumentasjon Informasjon – a controlled subject vocabulary used and maintained by the five Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden).	
INSPEC subject	24	Subject-INSPEC	Subject headings from Information Services for the Physics and Engineering Communities – the Information Services Division of the Institution of Electrical Engineers.	600i2, 610i2, 600i2, 610i2, 650i2, 651i2

Name	Val	Name1	Text	USmarc
MESH subject	25	Subject-MESH	Subject headings from Medical Subject Headings – maintained by the US National Library of Medicine.	600i2, 610i2, 611i2, 630i2, 650i2, 651i2
PA subject	26	Subject-PA	Subject headings from Thesaurus of Psychological Index Terms – maintained by the Retrieval Services Unit of the American Psychological Association.	600i2, 610i2, 611i2, 630i2, 650i2, 651i2
LC subject heading	27	Subject-LC	Subject headings from US Library of Congress Subject Headings.	600, 610, 611, 630, 650, 651
RVM subject heading	28	Subject-RVM	Subject headings from Répertoire des vedettes- matière – maintained by the Bibliothèque de l'Université de Laval.	600i6, 610i6, 611i6, 630i6, 650i6, 651i6
Local subject index	29	Subject-local	Subjects headings defined locally.	
Date	30	Date	The point of time at which a transaction or event takes place.	005, 008/00-05, 008/07-10, 260\$c, 008/11-14, 033, etc
Date of publication	31	Date-publication	The date (usually year) in which a document is published.	008/07-10, 260\$c 046, 533\$d
Date of acquisition	32	Date-acquisition	The date when a document was acquired.	541\$d
Title-key	33	Title-key	The unique name assigned to a serial by the International Serials Data System (ISDS).	222
Title collective	34	Title-collective	A title proper that is an inclusive title for an item containing several works.	243
Title parallel	35	Title-parallel	The title proper in another language and/or script.	246i1
Title cover	36	Title-cover	The title printed on the cover of an item as issued.	246i4

Name	Val	Name1	Text	USmarc
Title added- title-page	37	Title-added-title-page	A title on a title page preceding or following the title page chosen as the basis for the description of the item. It may be more general (e.g., a series title page), or equally general (e.g., a title page in another language).	246i5
Title caption	38	Title-caption	A title given at the beginning of the first page of the text.	246i6
Title running	39	Title-running	A title, or abbreviated title, that is repeated at the head or foot of each page or leaf.	246i7
Title spine	40	Title-spine	A title appearing on the spine of an item.	246i8
Title other variant	41	Title-other-variant	A variation from the title page title appearing elsewhere in the item (e.g., a variant cover title, caption title, running title, or title from another volume) or in another issue.	212 (obs.), 46i3, 247, 740
Title former	42	Title-former	A former title or title variation when one bibliographic record represents all issues of a serial that has changed title.	247, 780
Title abbreviated	43	Title-abbreviated	Shortened form of the title; either assigned by national centers under the auspices of the International Serials Data System, or a title (such as an acronym) that is popularly associated with the item.	210, 211 (obs.), 246
Title expanded	44	Title-expanded	An expanded (or augmented) title has been enlarged with descriptive words by the cataloger to provide additional indexing and searching capabilities.	214 (obs.), 246
Subject PRECIS	45	Subject-PRECIS	Subject headings from PREserved Context Index System – a string of indexing terms set down in a prescribed order, each term being preceded by a manipulation code which governs the production of pre-coordinated subject index entries under elected terms – maintained by the British Library.	
Subject RSWK	46	Subject-RSWK	Subject headings from Regeln für den Schlagwortkatalog – maintained by the Deutsches Bibliotheksinstitut.	
Subject subdivision	47	Subject-subdivision	An extension to a subject heading indicating the form, place, period of time treated, or aspect of the subject treated.	6XX\$, 6XX\$, 6XX\$z

Name	Val	Name1	Text	USmarc
Number natl bibliography	48	Identifier-national-bibliography	Character string that uniquely identifies a record in a national bibliography.	015
Number legal deposit	49	Identifier-legal-deposit	The copyright registration number that is assigned to an item when the item is deposited for copyright.	017
Number govt publication	50	Classification-government-publication	A classification number assigned to a government document by a government agency at any level (e.g., state, national, international).	086
Number publisher for music	51	Identifier-publisher-for-music	A formatted number assigned by a publisher to a sound recording or to printed music.	028
Number DB	52	Control-number-DB	Character string that uniquely identifies a record in the Deutsche Bibliothek.	015
Number local call	53	Identifier-local-call	Call number (e.g., shelf location) assigned by a local system (not a classification number).	
Code-language	54	Code-language	A code that indicates the language of the item. The codes are defined by the target.	008/35-37, 041
Code-geographic area	55	Code-geographic-area	A code that indicates the geographic area(s) that appear or are implied in the headings assigned to the item during cataloging.	043
Code-institution	56	Code-institution	An authoritative-agency symbol for an institution. The code space is defined by the target.	040
Name and title	57	Name and title	The name of a person, corporate body, conference, or meeting, and the title of an item. (Subject name headings are included.) The syntax of the name-title combination is up to the target, unless used with the Structure attribute Key (see below).	100/2XX, 110/2XX, 111/2XX, subfields \$a & \$t in following: 400,410 411, 600, 610, 611 700, 710, 711, 800 810, 811

Name	Val	Name1	Text	USmarc
Name geographic	58	Name-geographic	Name of a country, jurisdiction, region, or geographic feature.	651
Place publication	59	Name-geographic-place-publication	City or town where an item was published.	008/15-17, 260\$a
CODEN	60	Identifier-CODEN	A six-character, unique, alphanumeric code assigned to serial and monographic publications by the CODEN section of the Chemical Abstracts Service.	030
Microform generation	61	Code-microform-generation	The code specifying the generation of a microform.	007/11
Abstract	62	Abstract	An abbreviated, accurate representation of a work, usually without added interpretation or criticism.	520
Note	63	Note	A concise statement in which such information as extended physical description, relationship to other works, or contents may be recorded.	5XX
Author-title	1000	Author-name-and-title	A personal or corporate author, title or a conference or meeting name, and the title of the item. (No subject name following: headings are included.) The syntax of the name-title combination is up to the target, unless used with the Structure attribute Key (see below)	100/2XX, 110/2XX, 111/2XX, subfields \$a & \$t in following: 400,410, 411, 700, 710, 711, 800, 810, 811
Record type	1001	Code-record-type	A code that specifies the characteristics and defines the components of the record. The codes are target-specific.	Leader/06
Name	1002	Name	The name of a person, corporate body, conference, or meeting. (Subject name headings are included.)	100, 110, 111, 400 410, 411, 600, 610 611, 700, 710, 711 800, 810, 811

Name	Val	Name1	Text	USmarc
Author	1003	Author-name	A personal or corporate author, or a conference or meeting name. (No subject name headings are included.)	100, 110, 111, 400 410, 411, 700, 710, 711, 800, 810, 811
Author-name personal	1004	Author-name-personal	A person's real name, pseudonym, title of nobility nickname, or initials. (Differs from attribute "name-personal (1)" in that personal name subject headings are not included.)	100, 400, 700, 800
Author-name corporate	1005	Author-name-corporate	An organization or a group of persons that is identified by a particular name. (Differs from attribute "name-corporate (2)" in that corporate name subject headings are not included.)	110, 410, 710, 810
Author-name conference	1006	Author-name-conference	A meeting of individuals or representatives of various bodies for the purpose of discussing topics of common interest. (Differs from attribute "name-conference (3)" in that conference name subject headings are not included.)	111, 411, 711, 811
Identifier- standard	1007	Identifier-standard	Standard numbers such as ISBN, ISSN, music publishers numbers, CODEN, etc., that are indexed together in many online public-access catalogs.	010, 011, 015, 017 018, 020, 022, 023 024, 025, 027, 028 030, 035, 037
Subject-LC children's	1008	Subject-LC-children's	Subject headings, for use with children's literature, that conform to the formulation guidelines in the "AC Subject Headings" section of the Library of Congress Subject Headings.	600i1, 610i1, 611i1, 630i1, 650i1, 651i1
Subject name-personal	1009	Subject-name-personal	A person's real name, pseudonym, title of nobility nickname, or initials that appears in a subject heading.	600
Body of text	1010	Body of text	Used in full-text searching to indicate that the term is to be searched only in that portion of the record that the target considers the body of the text, as opposed to some other discriminated part such as a headline, title, or abstract.	
Date/time added to database	1011	Date/time added to database	The date and time that a record was added to the database.	008/00-05

Name	Val	Name1	Text	USmarc
Date/time last modified	1012	Date/time last modified	The date and time a record was last updated.	005
Authority / format identifier	1013	Identifier-authority / format	Used in full-text searching to indicate to the target system the format of the document that should be returned to the originating system. The attribute carries not only the format code, but also the authority (e.g., system) that assigned that code.	
Concept-text	1014	Concept-text	Used within Z39.50-1988; included here for historical reasons but its use is deprecated.	
Concept-reference	1015	Concept-reference	Used within Z39.50-1988; included here for historical reasons but its use is deprecated.	
Any	1016	Any		520
Server choice	1017	Server-choice	The target substitutes one or more access points. The origin leaves the choice to the target.	
Publisher	1018	Name-publisher	The organization responsible for the publication of the item.	260\$b
Record source	1019	Record-source	The USMARC code or name of the organization(s) that created the original record, assigned the USMARC content designation and transcribed the record into machine-readable form, or modified the existing USMARC record.	008/39, 040
Editor	1020	Name-editor	A person who prepared for publication an item that is not his or her own.	100 \$a or 700 \$a
Bib-level	1021	Bib-level	A one-character alphabetic code indicating the bibliographic level such as monograph, serial or collection of the record.	Leader/07
Geographic class	1022	Geographic-class	A code that represents the geographic area and if applicable the geographic subarea covered by an item. The codes are derived from the LC Classification-Class G and the expanded Cutter number list.	052
Indexed by	1023	Indexed-by	For serials, a publication in which the serial has been indexed and/or abstracted.	510

Name	Val	Name1	Text	USmarc
Map scale	1024	Map-scale	Coded form of cartographic mathematical data, including scale, projection and/or coordinates related to the item.	034
Music key	1025	Music-key	A statement of the key in which the music is written.	\$r in the following 130, 240, 243, 630
Related periodical	1026	Related-periodical	Serial titles related to this item, either the immediate predecessor or the immediate successor.	247, 780, 785
Report number	1027	Identifier-report	A report number assigned to the item. This number could be the STRN (Standard Technical Report Number) or another report number. Cf. ANSI/NISO Z39.23 and ISO 10444.	027, 088
Stock number	1028	Identifier-stock	A stock number that could be used for ordering the item.	037
Thematic number	1030	Identifier-thematic	The numeric designation for a part/section of a work such as the serial, opus or thematic index number.	\$n in the following 130, 240, 243, 630 700, 730
Material type	1031	Material-type	A free-form string, more specific than the one-letter code in Leader/06, that describes the material type of the item, e.g., cassette, kit, computer database, computer file.	derived value from Leader/06-07, 007, 008, and 502
Doc ID	1032	Identifier-document	A persistent identifier, or Doc-ID, assigned by a server, that uniquely identifies a document on that server.	
Host item	1033	Host-item	The item containing the part described in the record, for example, a journal title when the record describes an article in the journal.	773
Content type	1034	Content-type	The type of materials contained in the item or publication. For example: review, catalog, encyclopedia, directory.	derived value from 008/24-27
Anywhere	1035	Anywhere	The record is selected if the term value (as qualified by the other attributes) occurs anywhere in the record.	
Author-Title-Subject	1036	Author-Title-Subject	An author or a title or a subject.	

Name	Val	Text
SICI	1037	The Serial Item and Contribution Identifier; based on NISO Z39.56, which provides an extensible mechanism for the creation of a code which uniquely identifies either an issue of a serial title or a contribution (e.g., article) within a serial regardless of distribution medium (paper, electronic, microform, etc.)

Following (thru 1083 submitted by German Library):

Name	Val	Text
Abstract-language	1038	The language code or the language name for the language of the abstract. Entries are made in accordance with ISO 639-1988 'Code for the representation of names of language'. The 2-digit letter code is used as language code, the English term is used as language name.
Application-kind	1039	Information (code or full text) about the patent application.
Classification	1040	The code or the section heading of a classification system. The attribute is a general (broad) attribute for codes which are not existing as a search key.
Classification-basic	1041	The code or the section heading of the Dutch classification system.
Classification-local-record	1042	The code or the section heading of a classification system assigned to a holding record.
Enzyme	1043	The enzyme code or the enzyme nomenclature. Codes are assigned by the 'Enzyme Commission' in the 'International Union of Biochemistry'.
Possessing-institution	1044	A code (library symbol or other code) of the institution which possesses the document. The codes are agreed upon between the partners.
Record-linking	1045	An alphabetical code which determines the type of linking.
Record-status	1046	Information about the status of the record, e.g. new, corrected, deleted, revised.
Treatment	1047	A statement (code or full text) describing subject aspects of the content.
Control-number-GKD	1048	The identification number of a corporate body name from the German authority file for corporate body names "Gemeinsame Koerper-schaftsdatei" (GKD).
Control-number-linking	1049	The unique identification number of the linked record. Attribute 'code-record-type' contains a code for the type of record which contains the 'control-number-linking' (b – bibliographical record, e – record for copy-specific data, n – authority record, l – holding record, p – record for cross references, v – full text). Attribute 'code-record-linking' contains a code which determines the type of linking.

Control-number-PND	1050	The identification number of a personal name from the German authority file for personal names "Personennamendatei" (PND).
Control-number-SWD	1051	The identification number of a subject heading from the German authority file for subject headings "Schlagwortnormdatei" (SWD).
Control-number-ZDB	1052	The identification number of a document in the German database for serials "Zeitschriftendatenbank" (ZDB).
Country-publication (country of Publication)	1053	The country code or the country name of the country where the document has been published. Entries are made according to ISO 3166 'Codes for the Representation of Names of Countries'. As country code a 2-digit letter code is used, as country name the English country name.
Date-conference (meeting date)	1054	The date of the conference or of another meeting.
Date-record-status	1055	The date on which the record status was assigned.
Dissertation-information	1056	Information about a dissertation thesis, or another publication connected with an academic degree.
Meeting-organizer	1057	The name of the organizer or the sponsor of a conference.
Note-availability	1058	Information about the availability of a document (delivery information).
Number-CAS-registry (CAS registry number)	1059	The 'Chemical Abstract Registry' number of the substance described in a document.
Number-document (document number)	1060	The publication number of the document (e.g. the number of the abstract in a secondary publication, or the number of the manuscript in a primary journal) provided that this number is not used as 'internal key' in the database. A publication number used as 'internal key' is entered in Attribute 12.
Number-local-accounting	1061	The account number of the document assigned by the accounting system.
Number-local-acquisition	1062	Document acquisition number assigned by the system.
Number-local-call-copy-specific	1063	The document's shelf number.
Number-of-reference (reference count)	1064	The number of literature references cited in a document.
Number-norm	1065	The number of a norm or standard.
Number-volume	1066	The number of single volumes of a multivolume publication, (year's) issues of serials, parts of multivolume publications, serials journals etc.
Place-conference (meeting location)	1067	The place where the conference or meeting was held.

Reference (references and footnotes)	1068	A literature reference from a document. The bibliographic data of a reference consists e.g. of a title, author, journal, publication year, volume and page information of the cited document.
Referenced-journal (reference work)	1069	The title of a journal cited in a document.
Section-code	1070	The section code of a subject classification.
Section-heading	1071	The section heading of a subject classification
Subject-GOO	1072	A subject heading from the 'Gemeenschappelijke Onderwerps Ontsluiting' (GOO).
Subject-name-conference	1073	A conference name used as subject heading.
Subject-name-corporate	1074	A corporate body name used as a subject heading.
Subject-genre/form (Formerly "Subject-name-form" renamed August 2003)	1075	A formal topical subject heading, e.g. collection of articles, handbook, source.
Subject-name-geographical	1076	A geographical/ethnographical subject heading.
Subject-chronological (formerly "Subject-name-chronological" renamed August 2003)	1077	A chronological subject heading.
Subject-title (formerly "Subject-name-title" renamed August 2003)	1078	A title proper used as subject heading.
Subject-topical (formerly "Subject-name-topical" renamed August 2003)	1079	A topical subject heading.
Subject-uncontrolled	1080	An uncontrolled subject term. General (broad) attribute for subject terms not specified any further, existing as a search key.
Terminology-chemical (chemical name)	1081	The description of a chemical substance. This is either the name of the substance or a name from another classification system, e.g. enzyme code.
Title-translated	1082	The translation of a title.
Year-of-beginning	1083	The publication year of the first issue/volume of serial publications (journals, newspapers, etc.).

Following (thru 1096) submitted by Danish National Library Authority, approved at January 1998 ZIG meeting:

Name	Val	Text
Year-of-ending	1084	The publication year of the last issue/volume of serial publications (journals, newspapers, etc).
Subject-AGROVOC	1085	A subject heading from the multilingual agricultural thesaurus from FAO.
Subject-COMPASS	1086	A subject heading from Computer Aided Subject System from British Library.
Subject-EPT	1087	A subject heading from European Pedagogical Thesaurus.
Subject-NAL	1088	A subject heading from National Agricultural Library.
Classification-BCM	1089	A classification number from British Catalogue of Music.
Classification-DB	1090	A classification number from Deutsche Bibliothek.
Identifier-ISRC	1091	ISRC. International standard recording code (ISO 3901).
Identifier-ISMN	1092	International standard music number (ISO 10957). ISMN
Identifier-ISRN	1093	ISRN. International standard technical report number (ISO 10444).
Identifier-DOI	1094	Digital Object Identifier.
Code-language-original	1095	A code that indicates the original language of the item.
Title-later	1096	A later version of title.

Following (1097 thru 1111) are Dublin Core Use attributes, originally approved at June 1998 ZIG meeting. These were re-affirmed at October 98 meeting, however, according to the decision at the meeting, the semantics (as well as mappings) are whatever a server decides. For detailed semantics of the Dublin Core elements, see Description of Dublin Core Elements

Name	Val		Name	Val
DC-Title	1097		DC-Language	1105
DC-Creator	1098		DC-OtherContributor	1106
DC-Subject	1099		DC-Format	1107
DC-Description	1100		DC-Source	1108
DC-Publisher	1101		DC-Relation	1109
DC-Date	1102		DC-Coverage	1110
DC-ResourceType	1103		DC-RightsManagement	1111
DC-ResourceIdentifier	1104			

The following (1112 thru 1184) are GILS Use attributes, approved at the June 1998 ZIG meeting. The GILS Use Attributes correspond semantically to GILS Core Elements, described in the GILS Z39.50 Specification.

Name	Val	Name	Val
Controlled Subject Index	1112	Available Time Period	1149
Subject Thesaurus	1113	Available Time Textual	1150
Index Terms – Controlled	1114	Available Time Structured	1151
Controlled Term	1115	Available Linkage	1152
Spatial Domain	1116	Linkage Type	1153
Bounding Coordinates	1117	Linkage	1154
West Bounding Coordinate	1118	Sources of Data	1155
East Bounding Coordinate	1119	Methodology	1156
North Bounding Coordinate	1120	Access Constraints	1157
South Bounding Coordinate	1121	General Access Constraints	1158
Place	1122	Originator Dissemination Control	1159
Place Keyword Thesaurus	1123	Security Classification Control	1160
Place Keyword	1124	Use Constraints	1161
Time Period	1125	Point of Contact	1162
Time Period Textual	1126	Contact Name	1163
Time Period Structured	1127	Contact Organization	1164
Beginning Date	1128	Contact Street Address	1165
Ending Date	1129	Contact City	1166
Availability	1130	Contact State or Province	1167
Distributor	1131	Contact Zip or Postal Code	1168
Distributor Name	1132	Contact Country	1169
Distributor Organization	1133	Contact Network Address	1170
Distributor Street Address	1134	Contact Hours of Service	1171
Distributor City	1135	Contact Telephone	1172
Distributor State or Province	1136	Contact Fax	1173
Distributor Zip or Postal Code	1137	Supplemental Information	1174
Distributor Country	1138	Purpose	1175
Distributor Network Address	1139	Agency Program	1176
Distributor Hours of Service	1140	Cross Reference	1177
Distributor Telephone	1141	Cross Reference Title	1178
Distributor Fax	1142	Cross Reference Relationship	1179
Resource Description	1143	Cross Reference Linkage	1180
Order Process	1144	Schedule Number	1181
Order Information	1145	Original Control Identifier	1182
Cost	1146	Language of Record	1183
Cost Information	1147	Record Review Date	1184
Technical Prerequisites	1148		

Following were approved at the October '98 ZIG meeting.

These were proposed on behalf of Music Libraries Online Project (<http://www.musiconline.uce.ac.uk/>)

Name	Val	Semantics
Performer	1185	Names of the Performer/Conductor (e.g. Alfred Brendel)
Performer-Individual	1186	Individual name of a Performer/Conductor (e.g. Alfred Brendel)
Performer-Group	1187	Name of a performance group or ensemble (e.g. London Symphony Orchestra)
Instrumentation	1188	Musical instrument used in the work (e.g. Piano, Violin)
Instrumentation-Original	1189	Musical instrument used in the original arrangement of the piece
Instrumentation-Current	1190	Musical instrument used in the given arrangement of the piece
Arrangement	1191	Arrangement of the work (e.g. Piano Trio, String Quartet)
Arrangement-Original	1192	Arrangement of the original work (e.g. Piano Trio, String Quartet)
Arrangement-Current	1193	Arrangement of the given arrangement (e.g. Piano Trio, String Quartet)
Musical Key-Original	1194	Key of the original arrangement
Musical Key-Current	1195	Key of the current arrangement
Date-Composition	1196	Date of the original composition of the work
Date-Recording	1197	Date of the current recording of the work
Place-Recording	1198	Place where the recording was made (e.g. Royal Festival Hall)
Country-Recording	1199	Country where the recording was made (ISO)
Number-ISWC	1200	International Standard Work Code (ISO/TC 46/SC 9/WG)
Number-Matrix	1201	Matrix number of recording
Number-Plate	1202	Publisher's plate number
Classification-McColvin	1203	McColvin Classification Scheme
Duration	1204	Duration of the work (or estimated duration) in seconds
Number-Copies	1205	Number of copies available (in the case of searching for performance sets for an ensemble)
Musical Theme	1206	A representation of a musical theme either as string format agreed between target and origin (e.g. Plaine and Easie) or as an EXTERNAL
Instruments – total number	1207	Total number of instruments (e.g. for a quartet, value would be 4)

Instruments – distinct number	1208	Number of distinct instruments. (e.g. for a string quartet with two violins, 1 viola, 1 cello, value would be 3) Following was approved at the August '99 ZIG meeting; proposed by Juha Hakala, Helsinki University Library.
Identifier – URN	1209	Uniform Resource Name

Following was approved at the December 2000 ZIG meeting; proposed by TZIG.

Name	Val	Semantics
Sears Subject Heading	1210	Sears List of Subject Headings

Following was approved at the October 2001 ZIG meeting; proposed by Dana Dietz on behalf of NISO SC AV.

Name	Val	Semantics
OCLC Number	1211	A sequential accession number assigned by OCLC as records are entered into OCLC WorldCat (the OCLC Online Union Catalog). Numbers are not re-used, but when records are merged, the old number is kept as a cross-reference to the new record

The following are based on a proposal from NORZIG, March 2003, approved August 2003.

Name	Val	Definition	Syntax / reference	Area of use
Composition	1212	Type of composition [Mus]	Two-letter Code. See MARC 21 008 pos.18-19	Searching for music (scores and sound recordings)
Intellectual level	1213	Target audience such as toddlers, preschools, etc.	One character Code. See MARC 21 008 pos.22	Limiting factor in literary searches
EAN	1214	European article number (EAN)	A 13-digit number where 1-3: type of merchandise 4-12: ISBN minus check digit (if type is book) 13: check digit. See UNIMARC 073 \$a and NORMARC 025 \$a	Searching for specific items
NLC	1215	Canadian class. system	Code consisting of capital letter(s), digits and punctuation. See MARC 21 055	Subject search

CRCS	1216	CRCS (Computing Review Classification Scheme)	Code of form: X.n.m Where X is a capital letter, n and m etc. are digits and the full stops indicate the hierarchy. See MARC 21 084 \$a and \$2 and NORMARC 062 \$a	Subject search
Nationality	1217	nationality (author)	Code consisting of two lowercase letters. See UNIMARC 7xx \$c and NORMARC 1xx \$j	Qualifying names
Equinox	1218	equinox (Map)	Name of equinox. See MARC 21 255 \$e	Limiting factor in searching for maps
Compression	1219	specific type of compression	MARC 21 856 \$c	Limiting search result
Format	1220	electronic format	MARC 21 856 \$q	Limiting search result
Subject – occupation	1221	Subject added entry – occupation	Index term. See MARC 21 656 \$a	Subject search
Subject – function	1222	Subject added entry – function	Index term. See MARC 21 657 \$a	Subject search
Edition	1223	Edition of document	Text. See MARC 21 250 \$a	Limiting search result

Номера 5000-10000 предназначены для локального использования и никогда не будут назначаться Агенством. Однако этих номеров применение не позволяет информационным системам быть в полной мере интероперабельными.

A.1.2 Relation attributes

Value	Relation
1	Less than
2	Less than or equal
3	Equal
4	Greater or equal
5	Greater than
6	Not equal
100	Phonetic
101	Stem
102	Relevance
103	AlwaysMatches

A.1.3 Position attributes

Position	Val	Definition
First in field	1	Search term must be the first data in the field
First in subfield	2	Search term may appear in any subfield but must be the first data in the subfield in which it appears
Any position in field	3	Search term may appear any place in the field

A.1.4 Structure attributes

Structure	Val	Definition
Phrase	1	A phrase consists of one or more groups of characters separated by blanks (for example, ASCII hex "20"). The value to be searched is exactly as it appears in the search term with respect to order and adjacency. Word(s) in the phrase may be explicitly truncated. (See "Truncation") To indicate that additional words may appear in the access point, use the completeness attribute.
Word	2	A word consists of a group of non-blank characters. It specifies the exact text of the value to be searched, unless the word is explicitly truncated. (See "Truncation") A word search term contains no blanks.
Key	3	A key specifies a sequence of characters extracted from those characters contained in an indexed word but not necessarily representing complete words. In the term, key segments should be separated by a blank (ASCII hex "20"). Each key segment should be the length of the local key or the length of the word, to a maximum of 6 characters. (For example, an name/title derived key search term for "Copland, Aaron, 1900-Rodeo" could be "coplan rodeo".) A segment may be adjusted by the receiving agency to the length required for the receiver's indexes. For example, the following derived key searches are in use at LC and at OCLC (in Online System)
Year	4	A year search term is numeric and contains four digits.
Date (normalized)	5	The day, month, year and time when a transaction or event takes place. The date search term structure is YYYYMMDDHHMMSSZ (YYYY = year; MM = month; DD = day; HH = hour; MM = minute; SS = second; Z = the time represented is Universal Time). The first four characters (YYYY) are mandatory.
Word list	6	A word list consists of one or more words separated by blanks (for example, ASCII hex "20"). No order of the words is implied. The attributes (other than structure) that are associated with the search term apply to each word in the word list. Any words in a word list may be explicitly truncated. (See "Truncation")

Date (un-normalized)	100	The day, month, and year when a transaction or event takes place. The un-normalized search term is unstructured.
Name (normalized)	101	A name search term that is structured in a particular order (e.g., last_name, first_name). The resulting term is subject to special matching rules on the target system that differ from those applied to names structured as phrases or unstructured names.
Name (un-normalized)	102	A name search term that is unstructured (e.g., first_name last_name), however, the resulting term is subject to matching rules on the target system that differ from those applied to phrases or structured names (e.g., the term "john smith" might be searched by the target as "smith, j#").
Structure	103	The term has a structure that is either implied by the Use attribute or defined by the target.
Urx	104	The term is a document identifier, for example, an identifier extracted from a Z39.50 URL.
Free-form-text	105	The term is text, input by the end user. May be used, for example, for relevance feedback.
Document-text	106	The term is text, extracted from a document. May be used, for example, for relevancefeedback.
Local-number	107	A number significant to the target.
String	108	The entire term is to be treated as a string, rather than a sequence or set of individual words
Numeric string	109	The term is a character string that represents a number

A.1.5 Truncation attributes

Truncation	Val	Definition
Right truncation	1	Word or Phrase: Last word of term is right truncated String: Entire term is right truncated. Word list: Each word is right truncated.
Left truncation	2	Word or Phrase: First word of term is left truncated. String: Entire term is left truncated. Word list: Each word is left truncated.
Left and right truncation	3	Word or Phrase: First word of term is left truncated and last word of term is right truncated. String: Entire term is left and right truncated. Word list: Each word is left and right truncated.
Do not truncate	100	No truncation is to be applied.
Process # in search term	101	The search term contains the symbol "#" (ASCII hex "23") to show where truncation will take place (e.g., "National H# Institute", or "d#on").
RegExpr-1	102	The term is in the form of a regular expression as prescribed by IEEE 1003.2 Volume 1, Section 2.8 "Regular Expression Notation".

RegExpr-2	103	The term is in the form of a regular expression whose format is target-defined.
-----------	-----	---

A.1.6 Completeness attributes

Completeness	Val	Definition
Incomplete subfield	1	Words other than those in the search term may appear in the subfield or field in which the term appears.
Complete subfield	2	No words other than those in the search term should appear in the entire subfield in which the term appears, but additional words may appear in other subfields in the field.
Complete field	3	No words other than those in the search term should appear in the entire field in which the term appears.

A.2 Атрибуты Exp-1

A.2.1 Use attributes

Значение	Название	Поисковый терм
1	ExplainCategory	TargetInfo, DatabaseInfo, SchemaInfo, TagSetInfo, RecordSyntaxInfo, AttributeSetInfo, TermListInfo, extendedServicesInfo, AttributeDetails, ElementSetDetails, RetrievalRecordDetails, SortDetails, Processing, CategoryList, VariantSetInfo, UnitInfo
2	HumanStringLanguage	Трёхсимвольные коды языков из ANSI/NISO Z39.53-1994
3	DatabaseName	
4	TargetName	
5	AttributeSetOID	OID
6	RecordSyntaxOID	OID
7	TagSetOID	OID
8	ExtendedServiceOID	OID
9	DateAdded	
10	DateChanged	
11	DateExpires	
12	ElementSetName	
13	ProcessingContext	Access, Search, Retrieval, RecordPresentation, RecordHandling
14	ProcessingName	
15	TermListName	

16	SchemaOID	OID
17	Producer	
18	Supplier	
19	Availability	
20	Proprietary	
21	UserFree	
22	VariantSetOID	OID
23	UnitSystem	
25	Keyword	
26	ExplainDatabase	
27	ProcessingOID	OID

A.3 Атрибуты GILS

A.3.1 Use attributes

Значение	Название
2000	Distributor
2003	Purpose
2004	General Access Constraints
2005	Use Constraints
2006	Distributor Organization
2007	Distributor Street Address
2008	Distributor City
2009	Distributor State or Province
2010	Distributor Zip or Postal Code
2011	Distributor Country
2012	Distributor Network Address
2013	Distributor Hours of Service
2014	Distributor Telephone
2015	Distributor Fax
2016	Resource Description
2017	Order Information
2018	Technical Prerequisites
2019	Available Time Structured
2020	Available Time Textual
2021	Linkage
2022	Linkage Type
2023	Contact Name

2024	Contact Organization
2025	Contact Street Address
2026	Contact City
2027	Contact State or Province
2028	Contact Zip or Postal Code
2029	Contact Country
2030	Contact Network Address
2031	Contact Hours of Service
2032	Contact Telephone
2033	Contact Fax
2034	Agency Program
2035	Sources of Data
2036	Subject Thesaurus
2037	Methodology
2038	West Bounding Coordinate
2039	East Bounding Coordinate
2040	North Bounding Coordinate
2041	South Bounding Coordinate
2042	Place Keyword
2043	Place Keyword Thesaurus
2044	Time Period Structured
2045	Time Period Textual
2046	Cross Reference Title
2047	Cross Reference Linkage
2049	Original Control Identifier
2050	Supplemental Information
2051	Record Review Date
2052	Originator Dissemination Control
2053	Security Classification Control
2054	Cost
2055	Cost Information
2056	Schedule Number
2057	Controlled Subject Index
2058	Uncontrolled Term
2059	Spatial Domain
2060	Bounding Coordinates
2061	Place
2062	Time Period

2063	Availability
2064	Order Process
2065	Available Time Period
2066	Access Constraints
2067	Point of Contact
2068	Cross Reference
2069	Available Linkage
2070	Cross Reference Relationship
2071	Language of Record
2072	Beginning Date
2073	Ending Date
2074	Controlled Term

A.4 Атрибуты Ext-1

Value	Attribute Name
1	Use
2	Permissions

A.4.1 Use attributes

Value	Attribute Name
1	UserId
2	PackageName
3	CreationDatetime
4	TaskStatus
5	PackgeType
6	RetentionTime
7	TargetReference

A.4.2 Permission attributes

Атрибуты Permission используются только, если Use есть UserId.

Value	Attribute Name
1	Delete
2	Modify
3	ModifyPermissions
4	Present

5	Invoke
6	Any

A.5 Атрибуты XD-1

Набор атрибутов XD-1 [53] (Cross-Domain Attribute Set) соответствует атрибутам Dublin Core. Согласно новой архитектуре поисковых атрибутов [54] типу Use соответствуют тип Access Point.

A.5.1 Access Point (Use) attributes

Val	Attribute	Meaning
1	Title	Имя ресурса
2	Subject	Тема ресурса
3	Name	Персональное имя или название организации, ассоциированные с ресурсом
4	Description	Описание ресурса
5	Date	Дата, связанная с созданием или готовностью ресурса
6	Resource Type	Характер или жанр ресурса
7	Format	Физическое или цифровое представление ресурса
8	Resource Identifier	Однозначный идентификатор ресурса в пределах заданного контекста
9	Source	Идентификатор ресурса-источника
10	Language	Язык интеллектуального содержания ресурса
11	Relation	Идентификатор связанного ресурса
12	Coverage	Область ресурса
13	Rights Management	Информация относительно прав

A.6 Атрибуты Utility

Комментарий к атрибутам оставлен без перевода для более точной передачи смысла [81].

A.6.1 Access Point (Use) attributes

Val	Attribute	Meaning
1	Record Date/time	A date (or date and time) associated with the record. May be used in conjunction with a functional qualifier, such as: 'creation' (to mean "date/time of creation of record"), 'modification' (to mean "date/time record last modified"), etc., or it may be used unqualified, in which case it is a date/time assigned by the server to be associated with the record.
2	Record Agent	Identifies a person, institution, or process that acts upon the database record. May be qualified by a semantic qualifier to indicate "person", "institution", or "process" and may be qualified by a functional qualifier, for example, 'creation', to indicate the type of action.
3	Language of Record	Search terms used with this access point should be taken from "ISO 639-2:1998. Codes for the Representation of Names of Languages – Part 2: Alpha-3 Code."
4	Local Control Number of Record	An identifier string (integer or character string) assigned by the server that uniquely identifies a record in the database
5	Cost to Retrieve Record	The cost (expressed as IntUnit) that the client will incur if it subsequently retrieves the full record. This access point may be used for example to limit a search to records where the cost of retrieval is less than a specified amount. (When the server does not charge for retrieval a cost of zero may be assumed.)
6	Record Syntax	A record syntax (expressed as an object identifier) supported for the database record. This access point may be used to limit the results to include records deliverable in a specific syntax. (Thus a search might create a large result set, where all except a few records are available only in a format that the client cannot support; the client may use this access point to narrow the result set to records it can support.) Note that the use of this attribute in a search does not mean that the resulting records, when subsequently retrieved, will necessarily be supplied in the given syntax, unless the present request explicitly specifies that syntax
7	Database Schema	This access point may be used to limit the results to include records available in a specific database schema (expressed as an object identifier).
8	Score	See Rank

9	Rank	<p>Access points Score and Rank pertain to relevancy ranking for ranked retrieval. The server may assign a score and/or a rank to a result set record. The score or rank applies to the record relative to other records in the result set. A score is an integer between zero and 100. The rank of a record is an integer from 1 to N, where N is the result set size. (Note that a higher score means more relevant, while a lower rank means more relevant.) It is assumed that if record A has a better (higher) score than record B, then it will also have a better (lower) rank, however score and rank differ in the following respect: no two records in the result set have the same rank, and for every integer between 1 and N (where N is the result set size) there is exactly one record with that rank; on the other hand, more than one result set record may share the same score, and there need not be a record for every possible score. The reason for defining both attributes (Score and Rank) is that some ranked retrieval systems score records while others rank records. These two attributes may be used to restrict a result to some threshold value, for example "records with a score greater than 50" or "records with a rank less than 10".</p>
10	Result Set Position	<p>This attribute may be used to limit the size of the result set. For example the operand: Result Set Position "less than or equal" 10 could be included in the query to restrict the result set to 10 or less records. Note that the Rank access point can be used for the same purpose; this attribute is defined because a server might support this function though it may not support ranking. The rank of a record is not necessarily its result set position; furthermore, the result set position of a record might change (if the result set is sorted) but its rank should not change.</p>
11	All Access Points	<p>When the origin uses 'All Access Points' it is asking the server to search for the term via all supported access points</p>
12	Anywhere in Record	<p>When the origin uses 'Anywhere in Record' it is asking the server to search for the term anywhere it may occur within the record, subject to the server's search capability and interpretation of what "anywhere within the record" means. The server might search commonly used access points, or it might search the entire content of the record</p>
13	Server Choice	<p>When the origin uses 'Server-choice' it is asking the server to select an access point (which may be defined in any attribute set, not necessarily the Utility set), and to use its best judgment in making that selection</p>
15	Wildcard	<p>May be used for a single wild card (within a field path specified by nested access points). Thus: "field-1 within wildcard within field-3" would match "field-1 within field-2 within field-3" and it would match "field-1 within field-4 within field-3" but would not match "field-1 within field-2 within field-4 within field-3".</p>

16	Wildpath	May be used for an unanchored search. Thus: "field-1 within wildpath" would match "field-1", or "field-1 within field-2", or "field-1 within field-2 within field-3", etc.; "field-1 within wildpath within field-2" would match "field-1 within field-3 within field-4 within field-2".
----	----------	--

A.7 Атрибуты Zthes-1

Комментарий к атрибутам оставлен без перевода для более точной передачи смысла.

A.7.1 Access Point (Use) attributes

Val	Attribute	Description
1	termQualifier	searches in the termQualifier element of the top-level term record
2	termType	searches in the termType element of the top-level term record
3	thesAdmin	used for a variety of searches related to administrative details of thesaurus structure
4	relatedTermID	used in conjunction with a semantic qualifier (attribute type 2) with value equal to one of the relationTypes; searches for all records in the specified relation to the record whose termID is equal to the search term. For example, a search for abc123 with access point relatedTermID and semantic qualifier "NT" finds all the narrower terms of the record whose termID is abc123.

A.8 Атрибуты Collections-1

Комментарий к атрибутам оставлен без перевода для более точной передачи смысла.

A.8.1 Access Point (Use) attributes

Val	Attribute	Description
1	Record-id	
2	Record-type	0 = record is a Descriptive Record 1 = record is a Collection Descriptive Record 2 = record is a Collection Descriptive Record for a context collection 3 = record is an Object Descriptive Record 4 = record is a Digital Object
3	Object-type	1 = object is a Digital Object 2 = object is a Digital Object for which there is a Physical Object 3 = object is a Digital Object for which there is no Physical Object 4 = object is a Physical Object

4	Parent-collection	This attribute is used to search for Descriptive Records that list the collection (specified by the term) as a parent. Note: There is no requirement in this profile that a collection or object know all of its parents, but when it does not, full functionality might not be achieved. This attribute may be combined with Use attribute Record-type to restrict the search to Collection Descriptive Records or Object Descriptive Records.
5	Superior-collection	This attribute is similar to Parent-collection, used to search for subordinate, but not necessarily immediately subordinate, Descriptive Records
6	Related-collection	Used to search for Descriptive Records for objects or collections that list the collection (indicated by the term) as a related-collection
7	Collection-name	Used to search for the Collection Descriptive Record for the collection (indicated by the term)

A.9 Атрибуты СИМІ-1

Комментарий к атрибутам оставлен без перевода для более точной передачи смысла [66].

A.9.1 Access Point (Use) attributes

Val	Attribute	Val	Attribute
4	title	2041	creatorGeneral
7	ISBN	2042	associationGeneral
8	ISSN	2043	objectLanguage
31	date of publication	2044	condition
32	date of acquisition	2045	physicalDescription
54	code language	2046	who
58	name geographic	2047	what
62	abstract	2048	when
1003	author	2049	where
1004	personal author	2051	DC-title
1016	any	2052	DC-creator
1018	publisher	2053	DC-subject
1031	material type	2054	DC-description
1032	doc-id	2055	DC-publisher
2000	award	2056	DC-contributors
2002	collection	2057	DC-date
2004	copyrightRestriction	2058	DC-type
2005	creditLine	2059	DC-format

2007	inscriptionMark	2060	DC-identifier
2008	materialMedium	2061	DC-source
2009	creatorNationalCultRace	2062	DC-language
2012	processTechnique	2063	DC-relation
2014	creatorRole	2064	DC-coverage
2017	stylePeriod	2065	DC-rights
2020	image	2070	fieldCollector
2022	dateOfOrigin	2071	dateCollected
2023	placeOfOrigin	2072	agePeriod
2024	objectID	2073	typeSpecimen
2026	owner	2074	dimensions
2027	repositoryName	2075	quantity
2028	repositoryPlace	2076	relatedObjects
2029	provenance	2077	resource
2030	contentGeneral	2078	wallTextLabel
2032	objectName	2079	administrativeEventGeneral
2033	objectTitle	2080	administrator
2034	relatedTextualReferences	3000	protectionStatus
2035	creatorName	3001	protectionDate
2036	creatorDateOfBirth	3003	spatialReferencingSystem
2037	creatorDateOfDeath	3004	x-coordinateInReferSystem
2038	contextHistorical	3005	y-coordinateInReferSystem
2039	contextArchaeological	3007	address
2040	subjectContent	3009	periodName

A.9.2 Authority (Type-101) attributes

Val	Attribute	Description
1	Non-authoritative	The client explicitly states that the term is not taken from any authoritative list
2	Local-to-server	The term is known to the client to come from an authoritative source defined by the server
3	USMARC	The term is a code taken from USMARC manuals or associated documents, such as the set of coded values for countries, languages, etc.
4	LCSH	The term is from Library of Congress Subject Headings
5	AAT	The term is from Art & Architecture Thesaurus (AAT)

6	AAT Date	The term is from "Date and Geographic Name Guidelines" in Appendix A of Chapter 3, Guide to Indexing and Cataloging with the Art & Architecture Thesaurus (AAT)
7	ACRL/RBMS Binding	The term is from Binding Terms: A Thesaurus for Use in Rare Book and Special Collections Cataloging
8	ACRL/RBMS Genre	The term is from Genre Terms: A Thesaurus for Use in Rare Book and Special Collections Cataloging
9	ACRL/RBMS Paper	The term is from Paper Terms: A Thesaurus for Use in Rare Book and Special Collections Cataloging
10	ACRL/RBMS Printing	The term is from Printing and Publishing Evidence: A Thesaurus for Use in Rare Book and Special Collections Cataloging
11	ACRL/RBMS Type	The term is from Type Evidence: A Thesaurus for Use in Rare Book and Special Collections Cataloging
12	Base Merimee	The term is from BaseMerimee: Lexique d'interrogation pour les champs: denomination, parties constituantes, ouvrages remarquables
13	BGN	The term is from the Board on Geographic Names
14	British Archaeological	The term is from British Archaeological Thesaurus: For Use with British Archaeological Abstracts, and Other Publications in British Archaeology
15	Canadiana	The term is from Canadiana Authorities / Canadiana, vedettes d'autorite
16	Dictionarium Museologicum	The term is from Dictionarium Museologicum
17	Garnier	The term is from Thesaurus iconographique: syst?me descriptif des representations
18	Geosaurus	The term is from Geosystems' Thesaurus of Geoscience
19	Glass	The term is from A Subject Index for the Visual Arts
20	ICOM Costume	The term is from Vocabulary of Basic Terms for Cataloging Costume / Vocabulaire de base pour les fichiers de costume
21	ICONCLASS	The term is from ICONCLASS: An Iconographical Classification System
22	Jewish Art	The term is from Index of Jewish Art: An Iconographical Index of Hebrew Illuminated Manuscripts
23	ISO Language	The term is from ISO 639: Codes for the Representation of Names of Languages/ Codes pour la representation des noms de langue

24	ISO Documentation	The term is from ISO 5127-1: Documentation and Information. Vocabulary. Part 1, Basic Concepts / Documentation et information. Vocabulaire. Partie 1, Notions fondamentales
25	ISO Iconic	The term is from ISO 5127-3: Documentation and Information. Part 3, Iconic Documents / Documentation et information. Vocabulaire. Partie 3, Documents iconiques
26	ISO AV	The term is from ISO 5127-11: Documentation and Information. Vocabulary. Part 11, Audio-visual Documents / Documentation et information. Vocabulaire. Partie 11, Documents audiovisuels
27	ISO Date/Time	The term is from ISO 8601: Data Elements and Interchange Formats. Information Interchange. Representation of Dates and Times
28	LC Descriptive Graphic	The term is from Descriptive Terms for Graphic Materials: Genre and Physical Characteristic Headings
29	LC Name	The term is from Library of Congress Name Authorities
30	LC Thesaurus Graphic	The term is from LC Thesaurus for Graphic Materials: Topical Terms for Subject Access
31	Moving Image Materials	The term is from Moving Image Materials: Genre Terms
32	Nomenclature	The term is from The Revised Nomenclature for Museum Cataloging: A Revised and Expanded Version of Robert G. Chenhall's System for Classifying Man-made Objects
33	Reynies	The term is from Le Mobilier domestique; vocabulaire typologique
34	TGN	The term is from Thesaurus of Geographic Names
35	Tozzer	The term is from Tozzer Index to Anthropological Subject Headings, Harvard University
36	ULAN	The term is from Union List of Artist Names
37	Villard	The term is from systeme descriptif des antiquites classiques
38	Yale British Artists	The term is from British Artists Authority List
1000	RCHME	The term is from the Royal Commission on the Historical Monuments of England

A.9.3 Charset (Type-103) attributes

Val	Attribute	Description
0	7-bit US-ASCII	ANSI X3.4-1986 (R1997): Information Systems – Coded Character Sets – 7-Bit American National Standard Code for Information Interchange (7-Bit ASCII)
1	ISO 8859-1 (Latin-1)	ANSI/ISO 8859-1-1987: Information Processing – 8-Bit Single Byte Coded Graphic Character Sets – Part 1: Latin Alphabet No. 1
2	ISO 8859-2 (Latin-2)	ISO 8859-2:1987: Information processing – 8-bit single byte coded graphic character sets – Part 2: Latin alphabet No. 2
3	ISO 8859-3 (Latin-3)	ISO 8859-3:1988: Information processing – 8-bit single-byte coded graphic character sets – Part 3: Latin alphabet No. 3
4	ISO 8859-4 (Latin-4)	ISO/IEC 8859-4:1998: Information technology – 8-bit single-byte coded graphic character sets – Part 4: Latin alphabet No. 4
5	ISO 8859-5 (Cyrillic)	ISO/IEC 8859-5:1988: Information processing – 8-bit single-byte coded graphic character sets – Part 5: Latin/Cyrillic alphabet
6	ISO 8859-6 (Arabic)	ISO 8859-6:1987: Information processing – 8-Bit single-byte coded graphic character sets – Part 6: Latin/Arabic alphabet
7	ISO 8859-7 (Greek)	ISO 8859-7:1987: Information processing – 8-bit single-byte coded graphic character sets – Part 7: Latin/Greek alphabet
8	ISO 8859-8 (Hebrew)	ISO 8859-8:1988: Information processing – 8-bit single-byte coded graphic character sets – Part 8: Latin/Hebrew alphabet
9	ISO 8859-9 (Latin-5)	ISO/IEC 8859-9:1989: Information processing – 8-bit single-byte coded graphic character sets – Part 9: Latin alphabet No. 5
10	ISO/IEC 8859-10 (Latin-6)	ISO/IEC 8859-10:1998: Information technology – 8-bit single-byte coded graphic character sets – Part 10: Latin alphabet No. 6
11	ISO/IEC 10646-1/AMD2 (UTF-8)	ISO/IEC 10646-1/AMD2:1996: Amendment 2 to ISO/IEC 10646-1:1993 UCS Transformation Format 8 (UTF-8)
12	ISO/IEC 10646-1/AMD1 (UTF-16)	ISO/IEC 10646-1/AMD1:1996: Amendment 1 to ISO/IEC 10646-1:1993 Transformation Format for 16 planes of group 00 (UTF-16)

A.10 Атрибуты GEO-1

Комментарий к атрибутам оставлен без перевода для более точной передачи смысла [60].

A.10.1 Access Point (Use) attributes

Val	Name	Tag	Description	Type	Domain
4	Title	title	the name by which the data set is known	text	free text
5	Series Name	sername	the name of the series publication of which the dataset is a part	text	free text
31	Publication Date	pubdate	the date when the data set is published or otherwise made available for release	date	"Unknown", "Unpublished material", free date
59	Publication Place	pubplace	the name of the city (and state or province, and country, if needed to identify the city) where the data set was published or released	text	free text
62	Abstract	abstract	a brief narrative summary of the data set	text	free text
1005	Originator	origin	the name of an organization or individual that developed the data set. If the name of editors or compilers are provided, the name must be followed by "(ed.)" or "(comp.)" respectively	text	"Unknown", free text
1012	Metadata Date	metd	the date that the metadata were created or last updated	date	free date
1016	Any				
1018	Publisher	publish	the name of the individual or organization that published the data set	text	free text
1024	Source Scale Denominator	srescale	the denominator of the representative fraction on a map (for example, on a 1:24,000-scale map, the Source Scale Denominator is 24000)	integer	Source Scale Denominator > 1

1031	Type of Source Media	typesrc	the medium of the source data set	text	Возможные значения приведены в Примечании 1031
1035	Anywhere				
2000	Distributor	distrib	the party from whom the data set may be obtained		
2002	Theme Keyword	themekey	common-use word or phrase used to describe the subject of the data set	text	free text
2003	Purpose	purpose	a summary of the intentions with which the data set was developed	text	free text
2004	Access Constraints	accesscons	restrictions and legal prerequisites for accessing the data set. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the data set	text	"None", free text
2005	Use Constraints	useconst	restrictions and legal prerequisites for using the data set after access is granted. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the data set	text	"None", free text
2013	Hours of Service	hours	time period when individuals can speak to the organization or individual	text	free text
2016	Resource Description	resdesc	the identifier by which the distributor knows the data set	text	free text
2017	Ordering Instructions	ordering	general instructions and advice about, and special terms and services provided for, the data set by the distributor	text	free text
2018	Technical Prerequisites	techpreq	description of any technical capabilities that the consumer must have to use the data set in the form(s) provided by the distributor	text	free text
2021	Online Linkage	onlink	the name of an online computer resource that contains the data set. Entries should follow the Uniform Resource Locator convention of the Internet	text	free text

2023	Contact Person	cntper	the name of the individual to which the contact type applies	text	free text
2024	Contact Organization	cntorg	the name of the organization to which the contact type applies	text	free text
2025	Address	address	an address line for the address	text	free text
2026	City	city	the city of the address	text	free text
2027	State or Province	state	the state or province of the address	text	free text
2028	Postal Code	postal	the ZIP or other postal code of the address	text	free text
2029	Country	country	the country of the address	text	free text
2030	Contact Electronic Mail Address	cntemail	the address of the electronic mailbox of the organization or individual	text	free text
2032	Contact Voice Telephone	cntvoice	the telephone number by which individuals can speak to the organization or individual	text	free text
2033	Contact Facsimile Telephone	cntfax	the telephone number of a facsimile machine of the organization or individual	text	free text
2035	Source Contribution	srctr	brief statement identifying the information contributed by the source to the data set	text	free text
2036	Theme Keyword Thesaurus	themekt	reference to a formally registered thesaurus or a similar authoritative source of theme keywords	text	"None" free text
2038	West Bounding Coordinate	westbc	western-most coordinate of the limit of coverage expressed in longitude	real	-180.0 <= West Bounding Coordinate < 180.0

2039	East Bounding Coordinate	eastbc	eastern-most coordinate of the limit of coverage expressed in longitude	real	-180.0 <= East Bounding Coordinate <= 180.0
2040	North Bounding Coordinate	northbc	northern-most coordinate of the limit of coverage expressed in latitude	real	-90.0 <= North Bounding Coordinate <= 90.0; North Bounding Coordinate >= South Bounding Coordinate
2041	South Bounding Coordinate	southbc	southern-most coordinate of the limit of coverage expressed in latitude	real	-90.0 <= South Bounding Coordinate <= 90.0; South Bounding Coordinate <= North Bounding Coordinate
2042	Place Keyword	placekey	the geographic name of a location covered by a data set	text	free text
2043	Place Keyword Thesaurus	placekt	reference to a formally registered thesaurus or a similar authoritative source of place keywords	text	"None", "Geographic Names Information System", free text
2045	Temporal Keyword	tempkey	the name of a time period covered by a data set	text	free text
2050	Supplemental Information	supplinf	other descriptive information about the data set	text	free text
2055	fees	fees	the fees and terms for retrieving the data set	text	free text
2059	Spatial Domain	spdom	the geographic areal domain of the data set		

2060	Bounding Coordinates	bounding	the limits of coverage of a data set expressed by latitude and longitude values in the order western-most, eastern-most, northern-most, and southern-most. For data sets that include a complete band of latitude around the earth, the West Bounding Coordinate shall be assigned the value -180.0, and the East Bounding Coordinate shall be assigned the value 180.0		
2061	Place	place	geographic locations characterized by the data set		
2062	Time Period Information	timeinfo			
2065	Available Time Period	availabl	the time period when the data set will be available from the distributor		
2067	Point of Contact	ptcontac	contact information for an individual or organization that is knowledgeable about the data set		
2068	Cross Reference	crossref	information about other, related data sets that are likely to be of interest		
2072	Beginning Date	begdate	the first year (and optionally month, or month and day) of the event	date	"Unknown"; "Present"; free date
2073	Ending Date	enddate	the last year (and optionally month, or month and day) for the event	date	"Unknown"; "Present"; free date
3000	Contact Information	cntinfo	This section provides a means of identifying individuals and organizations, and is used by other sections of the metadata standard. This section is never used alone		
3004	Contact Person Primary	cntperp	the person, and the affiliation of the person, associated with the data set. Used in cases where the association of the person to the data set is more significant than the association of the organization to the data set		
3005	Contact Position	cntpos	the title of individual	text	free text

3006	Contact Address	cntaddr	the address for the organization or individual			
3007	Address Type	addrtype	the information provided by the address	text	"mailing address", "physical address", "mailing and physical address"	
3008	Contact Organization Primary	cntorgp	the organization, and the member of the organization, associated with the data set. Used in cases where the association of the organization to the data set is more significant than the association of the person to the data set			
3014	Contact TDD/TTY Telephone	cnttdd	the telephone number by which hearing-impaired individuals can contact the organization or individual	text	free text	
3018	Contact Instructions	cntinst	supplemental instructions on how or when to contact the individual or organization	text	free text	
3100	Identification Information	idinfo				
3101	Citation	citation	information to be used to reference the data set			
3102	Description	descript	a characterization of the data set, including its intended use and limitations			
3106	Currentness Reference	current	the basis on which the time period of content information is determined	text	"ground condition", "publication date", free text	
3107	Status	status	the state of and maintenance information for the data set			
3108	Progress	progress	the state of the data set	text	"Complete", "In work", "Planned"	

3109	Maintenance and Update Frequency	update	the frequency with which changes and additions are made to the data set after the initial data set is completed	text	"Continually", "Daily", "Weekly", "Monthly", "Annually", "Unknown", "As needed", "Irregular", "None planned", free text
3116	Data Set G-Polygon	dsgpoly	coordinates defining the outline of an area covered by a data set		
3117	Data Set G-Polygon Outer G-Ring	dsgpolyo	the closed nonintersecting boundary of an interior area		
3118	G-Ring Latitude	gringlat	the latitude of a point of the G-ring	real	-90.0 <= G-Ring Latitude <= 90.0
3119	G-Ring Longitude	gringlon	the longitude of a point of the G-ring	real	-180.0 <= G-Ring Latitude < 180.0
3120	Data Set G-Polygon Exclusion G-Ring	dsgpolyx	"the closed nonintersecting boundary of a void area (or 'hole') in an interior area"		
3121	Keywords	keywords	words or phrases summarizing an aspect of the data set		
3122	Theme	theme	subjects covered by the data set (for a list of some commonly-used thesauri, see Part IV: Subject/index term sources in Network Development and MARC Standards Office, 1988, USMARC code list for relators, sources, and description conventions: Washington, Library of Congress)		
3128	Stratum	stratum	layered, vertical locations characterized by the data set		
3129	Stratum Keyword Thesaurus	stratkt	reference to a formally registered thesaurus or a similar authoritative source of stratum keywords	text	"None" free text"

3130	Stratum Keyword	stratkey	the name of a vertical location used to describe the locations covered by a data set	text	free text
3131	Temporal	temporal	time period(s) characterized by the data set		
3132	Temporal Keyword Thesaurus	tempkey	reference to a formally registered thesaurus or a similar authoritative source of temporal keywords	text	"None", free text
3137	Browse Graphic	browse	a graphic that provides an illustration of the data set. The graphic should include a legend for interpreting the graphic		
3138	Browse Graphic File Name	browse	name of a related graphic file that provides an illustration of the data set	text	free text
3139	Browse Graphic File Description	browse	a text description of the illustration	text	free text
3140	Browse Graphic File Type	browse	graphic file type of a related graphic file	text	См. Приложение 3140
3141	Data Set Credit	datacred	recognition of those who contributed to the data set	text	free text
3142	Security Information	secinfo	handling restrictions imposed on the data set because of national security, privacy, or other concerns		
3143	Security Classification System	secsys	name of the classification system	text	free text
3144	Security Classification	secclass	name of the handling restrictions on the data set	text	"Top secret", "Secret", "Confidential", "Restricted", "Unclassified", "Sensitive", free text

3145	Security Handling Description	sechandl	additional information about the restrictions on handling the data set	text	free text
3146	Native Data Set Environment	native	a description of the data set in the producer's processing environment, including items such as the name of the software (including version), the computer operating system, file name (including host-, path-, and filenames), and the data set size	text	free text
3148	Extent	extent			
3200	Data Quality Information	dataqual			
3201	Attribute Accuracy	attracc	an assessment of the accuracy of the identification of entities and assignment of attribute values in the data set		
3202	Attribute Accuracy Report	attraccr	an explanation of the accuracy of the identification of the entities and assignments of values in the data set and a description of the tests used	text	free text
3203	Quantitative Attribute Accuracy Assessment	qatraccc	a value assigned to summarize the accuracy of the identification of the entities and assignments of values in the data set and the identification of the test that yielded the value		
3204	Attribute Accuracy Value	attraccv	an estimate of the accuracy of the identification of the entities and assignments of attribute values in the data set	text	"Unknown", free text
3205	Attribute Accuracy Explanation	attraccex	the identification of the test that yielded the Attribute Accuracy Value	text	free text
3206	Logical Consistency Report	logic	an explanation of the fidelity of the relationships in the data set and the tests used	text	free text

3207	Complete- ness Report	complete	information about omissions, selection criteria, generalization, definitions used, and other rules used to derive the data set	text	free text
3208	Positional Accuracy	posacc	an assessment of the accuracy of the positions of spatial objects		
3209	Horizontal Positional Accuracy	horizpa	an estimate of accuracy of the horizontal positions of the spatial objects		
3210	Horizontal Positional Accuracy Report	horizpar	an explanation of the accuracy of the horizontal coordinate measurements and a description of the tests used	text	free text
3211	Quantitative Horizontal Positional Accuracy Assessment	qhorizpa	numeric value assigned to summarize the accuracy of the horizontal coordinate measurements and the identification of the test that yielded the value		
3212	Horizontal Positional Accuracy Value	horizpav	an estimate of the accuracy of the horizontal coordinate measurements in the data set expressed in (ground) meters	real	free real
3213	Horizontal Positional Accuracy Explanation	horizpae	the identification of the test that yielded the Horizontal Positional Accuracy Value	text	free text
3214	Vertical Positional Accuracy	vertacc	an estimate of accuracy of the vertical positions in the data set		

3215	Vertical Positional Accuracy Report	vertaccr	an explanation of the accuracy of the vertical coordinate measurements and a description of the tests used	text	free text
3216	Quantitative Vertical Positional Accuracy Assessment	qvertpa	numeric value assigned to summarize the accuracy of vertical coordinate measurements and the identification of the test that yielded the value		
3217	Vertical Positional Accuracy Value	vertaccv	an estimate of the accuracy of the vertical coordinate measurement in the data set expressed in (ground) meters	real	free real
3218	Vertical Positional Accuracy Explanation	vertacce	the identification of the test that yielded the Vertical Positional Accuracy Value	text	free text
3219	Lineage	lineage	information about the events, parameters, and source data which constructed the data set, and information about the responsible parties		
3220	Source Information	srcinfo	list of sources and a short discussion of the information contributed by each		
3221	Source Citation	srccite	reference for a source data set		
3223	Source Time Period of Content	srctime	time period(s) for which the source data set corresponds to the ground		
3224	Source Currentness Reference	srccurr	the basis on which the source time period of content information of the source data set is determined	text	"ground condition", "publication date", free text

3225	Source Citation Abbreviation	srccitea	short-form alias for the source citation	text	free text
3227	Process Step	procestp	information about a single event		
3228	Process Description	procdesc	an explanation of the event and related parameters or tolerances	text	free text
3229	Source Used Citation Abbreviation	srcused	the Source Citation Abbreviation of a data set used in the processing step	text	Source Citation Abbreviations from the Source Information entries for the data set
3230	Process Date	procdat	the date when the event was completed	date	"Unknown", "Not complete", free date
3231	Process Time	proctime	the time when the event was completed	time	free time
3232	Source Produced Citation Abbreviation	srcprod	the Source Citation Abbreviation of an intermediate data set that (1) is significant in the opinion of the data producer, (2) is generated in the processing step, and (3) is used in later processing steps	text	Source Citation Abbreviations from the Source Information entries for the data set
3233	Process Contact	procont	the party responsible for the processing step information		
3234	Cloud Cover	cloud	area of a data set obstructed by clouds, expressed as a percentage of the spatial extent	integer	'0 <= Cloud Cover <= 100 "Unknown"
3300	Spatial Data Organization Information	spdoinfo			
3301	Indirect Spatial Reference	indspref	name of types of geographic features, addressing schemes, or other means through which locations are referenced in the data set	text	free text

3302	Direct Spatial Reference Method	direct	the system of objects used to represent space in the data set	text	"Point", "Vector", "Raster"
3303	SDTS Terms Description	sdtstern	"point and vector object information using the terminology and concepts from "Spatial Data Concepts," which is chapter 2 of part 1 in Department of Commerce, 1992, Spatial Data Transfer Standard (SDTS) (Federal Information Processing Standard 173): Washington, Department of Commerce, National Institute of Standards and Technology. (Note that this reference to the SDTS is used ONLY to provide a set of terminology for the point and vector objects.)"		
3304	SDTS Point and Vector Object Type	sdtstype	name of point and vector spatial objects used to locate zero-, one-, and two-dimensional spatial locations in the data set	text	См. Примечание 3304.
3305	Point and Vector Object Count	ptvctcnt	the total number of the point or vector object type occurring in the data set	integer	Point and Vector Object Count > 0
3306	VPF Terms Description	vpfterm	point and vector object information using the terminology and concepts from Department of Defense, 1992, Vector Product Format (MIL-STD-600006): Philadelphia, Department of Defense, Defense Printing Service Detachment Office. (Note that this reference to the VPF is used ONLY to provide a set of terminology for the point and vector objects.)		
3307	VPF Topology Level	vpflevel	the completeness of the topology carried by the data set. The levels of completeness are defined in Department of Defense, 1992, Vector Product Format (MIL-STD-600006): Philadelphia, Department of Defense, Defense Printing Service Detachment Office	integer	0 <= VPF Topology Level <= 3
3308	VPF Point and Vector Object Type	vpftype	name of point and vector spatial objects used to locate zero-, one-, and two-dimensional spatial locations in the data set	text	Смю Примечание 3308.

3309	Raster Object Information	rastinfo	the types and numbers of raster spatial objects in the data set		
3310	Raster Object Type	rasttype	raster spatial objects used to locate zero-, two-, or three-dimensional locations in the data set	text	См. Примечание 3310.
3311	Row Count	rowcount	the maximum number of raster objects along the ordinate (y) axis. For use with rectangular raster objects	integer	Row Count > 0
3312	Column Count	colcount	the maximum number of raster objects along the abscissa (x) axis. For use with rectangular raster objects	integer	Column Count > 0
3313	Vertical Count	vrtcount	the maximum number of raster objects along the vertical (z) axis. For use with rectangular volumetric raster objects (voxels)	integer	Depth Count > 0
3314	Point and Vector Object Information	ptvctinf	the types and numbers of vector or nongridded point spatial objects in the data set		
3400	Spatial Reference Information	spref			
3401	Horizontal Coordinate System Definition	horizsys	the reference frame or system from which linear or angular quantities are measured and assigned to the position that a point occupies		
3402	Geographic	geograph	the quantities of latitude and longitude which define the position of a point on the Earth's surface with respect to a reference spheroid		
3403	Latitude Resolution	latres	the minimum difference between two adjacent latitude values expressed in Geographic Coordinate Units of measure	real	Latitude Resolution > 0.0
3404	Longitude Resolution	longres	the minimum difference between two adjacent longitude values expressed in Geographic Coordinate Units of measure	real	Longitude Resolution > 0.0

3405	Geographic Coordinate Units	geogunit	units of measure used for the latitude and longitude values	text	См. Примечание 3405.
3406	Planar	planar	the quantities of distances, or distances and angles, which define the position of a point on a reference plane to which the surface of the Earth has been projected		
3407	Map Projection	mapproj	the systematic representation of all or part of the surface of the Earth on a plane or developable surface		
3408	Map Projection Name	mapprojn	name of the map projection	text	См. Примечание 3408.
3410	Standard Parallel	stdparll	line of constant latitude at which the surface of the Earth and the plane or developable surface intersect	real	$-90.0 \leq \text{Standard Parallel} \leq 90.0$
3411	Longitude of Central Meridian	longcm	the line of longitude at the center of a map projection generally used as the basis for constructing the projection	real	$-180.0 \leq \text{Longitude of Central Meridian} < 180.0$
3412	Latitude of Projection Origin	latprjo	latitude chosen as the origin of rectangular coordinates for a map projection	real	$-90.0 \leq \text{Latitude of Projection Origin} \leq 90.0$
3413	False Easting	feast	"the value added to all "x" values in the rectangular coordinates for a map projection. This value frequently is assigned to eliminate negative numbers. Expressed in the unit of measure identified in Planar Coordinate Units"	real	free real
3414	False Northing	fnorth	"the value added to all "y" values in the rectangular coordinates for a map projection. This value frequently is assigned to eliminate negative numbers. Expressed in the unit of measure identified in Planar Coordinate Units"	real	free real
3415	Scale Factor at Equator	sfequat	a multiplier for reducing a distance obtained from a map by computation or scaling to the actual distance along the equator	real	Scale Factor at Equator > 0.0

3416	Height of Perspective Point Above Surface	heightpt	height of viewpoint above the Earth, expressed in meters	real	Height of Perspective Point Above Surface > 0.0
3417	Longitude of Projection Center	longpc	longitude of the point of projection for azimuthal projections	real	-180.0 <= Longitude of Projection Center < 180.0
3418	Latitude of Projection Center	latprjc	latitude of the point of projection for azimuthal projections	real	-90.0 <= Latitude of Projection Center <= 90.0
3419	Scale Factor at Center Line	sfctrln	a multiplier for reducing a distance obtained from a map by computation or scaling to the actual distance along the center line	real	Scale Factor at Center Line > 0.0
3420	Oblique Line Azimuth	obqlazim	method used to describe the line along which an oblique mercator map projection is centered using the map projection origin and an azimuth		
3421	Azimuthal Angle	azimangl	angle measured clockwise from north, and expressed in degrees	real	0.0 <= Azimuthal Angle < 360.0
3422	Azimuth Measure Point Longitude	azimptl	longitude of the map projection origin	real	-180.0 <= Azimuth Measure Point Longitude < 180.0
3423	Oblique Line Point	obqlpt	method used to describe the line along which an oblique mercator map projection is centered using two points near the limits of the mapped region that define the center line		
3424	Oblique Line Latitude	obqllat	latitude of a point defining the oblique line	real	-90.0 <= Oblique Line Latitude <= 90.0
3425	Oblique Line Longitude	obqllong	longitude of a point defining the oblique line	real	-180.0 <= Oblique Line Longitude < 180.0

3426	Straight Vertical Longitude from Pole	svlong	longitude to be oriented straight up from the North or South Pole	real	-180.0 <= Straight Vertical Longitude from Pole < 180.0
3427	Scale Factor at Projection Origin	sfprjorg	a multiplier for reducing a distance obtained from a map by computation or scaling to the actual distance at the projection origin	real	Scale Factor at Projection Origin > 0.0
3428	Landsat Number	landsat	number of the Landsat satellite. (Note: This data element exists solely to provide a parameter needed to define the space oblique mercator projection. It is not used to identify data originating from a remote sensing vehicle.)	integer	0 < Landsat Number < 5
3429	Path Number	pathnum	number of the orbit of the Landsat satellite. (Note: This data element exists solely to provide a parameter needed to define the space oblique mercator projection. It is not used to identify data originating from a remote sensing vehicle.)	integer	'0 < Path Number < 251 for Landsats 1, 2, or 3; 0 < Path Number < 233 for Landsats 4 or 5"
3430	Scale Factor at Central Meridian	sfctrmer	a multiplier for reducing a distance obtained from a map by computation or scaling to the actual distance along the central meridian	real	Scale Factor at Central Meridian > 0.0
3431	Other Projection's Definition	otherprj	a complete description of a projection, not defined elsewhere in this standard, that was used for the data set. The information provided shall include the name of the projection, the names of the parameters and values used for the data set, and the citation of the specification for the algorithms that describe the mathematical relationship between the Earth and the plane or developable surface for the projection	text	free text
3432	Grid Coordinate System	gridsys	a plane-rectangular coordinate system usually based on, and mathematically adjusted to, a map projection so that geographic positions can be readily transformed to and from plane coordinates		

3433	Grid Coordinate System Name	gridsysn	name of the grid coordinate system	text	См. Примечание 3433.
3434	Universal Transverse Mercator	utm	a grid system based on the transverse mercator projection, applied between latitudes 84 degrees north and 80 degrees south on the Earth's surface		
3435	UTM Zone Number	utmzone	identifier for the UTM zone	integer	"1 <= UTM Zone Number <= 60 for the northern hemisphere; -60 <= UTM Zone Number <= -1 for the southern hemisphere"
3436	Universal Polar Stereographic	ups	a grid system based on the polar stereographic projection, applied to the Earth's polar regions north of 84 degrees north and south of 80 degrees south		
3437	UPS Zone Identifier	upszone	identifier for the UPS zone	text	"A", "B", "Y", "Z"
3438	State Plane Coordinate System	spcs	a plane-rectangular coordinate system established for each state in the United States by the National Geodetic Survey		
3439	SPCS Zone Identifier	spcszone	identifier for the SPCS zone	text	См. Примечание 3439.
3440	ARC Coordinate System	arcsys	the Equal Arc-second Coordinate System, a plane-rectangular coordinate system established in Department of Defense, 1990, Military specification ARC Digitized Raster Graphics (ADRG) (MIL-A-89007): Philadelphia, Department of Defense, Defense Printing Service Detachment Office		
3441	ARC System Zone Identifier	arczone	identifier for the ARC Coordinate System Zone	integer	1 <= ARC System Zone Identifier <= 18

3442	Other Grid System's Definition	othergrid	a complete description of a grid system, not defined elsewhere in this standard, that was used for the data set. The information provided shall include the name of the grid system, the names of the parameters and values used for the data set, and the citation of the specification for the algorithms that describe the mathematical relationship between the Earth and the coordinates of the grid system	text	free text
3443	Local Planar	localp	any right-handed planar coordinate system of which the z-axis coincides with a plumb line through the origin that locally is aligned with the surface of the Earth		
3444	Local Planar Description	localpd	a description of the local planar system	text	free text
3445	Local Planar Georeference Information	localpgi	a description of the information provided to register the local planar system to the Earth (e.g. control points, satellite ephemeral data, inertial navigation data)	text	free text
3446	Planar Coordinate Information	planci	information about the coordinate system developed on the planar surface		
3447	Planar Coordinate Encoding Method	plance	the means used to represent horizontal positions	text	"coordinate pair", "distance and bearing", "row and column"
3448	Coordinate Representation	coordrep	"the method of encoding the position of a point by measuring its distance from perpendicular reference axes (the "coordinate pair" and "row and column" methods)"		
3449	Abscissa Resolution	absres	"the (nominal) minimum distance between the "x" or column values of two adjacent points, expressed in Planar Distance Units of measure"	real	Abscissa Resolution > 0.0
3450	Ordinate Resolution	ordres	"the (nominal) minimum distance between the "y" or row values of two adjacent points, expressed in Planar Distance Units of measure"	real	Ordinate Resolution > 0.0

3451	Distance and Bearing Representation	distbrep	a method of encoding the position of a point by measuring its distance and direction (azimuth angle) from another point			
3452	Distance Resolution	distres	the minimum distance measurable between two points, expressed Planar Distance Units of measure	real	Distance Resolution > 0.0	
3453	Bearing Resolution	bearres	the minimum angle measurable between two points, expressed in Bearing Units of measure	real	Bearing Resolution > 0.0	
3454	Bearing Units	bearunit	units of measure used for angles	text	См. Примечание 3454.	
3455	Bearing Reference Direction	bearrefd	direction from which the bearing is measured	text	"North", "South"	
3456	Bearing Reference Meridian	bearrefm	axis from which the bearing is measured	text	"Assumed", "Grid", "Magnetic", "Astronomic", "Geodetic"	
3457	Planar Distance Units	plandu	units of measure used for distances	text	"meters", "international feet", "survey feet", free text	
3458	Local	local	a description of any coordinate system that is not aligned with the surface of the Earth			
3459	Local Description	localdes	a description of the coordinate system and its orientation to the surface of the Earth	text	free text	
3460	Local Georeference Information	localgeo	a description of the information provided to register the local system to the Earth (e.g. control points, satellite ephemeral data, inertial navigation data)	text	free text	
3461	Geodetic Model	geodetic	parameters for the shape of the earth			

3462	Horizontal Datum Name	horizdn	the identification given to the reference system used for defining the coordinates of points	text	"North American Datum of 1927", "North American Datum of 1983", free text
3463	Ellipsoid Name	ellips	identification given to established representations of the Earth's shape	text	"Clarke 1866", "Geodetic Reference System 80", free text
3464	Semi-major Axis	semiaxis	radius of the equatorial axis of the ellipsoid	real	Semi-major Axis > 0.0
3465	Denominator of Flattening Ratio	denflat	the denominator of the ratio of the difference between the equatorial and polar radii of the ellipsoid when the numerator is set to 1	real	Denominator of Flattening > 0.0
3466	Vertical Coordinate System Definition	vertdef	the reference frame or system from which vertical distances (altitudes or depths) are measured		
3467	Altitude System Definition	altsys	"the reference frame or system from which altitudes (elevations) are measured. The term "altitude" is used instead of the common term "elevation" to conform to the terminology in Federal Information Processing Standards 70-1 and 173"		
3468	Altitude Datum Name	altdatum	the identification given to the level surface taken as the surface of reference from which altitudes are measured	text	"National Geodetic Vertical Datum of 1929", "North American Vertical Datum of 1988", free text
3469	Altitude Resolution	altres	the minimum distance possible between two adjacent altitude values, expressed in Altitude Distance Units of measure	real	Altitude Resolution > 0.0
3470	Altitude Distance Units	altunits	units in which altitudes are recorded	text	"meters", "feet", free text

3471	Altitude Encoding Method	altenc	the means used to encode the altitudes	text	"Explicit elevation coordinate included with horizontal coordinates", "Implicit coordinate", "Attribute values"
3472	Depth System Definition	depthsys	the reference frame or system from which depths are measured		
3473	Depth Datum Name	depthdn	the identification given to surface of reference from which depths are measured	text	См. Примечание 3473.
3474	Depth Resolution	depthres	the minimum distance possible between two adjacent depth values, expressed in Depth Distance Units of measure	real	Depth Resolution > 0.0
3475	Depth Distance Units	depthdu	units in which depths are recorded	text	"meters", "feet", free text
3476	Depth Encoding Method	depthem	the means used to encode depths	text	"Explicit depth coordinate included with horizontal coordinates", "Implicit coordinate", "Attribute values"
3477	Albers Conical Equal Area	albers			
3478	Azimuthal Equidistant	azimequi			
3479	Equidistant Conic	equicon			
3480	Equirectangular	equirect			

3481	General Vertical Near-sided Projection	gvnsp			
3482	Gnomonic	gnomonic			
3483	Lambert Azimuthal Equal Area	lamberta			
3484	Lambert Conformal Conic	lambertc			
3485	Mercator	mercator			
3486	Modified Stereographic for Alaska	modsak			
3487	Miller Cylindrical	miller			
3488	Oblique Mercator	obqmerc			
3491	Orthographic	orthogr			
3492	Polar Stereographic	polarst			
3493	Polyconic	polycon			
3494	Robinson	robinson			
3495	Sinusoidal	sinusoid			

3496	Space Oblique Mercator (Landsat)	spaceobq			
3497	Stereographic	stereo			
3498	Transverse Mercator	transmer			
3499	van der Grinten	vdgrin			
3500	Entity and Attribute Information	eainfo			
3501	Detailed Description	detailed			description of the entities, attributes, attribute values, and related characteristics encoded in the data set
3502	Entity Type	enttyp			the definition and description of a set into which similar entity instances are classified
3503	Entity Type Label	enttyp1		text	the name of the entity type free text
3504	Entity Type Definition	enttypd		text	the description of the entity type free text
3505	Entity Type Definition Source	enttypds		text	the authority of the definition free text
3506	Attribute	attr			a defined characteristic of an entity
3507	Attribute Label	attrlabl		text	the name of the attribute free text
3508	Attribute Definition	attrdef		text	the description of the attribute free text

3509	Attribute Definition Source	attrdefs	the authority of the definition	text	free text
3510	Attribute Domain Values	attrdomv	the valid values that can be assigned for an attribute		
3511	Enumerated Domain	edom	the members of an established set of valid values		
3512	Enumerated Domain Value	edomv	the name or label of a member of the set	text	free text
3513	Enumerated Domain Value Definition	edomvd	the description of the value	text	free text
3514	Enumerated Domain Value Definition Source	edomvds	the authority of the definition	text	free text
3515	Range Domain	rdom	the minimum and maximum values of a continuum of valid values		
3516	Range Domain Minimum	rdommin	the least value that the attribute can be assigned	text	free text
3517	Range Domain Maximum	rdommax	the greatest value that the attribute can be assigned	text	free text

3518	Codeset Domain	codesetd	reference to a standard or list which contains the members of an established set of valid values			
3519	Codeset Name	codesetn	the title of the codeset	text	text	free text
3520	Codeset Source	codesets	the authority for the codeset	codesets	text	free text
3521	Unrepresentable Domain	udom	description of the values and reasons why they cannot be represented	udom	text	free text
3522	Attribute Units of Measurement	attrunit	the standard of measurement for an attribute value	attrunit	text	free text
3523	Attribute Measurement Resolution	attrmres	the smallest unit increment to which an attribute value is measured	attrmres	real	Attribute Measurement Resolution > 0.0
3524	Beginning Date of Attribute Values	begdatea	earliest or only date for which the attribute values are current. In cases when a range of dates are provided, this is the earliest date for which the information are valid	begdatea	date	free date
3525	Ending Date of Attribute Values	enddatea	latest date for which the information are current. Used in cases when a range of dates are provided	enddatea	date	free date
3526	Attribute Value Accuracy Information	attrvai	an assessment of the accuracy of the assignment of attribute values	attrvai		
3527	Attribute Value Accuracy	attrva	an estimate of the accuracy of the assignment of attribute values	attrva	real	free real

3528	Attribute Value Accuracy Explanation	attrvae	the definition of the Attribute Value Accuracy measure and units, and a description of how the estimate was derived	text	free text
3529	Attribute Measurement Frequency	attrmfreq	the frequency with which attribute values are added	real	"Unknown", "As needed", "Irregular", "None planned", free text
3530	Overview Description	overview	summary of, and citation to detailed description of, the information content of the data set		
3531	Entity and Attribute Overview	eaover	detailed summary of the information contained in a data set	text	free text
3532	Entity and Attribute Detail Citation	eadetcit	reference to the complete description of the entity types, attributes, and attribute values for the data set	text	free text
3600	Distribution Information	distinfo			
3603	Distribution Liability	distliab	statement of the liability assumed by the distributor	text	free text
3604	Standard Order Process	stdorder	the common ways in which the data set may be obtained or received, and related instructions and fee information		
3605	Non-digital Form	nondig	the description of options for obtaining the data set on non-computer-compatible media	text	free text
3606	Digital Form	digform	the description of options for obtaining the data set on computer-compatible media		

3607	Digital Transfer Information	digitinfo	description of the form of the data to be distributed			
3608	Format Name	formname	the name of the data transfer format	text	См. Примечание 3473.	
3609	Format Version Number	formvern	version number of the format	text	free text	
3610	Format Version Date	formverd	date of the version of the format	date	free date	
3611	Format Specification	formspec	name of a subset, profile, or product specification of the format	text	free text	
3612	Format Information Content	formcont	description of the content of the data encoded in a format	text	free text	
3613	File Decompression Technique	filedec	recommendations of algorithms or processes (including means of obtaining these algorithms or processes) that can be applied to read or expand data sets to which data compression techniques have been applied	text	"No compression applied", free text	
3614	Transfer Size	transize	the size, or estimated size, of the transferred data set in megabytes	real	Transfer Size > 0.0	
3615	Digital Transfer Option	digtopt	the means and media by which a data set is obtained from the distributor			
3616	Online Option	onlinopt	information required to directly obtain the data set electronically			
3617	Computer Contact Information	computer	instructions for establishing communications with the distribution computer			

3618	Network Address	networka	the electronic address from which the data set can be obtained from the distribution computer			
3619	Network Resource Name	networkr	the name of the file or service from which the data set can be obtained	text	text	free text
3620	Dialup Instructions	dialinst	information required to access the distribution computer remotely through telephone lines			
3621	Lowest BPS	lowbps	lowest or only speed for the connection's communication, expressed in bits per second	integer	integer	Lowest BPS ≥ 110
3622	Highest BPS	highbps	highest speed for the connection's communication, expressed in bits per second. Used in cases when a range of rates are provided	integer	integer	Highest BPS $>$ Lowest BPS
3623	Number DataBits	numdata	number of data bits in each character exchanged in the communication	integer	integer	$7 < \leq$ Number DataBits $< \leq 8$
3624	Number StopBits	numstop	number of stop bits in each character exchanged in the communication	integer	integer	$1 < \leq$ Number StopBits $< \leq 2$
3625	Parity	parity	parity error checking used in each character exchanged in the communication	text	text	"None", "Odd", "Even", "Mark", "Space"
3626	Compression Support	compress	data compression available through the modem service to speed data transfer	text	text	"V.32", "V.32bis", "V.42", "V.42bis", free text
3627	Dialup Telephone	dialtel	the telephone number of the distribution computer	text	text	free text
3628	Dialup File Name	dialfile	the name of a file containing the data set on the distribution computer	text	text	free text
3629	Access Instructions	accessinst	instructions on the steps required to access the data set	text	text	free text

3630	Online Computer and Operating System	oncomp	the brand of distribution computer and its operating system	text	free text
3631	Offline Option	offoptn	information about media-specific options for receiving the data set		
3632	Offline Media	offmedia	name of the media on which the data set can be received	text	См. Примечание 3632.
3633	Recording Capacity	reccap	the density of information to which data are written. Used in cases where different recording capacities are possible		
3634	Recording Density	recden	the density in which the data set can be recorded	real	Recording Density > 0.0
3635	Recording Density Units	recdenu	the units of measure for the recording density	text	free text
3636	Recording Format	refint	the options available or method used to write the data set to the medium	text	См. Примечание 3636.
3637	Compatibility Information	compat	description of other limitations or requirements for using the medium	text	free text
3640	Turnaround	turnarnd	typical turnaround time for the filling of an order	text	free text
3641	Custom Order Process	custom	description of custom distribution services available, and the terms and conditions for obtaining these services	text	free text
3700	Metadata Reference Information	metainfo			
3702	Metadata Review Date	metrd	the date of the latest review of the metadata entry	date	"free date; Metadata Review Date later than Metadata Date"

3703	Metadata Future Review Date	metfrd	the date by which the metadata entry should be reviewed	date	"free date; Metadata Future Review Date later than Metadata Review Date"
3704	Metadata Contact	metc	the party responsible for the metadata information		
3705	Metadata Standard Name	metstdn	the name of the metadata standard used to document the data set	text	"FGDC Content Standards for Digital Geospatial Metadata", free text
3706	Metadata Standard Version	metstdv	identification of the version of the metadata standard used to document the data set	text	free text
3707	Metadata Time Convention	mettc	form used to convey time of day information in the metadata entry. Used if time of day information is included in the metadata for a data set	text	"local time", "local time with time differential factor", "universal time"
3708	Metadata Access Constraints	metac	restrictions and legal prerequisites for accessing the metadata. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the metadata	text	free text
3709	Metadata Use Constraints	metuc	restrictions and legal prerequisites for using the metadata after access is granted. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the metadata	text	free text
3710	Metadata Security Information	metssi	handling restrictions imposed on the metadata because of national security, privacy, or other concerns		

3711	Metadata Security Classification System	metscs	name of the classification system for the metadata	text	free text
3712	Metadata Security Classification	metsc	name of the handling restrictions on the metadata	text	"Top secret", "Secret", "Confidential", "Restricted", "Unclassified", "Sensitive", free text
3713	Metadata Security Handling Description	metshd	additional information about the restrictions on handling the metadata	text	free text
3800	Citation Information	citeinfo			
3803	Publication Time	pubtime	the time of day when the data set is published or otherwise made available for release	time	"Unknown", free time
3805	Geospatial Data Presentation Form	geoform	the mode in which the geospatial data is represented	text	См. Примечание 3805.
3806	Series Information	serinfo	the identification of the series publication of which the data set is a part		
3807	Edition	edition	the version of the title	text	free text
3808	Issue Identification	issue	information identifying the issue of the series publication of which the data set is a part	text	free text
3809	Publication Information	pubinfo	publication details for published data sets		

3812	Other Citation Details	othercit	other information required to complete the citation	text	free text
3814	Larger Work Citation	lworkcit	the information identifying a larger work in which the data set is included		
3901	Time Period of Content	timeperd	time period(s) for which the data set corresponds to the ground		
3902	Single Date/Time	sngdate	means of encoding a single date and time		
3903	Calendar Date	caldate	the year (and optionally month, or month and day)	date	"Unknown", free date
3904	Time of Day	time	the hour (and optionally minute, or minute and second) of the day	time	"Unknown", free time
3905	Multiple Dates/Times	mdattim	means of encoding multiple individual dates and times		
3906	Range of Dates/Times	rngdates	means of encoding a range of dates and times		
3908	Beginning Time	begtime	the first hour (and optionally minute, or minute and second) of the day for the event	time	"Unknown", free time
3910	Ending Time	endtime	the last hour (and optionally minute, or minute and second) of the day for the event	time	"Unknown", free time
3999	Metadata	metadata			

Примечание 1031: "paper" "stable-base material", "microfiche", "microfilm", "audiocassette", "chart", "filmstrip", "transparency", "videocassette", "videodisc", "videotape", "physical model", "computer program", "disc", "cartridge tape", "magnetic tape", "online", "CD-ROM", "electronic bulletin board", "electronic mail system", free text

Примечание 3140: "domain values in the table below; free text", "CGM" – Computer Graphics Metafile, "EPS" – Encapsulated Postscript format, "GIF" – Graphic Interchange Format, "JPEG" – Joint Photographic Experts Group format, "PBM" – Portable Bit Map format, "PS" – Postscript format, "TIFF" – Tagged Image File Format, "XWD" – X-Windows Dump

Примечание 3140: (The domain is from 'Spatial Data Concepts', which is chapter 2 of part 1 in Department of Commerce, 1992, Spatial Data Transfer Standard (SDTS) (Federal Information Processing Standard 173): Washington, Department of Commerce, National Institute of Standards and Technology): "Point", "Entity point", "Label point", "Area point", "Node, planar graph", "Node, network", "String", "Link", "Complete chain", "Area chain", "Network chain, planar graph", "Network chain, nonplanar graph", "Circular arc, three point center", "Elliptical arc", "Uniform B spline", "Piecewise Bezier", "Ring with mixed, composition", "Ring composed of strings", "Ring composed of chains", "Ring composed of arcs", "G-polygon", "GT-polygon composed of rings", "GT-polygon composed of chains", "Universe polygon composed of rings", "Universe polygon composed of chains", "Void polygon composed of rings", "Void polygon composed of chains",

Примечание 3308: The domain is from Department of Defense, 1992, Vector Product Format (MIL-STD-600006): Philadelphia, Department of Defense, Defense Printing Service Detachment Office: "Node", "Edge", "Face", "Text".

Примечание 3310: With the exception of "voxel", the domain is from 'Spatial Data Concepts,' which is chapter 2 of part 1 in Department of Commerce, 1992, Spatial Data Transfer Standard (SDTS) (Federal Information Processing Standard 173): Washington, Department of Commerce, National Institute of Standards and Technology): "Point", "Pixel", "Grid Cell", "Voxel".

Примечание 3405: "Decimal degrees", "Decimal minutes", "Decimal seconds", "Degrees and decimal minutes", "Degrees, minutes, and decimal seconds", "Radians", "Grads".

Примечание 3408: "Albers Conical Equal Area", "Azimuthal Equidistant", "Equidistant Conic", "Equirectangular", "General Vertical Near-sided Projection", "Gnomonic", "Lambert Azimuthal Equal Area", "Lambert Conformal Conic", "Mercator", "Modified Stereographic for Alaska", "Miller Cylindrical", "Oblique Mercator", "Orthographic", "Polar

Stereographic", "Polyconic", "Robinson", "Sinusoidal", "Space Oblique Mercator", "Stereographic", "Transverse Mercator", "van der Grinten", "other projection"

Примечание 3433: "Universal Transverse Mercator", "Universal Polar Stereographic", "State Plane Coordinate System 1927", "State Plane Coordinate System 1983", "ARC Coordinate System", "other grid system".

Примечание 3439: Four-digit numeric codes for the State Plane Coordinate Systems based on the North American Datum of 1927 are found in Department of Commerce, 1986, Representation of geographic point locations for information interchange (Federal Information Processing Standard 70-1): Washington: Department of Commerce, National Institute of Standards and Technology. Codes for the State Plane Coordinate Systems based on the North American Datum of 1983 are found in Department of Commerce, 1989 (January), State Plane Coordinate System of 1983 (National Oceanic and Atmospheric Administration Manual NOS NGS 5): Silver Spring, Maryland, National Oceanic and Atmospheric Administration, National Ocean Service, Coast and Geodetic Survey.

Примечание 3454: "Decimal degrees", "Decimal minutes", "Decimal seconds", "Degrees and decimal minutes", "Degrees, minutes, and decimal seconds", "Radians", "Grads".

Примечание 3473: "Local surface", "Chart datum; datum for sounding reduction", "Lowest astronomical tide", "Highest astronomical tide", "Mean low water", "Mean high water", "Mean sea level", "Land survey datum", "Mean low water springs", "Mean high water springs", "Mean low water neap", "Mean high water neap", "Mean lower low water", "Mean lower low water springs", "Mean higher high water", "Mean higher low water", "Mean lower high water", "Spring tide", "Tropic lower low water", "Neap tide", "High water", "Higher high water", "Low water", "Low-water datum", "Lowest low water", "Lower low water", "Lowest normal low water", "Mean tide level", "Indian spring low water", "High-water full and charge", "Low-water full and charge", "Columbia River datum", "Gulf Coast low water datum", "Equatorial springs low water", "Approximate lowest astronomical tide", "No correction", free text.

Примечание 3608: Domain values from the table below; "free text", "ARCE" – ARC/INFO Export format, "ARCG" – ARC/INFO Generate format, "ASCII" – ASCII file, formatted for text attributes, declared format, "BIL" – Imagery, band interleaved by line, "BIP" – Imagery, band interleaved by pixel, "BSQ" – Imagery, band interleaved sequential, "CDF" – Common Data Format, "CFF" – Cartographic Feature File (U.S. Forest Service), "COORD" – User-created coordinate file, declared format, "DEM" – Digital Elevation Model format (U.S. Geological Survey), "DFAD" – Digital Feature Analysis Data (Defense Mapping Agency), "DGN" – Microstation format (Intergraph Corporation), "DIGEST" –

Digital Geographic Information Exchange Standard, "DLG" – Digital Line Graph (U.S. Geological Survey), "DTED" – Digital Terrain Elevation Data (MIL-D-89020), "DWG" – AutoCAD Drawing format, "DX90" – Data Exchange '90", "DXF" – AutoCAD Drawing Exchange Format, "ERDAS" – ERDAS image files (ERDAS Corporation), "GRASS" – Geographic Resources Analysis Support System, "HDF" – Hierarchical Data Format, "IGDS" – Interactive Graphic Design System format (Intergraph Corporation), "IGES" – Initial Graphics Exchange Standard, "MOSS" – Multiple Overlay Statistical System export file, "netCDF" – network Common Data Format, "NITF" – National Imagery Transfer Format, "RPF" – Raster Product Format (Defense Mapping Agency), "RVC" – Raster Vector Converted format (MicroImages), "RVF" – Raster Vector Format (MicroImages), "SDTS" – Spatial Data Transfer Standard (Federal Information Processing Standard 173), "SIF" – Standard Interchange Format (DOD Project 2851), "SLF" – Standard Linear Format (Defense Mapping Agency), "TIFF" – Tagged Image File Format, "TGRLN" – Topologically Integrated Geographic Encoding and Referencing (TIGER) – Line format (Bureau of the Census), "VPF" – Vector Product Format (Defense Mapping Agency)

Примечание 3632: "CD-ROM", "3-1/2 inch floppy disk", "5-1/4 inch floppy disk", "9-track tape", "4 mm cartridge tape", "8 mm cartridge tape", "1/4-inch cartridge tape", free text

Примечание 3636: "cpio", "tar", "High Sierra", "ISO 9660", "ISO 9660 with Rock Ridge extensions", "ISO 9660 with Apple HFS extensions", free text

Примечание 3805: The domain is from pp. 88-91 in Anglo-American Committee on Cataloguing of Cartographic Materials, 1982, Cartographic materials: A manual of interpretation for AACR2: Chicago, American Library Association: "atlas", "diagram", "globe", "map", "model", "profile", "remote-sensing image", "section", "view".

Приложение В

Наборы меток (tagSet)

В.1 Metadata Tag Set (tagSet-M)

№	Название тэга	Тип данных
1	schemaIdentifier	OBJECT IDENTIFIER
2	elementsOrdered	BOOLEAN
3	elementOrdering	INTEGER
4	defaultTagType	INTEGER
5	defaultVariantSetId	OBJECT IDENTIFIER
6	defaultVariantSpec	VariantSpec
7	processingInstructions	InternationalString
8	recordUsage	INTEGER
9	restriction	InternationalString
10	rank	INTEGER
11	userMessage	InternationalString
12	url	InternationalString
13	record	structured
14	localControlNumber	same as tagSet-G element dateTime
15	creationDate	same as tagSet-G element dateTime
16	dateOfLastModification	same as tagSet-G element dateTime
17	dateOfLastReview	same as tagSet-G element dateTime
18	score	INTEGER
19	wellKnown	Defined by schema; default InternationalString
20	recordWrapper	structured
21	defaultTagSetId	OBJECT IDENTIFIER
22	languageOfRecord	Same as tagSet-G element 'language'
23	type	INTEGER or InternationalString
24	Scheme	INTEGER or InternationalString
25	costInfo	InternationalString
26	costFlag	BOOLEAN ('true' means there is a cost)

27	RecordCreatedBy	InternationalString
28	RecordModifiedBy	InternationalString

B.2 Generic Tag Set (tagSet-G)

№	Название тэга	Тип данных
1	title	InternationalString or structured into following sub-elements: 'wellKnown', tagSet-M element 19, dataType InternationalString, 'type' tagSet-M element 23 'scheme' tagSet-M element 24
2	author	Same datatype definition as title
3	publicationPlace	InternationalString
4	publicationDate	InternationalString
5	documentId	InternationalString
6	abstract	internationalString
7	name	Same datatype definition as title
8	dateTime	EXTERNAL (Z3950DateTime) or GeneralizedTime, or InternationalString, or structured into following sub-elements: 'wellKnown', tagSet-M element 19, dataType 1, 2, or 3 above. 'type' tagSet-M element 23 'scheme' tagSet-M element 24
9	displayObject	OCTET STRING
10	organization	InternationalString
11	postalAddress	InternationalString
12	networkAddress	InternationalString
13	eMailAddress	InternationalString
14	phoneNumber	InternationalString
15	faxNumber	InternationalString
16	country	InternationalString, or structured into following sub-elements: 'wellKnown', tagSet-M element 19, dataType InternationalString 'scheme' tagSet-M element 24
17	description	InternationalString, or structured into following sub-elements: 'wellKnown', tagSet-M element 19, dataType InternationalString 'type' tagSet-M element 23
18	time	
19	DocumentContent	OCTET STRING
20	language	InternationalString, or structured into following sub-elements: 'wellKnown', tagSet-M element 19, dataType InternationalString 'scheme' tagSet-M element 24
21	subject	Same dataType definition as title

22	resourceType	InternationalString, or structured into following sub-elements: 'wellKnown', tagSet-M element 19, dataType InternationalString 'scheme' tagSet-M element 24
23	city	InternationalString
24	stateOrProvince	InternationalString
25	zipOrPostalCode	InternationalString
26	cost	InternationalString, or IntUnit, or structured
27	format	InternationalString or structured
28	identifier	Same dataType definition as title
29	rights	Same dataType definition as title
30	relation	Same dataType definition as title
31	publisher	Same dataType definition as title
32	contributor	Same dataType definition as title
33	source	Same dataType definition as title
34	coverage	Same dataType definition as title
35	private	dataType definition defined by schema
36	databaseName	InternationalString
37	recordId	InternationalString

B.3 GILS Tag Set

№	Название тэга	Тип данных
1	controlIdentifier	InternationalString
2	streetAddress	InternationalString
3	city	InternationalString
4	stateOrProvince	InternationalString
5	zipOrPostalCode	InternationalString
6	hoursOfService	InternationalString
7	resourceDescription	InternationalString
8	technicalPrerequisites	InternationalString
9	westBoundingCoordinate	intUnit
10	eastBoundingCoordinate	intUnit
11	northBoundingCoordinate	intUnit
12	southBoundingCoordinate	intUnit
13	placeKeyword	InternationalString
14	placeKeywordThesaurus	InternationalString
15	beginningDate	GeneralizedTime
16	timePeriodTextual	InternationalString
17	linkage	InternationalString
18	linkageType	InternationalString
19	recordSource	InternationalString
20	controlledTerm	InternationalString
21	subjectThesaurus	InternationalString

22	uncontrolledTerm	InternationalString
23	originalControlIdentifier	InternationalString
24	recordReviewDate	GeneralizedTime
25	generalAccessConstraints	InternationalString
26	originatorDisseminationControl	InternationalString
27	securityClassificationControl	InternationalString
28	orderInformation	InternationalString
29	cost	Boolean
30	costInformation	InternationalString
31	scheduleNumber	InternationalString
32	languageOfResource	InternationalString
33	medium	InternationalString
34	languageOfRecord	InternationalString
35	relationship	InternationalString
36	endingDate	GeneralizedTime
51	purpose	structured
52	originator	structured
53	accessConstraints	structured
54	useConstraints	structured
55	agencyProgram	structured
56	orderProcess	structured
57	sourcesOfData	structured
58	methodology	structured
59	supplementalInformation	structured
70	availability	structured
71	spatialDomain	structured
90	distributor	structured
91	boundingCoordinates	structured
92	place	structured
93	timePeriod	structured
94	pointOfContact	structured
95	controlledSubjectIndex	structured
96	subjectTermsControlled	structured
97	subjectTermsUncontrolled	structured
98	crossReference	structured
99	availableLinkage	structured
100	crossReferenceLinkage	structured
101	timePeriodStructured	structured
102	availableTimeStructured	structured

B.4 Zthes Tag Set

Набор меток для определения схемы данных навигации по тезаурусам на момент написания этих строк не имеет глобального OID. Во всех приложениях используется локально определенный набор (OID:

1.2.840.10003.14.1000.136.1):

№	Название тэга	Тип данных
1	termQualifier	InternationalString
2	termType	InternationalString
3	relationType	InternationalString
4	postings	structured
5	fieldName	InternationalString
6	hitCount	int

B.4.1 Collections Tag Set

№	Название тэга	Тип данных
1	typeOfDescriptiveRecord	INTEGER
2	briefDescription	BriefTextDescription (InternationalString)
3	collectionInfo	CollectionInfo
4	objectInfo	ObjectInfo
5	associatedDescription	AssociatedDescription
6	relatedCollection	relatedCollection
7	collectionName	InternationalString
8	database	ServerAndDb
9	enumeratedMember	EnumeratedMember
10	fullyEnumerated	BOOLEAN
11	childrenKnowThisParent	BOOLEAN
12	typeOfObject	INTEGER
13	categoryOfObject	InternationalString
14	digitalObject	DigitalObject
15	briefDescriptionOfAD	BriefTextDescription (InternationalString)
16	categoryOfAD	InternationalString
17	description	Description
18	collection	CollectionAndDb
19	relationship	INTEGER
20	relativeLevel	INTEGER
21	descriptionOfRelationship	InternationalString
22	collectionName	InternationalString
23	database	ServerAndDb
24	server	InternationalString
25	db	InternationalString
26	briefDescriptionOfMember	BriefTextDescription (InternationalString)
27	pointer	RecordPointer
28	whatPointerPointsTo	INTEGER

29	actualDO	variant dependent; default is OCTET STRING
30	briefDescriptionOfDOVariant	BriefTextDescription (InternationalString)
31	pointerToDo	RecordPointer
32	authoritative	BOOLEAN
33	briefDescriptionOfADVariant	BriefTextDescription (InternationalString)
34	actualAD	variant dependent; default is InternationalString
35	pointerToAD	RecordPointer
36	database	ServerAndDb
37	recordId	OCTET STRING
38	alternativeIdentifier	OCTET STRING
39	typeOfIdentifier	OBJECT IDENTIFIER or InternationalString

B.5 CIMI Tag Set

№	Название тэга	Тип данных
1	repositoryName	InternationalString
2	subject	InternationalString
3	objectID	InternationalString
4	nationalityCultureRace	InternationalString
5	materialMedium	InternationalString
7	creditLine	InternationalString
8	dateOfBirth	InternationalString
9	dateOfDeath	InternationalString
10	role	InternationalString
11	placeOfOrigin	InternationalString
12	processTechnique	InternationalString
13	dimensions	InternationalString
14	stylePeriod	InternationalString
15	provenance	InternationalString
16	relatedObjects	InternationalString
17	quantity	InternationalString
18	award	InternationalString
20	collection	InternationalString
22	inscriptionMark	InternationalString
24	association	structured
25	content	structured
26	repositoryPlace	InternationalString
28	mrObject	structured
29	rendition	structured

30	resource	OCTET String, External, or InternationalString
31	objectName	InternationalString
32	objectTitle	InternationalString
33	bibliographicTitle	InternationalString
35	relatedTextualReferences	InternationalString
36	creatorInfo	structured
38	owner	InternationalString
39	contentGeneral	InternationalString
41	place	InternationalString
42	event	InternationalString
43	activity	InternationalString
45	dateOfOrigin	InternationalString
46	contextHistorical	InternationalString
47	contextArchaeological	InternationalString
48	copyrightRestriction	InternationalString
49	creatorGeneral	InternationalString
50	associationGeneral	InternationalString
51	objectLanguage	InternationalString
52	condition	InternationalString
53	physicalDescription	InternationalString
54	wallTextLabel	InternationalString
55	protectionStatus	InternationalString
56	protectionDate	InternationalString
57	spatialReferencingSystem	InternationalString
58	xCoordinateInSpatialReferencingSystem	InternationalString
59	yCoordinateInSpatialReferencingSystem	InternationalString
60	fieldCollector	InternationalString
61	dateCollected	InternationalString
62	agePeriod	InternationalString
63	typeSpecimen	InternationalString
64	address	InternationalString
65	periodName	InternationalString
66*	administrativeEvent	structured
67*	administrativeEventType	InternationalString
68	administrativeEventGeneral	InternationalString
69	administrator	InternationalString

B.6 UIGGM Tag Set

№	Название тэга	Тип данных
0	Индекс	InternationalString
1	Автор	InternationalString
2	Перевод_заглавия	InternationalString

3	Сокр_заглавие_издания	InternationalString
4	Язык_осн_текста	InternationalString
6	Ключ	InternationalString
7	Дата_издания	InternationalString
11	Составитель	InternationalString
13	Редактор	InternationalString
14	Заявитель_изобретения	InternationalString
15	Библ_ссылки	InternationalString
16	Параллельное_заглавие	InternationalString
17	Номер_депонирования	InternationalString
21	Основное_заглавие	InternationalString
22	Коллект_автор_сокр	InternationalString
24	Обозначение_части_и_секции	InternationalString
25	Заглавие_части_статьи_и_секции	InternationalString
33	Издательство	InternationalString
36	Код_рубрики_ВИНИТИ	InternationalString
37	Индекс_УДК	InternationalString
38	Заглавие_несамос_приложения	InternationalString
40	Страна_подачи_заявки_изобретения	InternationalString
41	Язык_вспом_текста	InternationalString
42	Место_издания_страна	InternationalString
43	Объем_пагинация	InternationalString
44	Иллюстр_материал	InternationalString
47	Сведения_отн_к_заглавию	InternationalString
48	Место_издания_город	InternationalString
49	Вспомогательный_текст	InternationalString
52	ISSN	InternationalString
53	ISBN	InternationalString
56	Дата_утверждения_НТД	InternationalString
57	Дата_введения_НТД_в_действие	InternationalString
58	Дата_окончания_действия_НТД	InternationalString
61	Дата_депонирования_работы	InternationalString
72	Обозначение_номер_подсерии	InternationalString
74	Том	InternationalString
76	Номер_выпуск	InternationalString
78	Дата_защиты_диссертации_год	InternationalString
79	Обозначение_НТД	InternationalString
80	Обозначение_класс_стандарта	InternationalString
81	Периодичность	InternationalString
82	Цена	InternationalString
83	Дата_подачи_заявки_изобретения	InternationalString
84	Организация_хранитель_документа	InternationalString
86	Дата_приоритета	InternationalString
89	Индекс_нац_класс_стандартов	InternationalString
90	Индекс_МКИ	InternationalString

91	Ученая_степень	InternationalString
92	Номер_патентного_документа	InternationalString
93	Сведения_о_переиздании	InternationalString
94	Регистрационный_номер_заявки	InternationalString
95	Дата_публикации_заявки_изобретения	InternationalString
97	Страна_приоритета	InternationalString
98	Номер_приоритетной_заявки	InternationalString
99	Индекс_НКИ	InternationalString
100	Реферат	InternationalString
201	Заглавие_серии	InternationalString
202	Заглавие_подсерии	InternationalString
214	Автор_ответа	InternationalString
501	Имя_индекс_БД	InternationalString
502	Имя_индекс_РЖ	InternationalString
504	Коды_рубрики_ГРНТИ	InternationalString
507	Номер_выпуска_РЖ_БД	InternationalString
580	Номер_записи	InternationalString
581	Организация_создатель_записи	InternationalString
582	Вид_документа	InternationalString
583	Дата_записи	InternationalString
584	Отр_реферат_службами	InternationalString
981	Авторы	structured
982	Ключевые_слова	structured
983	Оснащение	structured
991	Заглавие	structured
992	Список	structured
993	Издание	structured

Приложение С

Схемы данных

С.1 Схема GILS

В таблице представлена схема *GILS* [47], которая определяет иерархическую абстрактную структуру записи в терминах *tagsetM* (1), *tagsetG* (2) и *GILS-tagset* (4) (см. Приложение В).

Tag Path	Element Name	Occur.
(1,1)	schemaIdentifier	
(1,10)	rank	
(1,12)	url	
(1,14)	localControlNumber	
(2,1)	title	
(4,52)	originator	R
(4,52)/(2,7)	originatorName	
(4,52)/(2,10)	originatorOrganization	
(4,52)/(4,2)	originatorStreetAddress	
(4,52)/(4,3)	originatorCity	
(4,52)/(4,4)	originatorStateOrProvince	
(4,52)/(4,5)	originatorZipOrPostalCode	
(4,52)/(2,16)	originatorCountry	
(4,52)/(2,12)	originatorNetworkAddress	R
(4,52)/(4,6)	originatorHoursofService	R
(4,52)/(2,14)	originatorTelephone	R
(4,52)/(2,15)	originatorFax	R
(2,2)	author	R
(2,2)/(2,7)	authorName	
(2,2)/(2,10)	authorOrganization	
(2,2)/(4,2)	authorStreetAddress	
(2,2)/(4,3)	authorCity	
(2,2)/(4,4)	authorStateOrProvince	
(2,2)/(4,5)	authorZipOrPostalCode	

(2,2)/(2,16)	authorCountry	
(2,2)/(2,12)	authorNetworkAddress	R
(2,2)/(4,6)	authorHoursofService	R
(2,2)/(2,14)	authorTelephone	R
(2,2)/(2,15)	authorFax	R
(2,4)	dateOfPublication	
(2,3)	placeOfPublication	
(4,32)	languageOfResource	R
(2,6)	abstract	
(4,95)	controlledSubjectIndex	R
(4,95)/(4,21)	subjectThesaurus	
(4,95)/(4,96)	subjectTermsControlled	
(4,95)/(4,96)/(4,20)	controlledTerm	R
(4,97)	subjectTermsUncontrolled	
(4,97)/(4,22)	uncontrolledTerm	R
(4,71)	spatialDomain	
(4,71)/(4,91)	boundingCoordinates	
(4,71)/(4,91)/(4,9)	westBoundingCoordinate	
(4,71)/(4,91)/(4,10)	eastBoundingCoordinate	
(4,71)/(4,91)/(4,11)	northBoundingCoordinate	
(4,71)/(4,91)/(4,12)	southBoundingCoordinate	
(4,71)/(4,92)	place	R
(4,71)/(4,92)/(4,14)	placeKeywordThesaurus	
(4,71)/(4,92)/(4,13)	placeKeyword	R
(4,93)	timePeriod	R
(4,93)/(4,16)	timePeriodTextual	
(4,93)/(4,101)	timePeriodStructured	R
(4,93)/(4,101)/(4,15)	beginningDate	
(4,93)/(4,101)/(4,36)	endingDate	
(4,70)	availability	R
(4,70)/(4,33)	medium	
(4,70)/(4,90)	distributor	
(4,70)/(4,90)/(2,7)	distributorName	
(4,70)/(4,90)/(2,10)	distributorOrganization	
(4,70)/(4,90)/(4,2)	distributorStreetAddress	
(4,70)/(4,90)/(4,3)	distributorCity	
(4,70)/(4,90)/(4,4)	distributorStateOrProvince	
(4,70)/(4,90)/(4,5)	distributorZipOrPostalCode	
(4,70)/(4,90)/(2,16)	distributorCountry	
(4,70)/(4,90)/(2,12)	distributorNetworkAddress	R
(4,70)/(4,90)/(4,6)	distributorHoursofService	R
(4,70)/(4,90)/(2,14)	distributorTelephone	R
(4,70)/(4,90)/(2,15)	distributorFax	R
(4,70)/(4,7)	resourceDescription	R
(4,70)/(4,55)	orderProcess	

(4,70)/(4,55)/(4,28)	orderInformation	
(4,70)/(4,55)/(4,29)	cost	
(4,70)/(4,55)/(4,30)	costInformation	
(4,70)/(4,8)	technicalPrerequisites	
(4,70)/(4,93)	availableTimePeriod	R
(4,70)/(4,93)/(4,16)	availableTimeTextual	
(4,70)/(4,93)/(4,102)	availableTimeStructured	R
(4,70)/(4,93)/(4,102)/(4,15)	beginningDate	
(4,70)/(4,93)/(4,102)/(4,36)	endingDate	
(4,70)/(4,99)	availableLinkage	R
(4,70)/(4,99)/(4,18)	linkageType	
(4,70)/(4,99)/(4,17)	linkage	R
(4,57)	sourcesOfData	
(4,58)	methodology	
(4,53)	accessConstraints	
(4,53)/(4,25)	generalAccessConstraints	
(4,53)/(4,26)	originatorDisseminationControl	
(4,53)/(4,27)	securityClassificationControl	
(4,54)	useConstraints	
(4,94)	pointOfContact	
(4,94)/(2,7)	contactName	
(4,94)/(2,10)	contactOrganization	
(4,94)/(4,2)	contactStreetAddress	
(4,94)/(4,3)	contactCity	
(4,94)/(4,4)	contactStateOrProvince	
(4,94)/(4,5)	contactZipOrPostalCode	
(4,94)/(2,16)	contactCountry	
(4,94)/(2,12)	contactNetworkAddress	R
(4,94)/(4,6)	contactHoursOfService	R
(4,94)/(2,14)	contactTelephone	R
(4,94)/(2,15)	contactFax	R
(4,59)	supplementalInformation	
(4,51)	purpose	
(4,56)	agencyProgram	
(4,98)	crossReference	R
(4,98)/(2,1)	crossReferenceTitle	
(4,98)/(4,35)	crossReferenceRelationship	R
(4,98)/(4,100)	crossReferenceLinkage	R
(4,98)/(4,100)/(4,18)	linkageType	
(4,98)/(4,100)/(4,17)	linkage	R
(4,31)	scheduleNumber	
(4,1)	controlIdentifier	
(4,23)	originalControlIdentifier	
(4,19)	recordSource	
(4,34)	languageOfRecord	

(1,16)	dateOfLastModification	
(4,24)	recordReviewDate	

C.2 Схема Zthes

В таблице представлена схема *Zthes* [74], которая определяет иерархическую абстрактную структуру записи в терминах *tagsetM* (1), *tagsetG* (2) и *Zthes-tagset* (4) (см. Приложение В)

Tag Path	Element	Occur.
(1,14)	termId	M
(2,1)	termName	M
(4,1)	termQualifier	O
(4,2)	termType	O
(2,17)	termNote	O
(2,20)	termLanguage	O
(1,15)	termCreatedDate	O
(1,27)	termCreatedBy	O
(1,16)	termModifiedDat	O
(1,28)	termModifiedBy	O
(4,4)	postings	O, R
(4,4)/(2,36)	sourceDb	M
(4,4)/(4,5)	fieldName	O
(4,4)/(4,6)	hitCount	M
(2,30)	relation	O, R
(2,30)/(4,3)	relationType	M
(2,30)/(2,36)	sourceDb	O
(2,30)/(1,14)	termId	M
(2,30)/(2,1)	termName	M
(2,30)/(4,1)	termQualifier	O
(2,30)/(4,2)	termType	O
(2,30)/(2,20)	termLanguage	O

C.3 Схема Collections

В таблице представлена схема *Collections* [63], которая определяет иерархическую абстрактную структуру записи в терминах *tagsetM* (1), *tagsetG* (2) и *Collections tagset* (4) (см. Приложение В).

Path	Element	Occur.	Description
(4,1)	typeOfDescriptiveRecord	M	INTEGER 1 = Describes a collection 2 = Describes an object 3 = Unspecified

(4,2)	briefDescription	O	InternationalString
(4,3)	collectionInfo	O	CollectionInfo occurs only if typeOfDescriptiveRecord is 1
(4,4)	objectInfo	O	ObjectInfo occurs only if typeOfDescriptiveRecord is 2
(4,5)	associatedDescription	O, R	AssociatedDescription
(4,6)	relatedCollection	O, R	RelatedCollection

CollectionInfo:

Path	Element	Occur.	Description
(4,7)	collectionName	M	InternationalString
(4,8)	database	O, R	ServerAndDb
(4,9)	enumeratedMember	O, R	EnumeratedMember
(4,10)	fullyEnumerated	O	BOOLEAN
(4,11)	childrenKnowThisParent	O	BOOLEAN

ObjectInfo:

Path	Element	Occur.	Description
(4,12)	typeOfObject	M	INTEGER 0 = Unspecified. 1 = Object is a Digital Object. 2 = Object is a Digital Object for which there is a Physical Object. 3 = Object is a Digital Object for which there is no Physical Object (for example when an object is originally created electronically). 4 = Object is a Physical Object. 5 = There is no (separate) object.
(4,13)	categoryOfObject	O	InternationalString
(4,14)	digitalObject	O, R	DigitalObject may occur only if typeOfObject less than 4

AssociatedDescription:

Path	Element	Occur.	Description
(4,15)	briefDescriptionOfAD	O	InternationalString
(4,16)	categoryOfAD	O	InternationalString
(4,17)	description	M, R	Description

RelatedCollection:

Path	Element	Occur.	Description
(4,18)	collection	M	CollectionAndDb
(4,19)	relationship	O, R	INTEGER 1 = superior collection 2 = context collection 3 = collection may be of interest (no familial relationship implied)
(4,20)	relativeLevel	O	INTEGER
(4,21)	descriptionOfRelationship	O	InternationalString

CollectionAndDb:

Path	Element	Occur.	Description
(4,22)	collectionName	M	InternationalString
(4,23)	database	M	ServerAndDb

ServerAndDb:

Path	Element	Occur.	Description
(4,24)	server	O	InternationalString if omitted, the indicated database is assumed to be on this server
(4,25)	db	M	InternationalString

EnumeratedMember:

Path	Element	Occ.	Description
(4,26)	briefDescriptionOfMember	O	InternationalString
(4,27)	pointer	O	RecordPointer
(4,28)	whatPointerPointsTo	occurs if and only if pointer occurs	INTEGER 1 = element pointer points to an Object Descriptive Record 2 = element pointer points to a Digital Object 3 = element pointer points to a Collection Descriptive Record 4 = type of pointer is Unspecified

DigitalObject:

Path	Element	Occ.	Description
(4,30)	briefDescriptionOf-DOVariant	O	InternationalString

(4,29)	actualDO	occurs if and only if pointerToDo does not occur	variant or category dependent; default is OCTET STRING
(4,31)	pointerToDo	occurs if and only if actualDO does not occur	RecordPointer
(4,32)	authoritative	O	BOOLEAN

Description:

Path	Element	Occ.	Description
(4,33)	briefDescriptionOf-ADVariant	O	InternationalString
(4,34)	actualAD	occurs if and only if pointerToAD does not occur	variant or category dependent; default is InternationalString
(4,32)	pointerToAD	occurs if and only if actualAD does not occur	RecordPointer
(4,32)	authoritative	O	BOOLEAN

RecordPointer:

Path	Element	Occ.	Description
(4,36)	database	occurs (in conjunction with recordId) if and only if alternative Identifier does not occur	ServerAndDb
(4,37)	recordId	occurs (in conjunction with database) if and only if alternative Identifier does not occur	OCTET STRING
(4,38)	alternative-Identifier	occurs if and only if Database does not occur	OCTET STRING
(4,39)	typeOfIdentifier	O occurs only if alternativeIdentifier occurs	OBJECT IDENTIFIER or InternationalString

С.4 Схема *CIMI*

В таблице представлена схема *CIMI* [66], которая определяет иерархическую абстрактную структуру записи в терминах *tagsetM* (1), *tagsetG* (2), *Collection-tagset* (4) и *CIMI-tagset* (5) (см. Приложение В).

Path	Element	Occ.	Description
(1,14)	localControlNumber	1	
(2,1)	title	0+	
(2,2)	creator	0+	
(2,32)	contributor	0+	
(2,8)	date	0+	
(2,17)	description	0+	
(2,28)	identifier	0+	the server can choose the data to return in this element. For example, it may return a URL for a rendition of an image or a pointer to other pages that are appropriate from the perspective of the server
(2,22)	type	0+	
(2,20)	language	0+	
(2,21)	subject	0+	
(2,31)	publisher	0+	
(2,27)	format	0+	
(2,33)	source	0+	
(2,30)	relation	0+	
(2,34)	coverage	0+	
(2,29)	rights	0+	
(1,1)	schemaIdentifier	1	First occurrence of this element in a retrieval record has Required Value = OID of Digital Collections Schema (1.2.840.10003.13.3).
(4,1)	typeOfDescriptive-Record	1	Required Value = 2 (object descriptive record)
(4,4)	objectInfo	1	
(4,4)(4,12)	typeOfObject	1	Required Value = 1 (object is a digital object)

(4,4)(4,13)	categoryOfObject	[0,1]	Permissible Values cimi:unspecified cimi:cataloging record cimi:image record cimi:object record
(4,4)(4,14)	digitalObject	1	
(4,4)(4,14)(4,29)	actualDO	1	ActualDO

ActualDO:

Path	Element	Occ.	Description
(1,1)	schemaIdentifier	1	Second occurrence of this element in a retrieval record has Required Value = OID of CIMI Schema (1.2.840.10003.13.5).
(5,31)	objectName	0+	
(5,32)	objectTitle	1+	In each retrieval record, the element objectTitle must occur if the record is an object record or an image with text record OR if there is no occurrence of the element bibliographicTitle
(5,33)	bibliographicTitle	1+	In each retrieval record, the element bibliographicTitle must occur if the record is a bibliographic record OR if there is no occurrence of the element objectTitle
(5,49)	creatorGeneral	0+	
(5,36)	creatorInfo	0+	CreatorInfo
(5,60)	fieldCollector	0+	
(5,1)	repositoryName	0+	
(5,26)	repositoryPlace	0+	
(5,38)	owner	0+	
(5,7)	creditLine	0+	
(5,2)	subject	0+	
(5,3)	objectID	0+	
(5,5)	materialMedium	0+	
(5,12)	processTechnique	0+	
(5,13)	dimensions	0+	
(5,11)	placeOfOrigin	0+	
(5,45)	dateOfOrigin	0+	
(5,61)	dateCollected	0+	
(5,62)	agePeriod	0+	
(5,63)	typeSpecimen	0+	

(5,14)	stylePeriod	0+	
(5,65)	periodName	0+	
(5,15)	provenance	0+	
(5,17)	quantity	0+	
(5,18)	award	0+	
(5,20)	collection	0+	
(5,22)	inscriptionMark	0+	
(5,51)	objectLanguage	0+	
(5,52)	condition	0+	
(5,53)	physicalDescription	0+	
(5,55)	protectionStatus	0+	
(5,56)	protectionDate	0+	
(5,57)	spatialReferencing System	0+	
(5,58)	xCoordinateInSpatial ReferencingSystem	0+	
(5,59)	yCoordinateInSpatial ReferencingSystem	0+	
(5,64)	address	0+	
(5,16)	relatedObjects	0+	
(5,35)	relatedTextual References	0+	
(5,50)	associationGeneral	0+	
(5,24)	association	0+	MoreInfo
(5,39)	contentGeneral	0+	
(5,25)	content	0+	MoreInfo
(5,46)	contextHistorical	0+	
(5,47)	contextArchaeological	0+	
(5,48)	copyrightRestriction	0+	
(5,54)	wallTextLabel	0+	
(2,9)	displayObject	0+	appliedVariant may be used with element; default is text
(5,68)	administrativeEvent General	0+	
(5,66)	administrativeEvent	0+	AdministrativeEvent
(5,69)	administrator	0+	
(5,28)	mrObject	0+	MrObject. Occurs for each "image" associated with record. If the element mrObject occurs, at least one occurrence of rendition and its subelement resource are mandatory.

CreatorInfo:

Path	Element	Occur.	Description
(2,7)	name	1	

(5,4)	nationalityCultureRace	0+	
(5,8)	dateOfBirth	0+	
(5,9)	dateOfDeath	0+	
(5,10)	role	0+	

MrObject:

Path	Element	Occur.	Description
(2,1)	title	0+	
(2,2)	creator	0+	
(2,32)	contributor	0+	
(2,8)	date	0+	The element date may be returned as either GeneralizedTime or EXTERNAL datatypes.
(2,17)	description	0+	
(2,22)	type	0+	
(2,20)	language	0+	
(2,21)	subject	0+	
(2,31)	publisher	0+	
(2,27)	format	0+	
(2,33)	source	0+	
(2,30)	relation	0+	
(2,34)	coverage	0+	
(2,29)	rights	0+	
(5,29)	rendition	1+	Rendition. If the element mrObject occurs, at least one occurrence of rendition and its subelement resource is mandatory. Renditions are to be ordered in the GRS record with the smallest size/lowest resolution listed in the first occurrence of rendition.

Rendition:

Path	Element	Occ.	Description
(5,30)	resource	1	appliedVariant must be used with element
(2,1)	title	0+	
(2,2)	creator	0+	
(2,32)	contributor	0+	
(2,8)	date	0+	The element date may be returned as either GeneralizedTime or EXTERNAL datatypes
(2,17)	description	0+	
(2,22)	type	0+	
(2,20)	language	0+	
(2,21)	subject	0+	

(2,31)	publisher	0+	
(2,27)	format	0+	
(2,33)	source	0+	
(2,30)	relation	0+	
(2,34)	coverage	0+	
(2,29)	rights	0+	

MoreInfo:

Path	Element	Occ.	Description
(2,7)	name	[0,1]	
(5,41)	place	[0,1]	
(5,42)	event	[0,1]	
(5,43)	activity	[0,1]	
(2,17)	description	[0,1]	

AdministrativeEvent:

Path	Element	Occ.	Description
(5,67)	administrativeEventType	1	Takes on values such as "Creation" and "Fst-publ"
(2,8)	date	1	Date of event
(2,2)	creator	1	Person responsible for event
(5,10)	role	1	Takes on values such as "Editor," "Publisher" (sic), and "Reviewer"

C.5 Схема *UIGGM*

В таблице представлена локально определенная схема *UIGGM* в терминах *UIGGMtagset* (89) (см. Приложение В).

Tag Path	Element	Occur.
(89,0)	Индекс	M
(89,580)	Номер_записи	M
(89,981)	Авторы	O
(89,981)/(89,1)	Автор	O
(89,981)/(89,992)	Список	O
(89,022)	Коллект_автор_сокр	O
(89,991)	Заглавие	O
(89,991)/(89,21)	Основное_заглавие	O
(89,991)/(89,16)	Параллельное_заглавие	O
(89,991)/(89,2)	Перевод_заглавия	O
(89,991)/(89,047)	Сведения_отн_к_заглавию	O
(89,991)/(89,024)	Обозначение_части_и_секции	O

(89,991)/(89,025)	Заглавие_части_статьи_и_секции	О
(89,993)	Издание	О
(89,993)/(89,201)	Заглавие_серии	О
(89,993)/(89,202)	Заглавие_подсерии	О
(89,993)/(89,072)	Обозначение_номер_подсерии	О
(89,993)/(89,003)	Сокр_заглавие_издания	О
(89,993)/(89,007)	Дата_издания	О
(89,993)/(89,074)	Том	О
(89,993)/(89,076)	Номер_выпуск	О
(89,993)/(89,43)	Объем_пагинация	О
(89,993)/(89,053)	ISBN	О
(89,993)/(89,052)	ISSN	О
(89,993)/(89,048)	Место_издания_город	О
(89,993)/(89,033)	Издательство	О
(89,993)/(89,042)	Место_издания_страна	О
(89,993)/(89,081)	Периодичность	О
(89,993)/(89,093)	Сведения_о_переиздании	О
(89,993)/(89,098)	Номер_приоритетной_заявки	О
(89,993)/(89,097)	Страна_приоритета	О
(89,993)/(89,086)	Дата_приоритета	О
(89,993)/(89,094)	Регистрационный_номер_заявки	О
(89,993)/(89,040)	Страна_подачи_заявки_изобретения	О
(89,993)/(89,083)	Дата_подачи_заявки_изобретения	О
(89,993)/(89,095)	Дата_публикации_заявки_изобретения	О
(89,993)/(89,092)	Номер_патентного_документа	О
(89,993)/(89,061)	Дата_депонирования_работы	О
(89,993)/(89,017)	Номер_депонирования	О
(89,013)	Редактор	О
(89,014)	Заявитель_изобретения	О
(89,100)	Реферат	О
(89,507)	Номер_выпуска_РЖ_БД	О
(89,982)	Ключевые_слова	О
(89,982)/(89,6)	Ключ	О
(89,982)/(89,992)	Список	О
(89,079)	Обозначение_НТД	О
(89,581)	Организация_создатель_записи	О
(89,582)	Вид_документа	О
(89,583)	Дата_записи	О
(89,004)	Язык_осн_текста	О
(89,041)	Язык_вспом_текста	О
(89,507)	Номер_выпуска_РЖ_БД	О
(89,502)	Имя_индекс_РЖ	О
(89,501)	Имя_индекс_БД	О
(89,037)	Индекс_УДК	О
(89,090)	Индекс_МКИ	О

(89,099)	Индекс_НКИ	О
(89,504)	Коды_рубрики_ГРНТИ	О
(89,036)	Код_рубрики_ВИНИТИ	О
(89,089)	Индекс_нац_класс_стандартов	О
(89,080)	Обозначение_класс_стандарта	О
(89,214)	Автор_ответа	О
(89,056)	Дата_утверждения_НТД	О
(89,057)	Дата_введения_НТД_в_действие	О
(89,058)	Дата_окончания_действия_НТД	О
(89,091)	Ученая_степень	О
(89,078)	Дата_защиты_диссертации_год	О
(89,584)	Отр_реферат_службами	О
(89,084)	Организация_хранитель_документа	О
(89,983)	Оснащение	О
(89,983)/(89,044)	Иллюстр_материал	О
(89,983)/(89,049)	Вспомогательный_текст	О
(89,983)/(89,015)	Библ_ссылки	О
(89,992)	Список	О

С.6 Схема GEO

Схема GEO приведена в виде *DTD (Document Type Declaration)*. Названия элементов совпадают с названиями поисковых атрибутов *GEO-1* (см. Приложение А.10).

```

<!-- FGDC Metadata DTD 1.0 19960501 -->

<!-- This is the Document Type Declaration for formal metadata, -->
<!-- metadata conforming to the Content Standards for Digital -->
<!-- Geospatial Metadata of the Federal Geographic Data -->
<!-- Committee. This DTD corresponds to the 19940608 (June 8, -->
<!-- 1994) version of the standard. -->

<!-- Tags: -->
<!-- Tags are a maximum of 8-characters long, to coincide -->
<!-- with the Reference Concrete Syntax. -->

<!-- Entity sets: -->
<!-- Scalar values (meaning the values of elements that are -->
<!-- not compound) are here declared #PCDATA to allow XML -->
<!-- parsers to recognize and support entities representing -->
<!-- special characters such as the degree symbol, less, -->
<!-- and greater. -->

<!-- Element ordering: -->
<!-- Generally the order of elements is not significant. -->
<!-- However, repeated groups of elements must be in order -->
<!-- (e.g. (A, B)+) and repeated elements must appear -->
<!-- together (e.g. (A, B+, C)) -->

<!-- Authors: -->

```



```

<!-- Peter N. Schweitzer (USGS, Reston, VA 20192) with      -->
<!-- assistance from Doug Nebert (USGS), Eric Miller (OCLC), -->
<!-- Quinn Hart (CERES), Jim Frew (UCSB), and              -->
<!-- Archie Warnock (AWWW).                                -->

<!--=====>

<!ENTITY % ISOnum PUBLIC
  "ISO 8879:1986//ENTITIES Numeric and Special Graphic//EN">
<!ENTITY % ISolat1 PUBLIC
  "ISO 8879:1986//ENTITIES Added Latin 1//EN">

%ISOnum;
%ISolat1;

<!--=====>

<!ELEMENT metadata (idinfo, dataqual?, spdoinfo?, spref?, eainfo?,
  distinfo?, extinfo*, metainfo)>
<!ELEMENT idinfo (citation, descript, timeperd, status, spdom,
  keywords, accconst, useconst, ptcontac?,
  browse*, datacred?, secinfo?, native?, crossref*)>
<!ELEMENT citation (citeinfo)>
<!ELEMENT descript (abstract, purpose, supplinf?)>
<!ELEMENT timeperd (timeinfo, current)>
<!ELEMENT status (progress, update)>
<!ELEMENT spdom (bounding, dsgpoly*, extent)>
<!ELEMENT bounding (westbc, eastbc, northbc, southbc)>
<!ELEMENT dsgpoly (dsgpolyo, dsgpolyx*)>
<!ELEMENT dsgpolyo ((gringlat, gringlon),(gringlat, gringlon),
  (gringlat, gringlon),(gringlat, gringlon)+)>
<!ELEMENT dsgpolyx ((gringlat, gringlon),(gringlat, gringlon),
  (gringlat, gringlon),(gringlat, gringlon)+)>
<!ELEMENT keywords theme, place?, stratum?, temporal?)>
<!ELEMENT theme ((themekt, themekey+)+)>
<!ELEMENT place ((placekt, placekey+)+)>
<!ELEMENT stratum ((stratkt, stratkey+)+)>
<!ELEMENT temporal ((tempkt, tempkey+)+)>
<!ELEMENT ptcontac (cntinfo)>
<!ELEMENT browse (brosen, browsed, browsset)>
<!ELEMENT secinfo (secsys, secclass, sechandl)>
<!ELEMENT crossref (citeinfo)>
<!ELEMENT dataqual (attracc?, logic, complete, posacc?, lineage,
  cloud?)>
<!ELEMENT attracc (attraccr, qattracc*)>
<!ELEMENT qattracc (attraccv, attracce)>
<!ELEMENT posacc (horizpa?, vertacc?)>
<!ELEMENT horizpa (horizpar, qhorizpa*)>
<!ELEMENT qhorizpa (horizpav, horizpae)>
<!ELEMENT vertacc (vertaccr, qvertpa*)>
<!ELEMENT qvertpa (vertaccv, vertacce)>
<!ELEMENT lineage (srcinfo*, procstep+)>
<!ELEMENT srcinfo (srccite, srcscale?, typesrc, srctime,
  srccitea, srccontr)>
<!ELEMENT srccite (citeinfo)>
<!ELEMENT srctime (timeinfo, srccurr)>
<!ELEMENT procstep (procdesc, srcused*, procddate, proctime?,
  srcprod*, procont?)>
<!ELEMENT procont (cntinfo)>

```

```

<!ELEMENT spdoinfo (indspref?, (direct,
    ((ptvctinf | rastinfo))?)?)>
<!ELEMENT ptvctinf ((sdtstern | vpfterm))>
<!ELEMENT sdtstern ((sdtstype, ptvctcnt?)+)>
<!ELEMENT vpfterm (vpflevel, (vpftype, ptvctcnt?)+)>
<!ELEMENT rastinfo (rasttype, (rowcount, colcount, vrtcount?)?)>
<!ELEMENT spref (horizsys?, vertdef?)>
<!ELEMENT horizsys ((geograph | planar+ | local), geodetic?)>
<!ELEMENT geograph (latres, longres, geogunit)>
<!ELEMENT planar ((mapproj | gridsys | localp), planci)>
<!ELEMENT mapproj (mapprojn, (albers | azimequi | equicon |
    equirect | gvnsp | gnomonic | lamberta | lambertc |
    mercator | modsak | miller | obqmerc | orthogr | polarst |
    polycon | robinson | sinusoid | spaceobq | stereo |
    transmer | vdgrin | otherprj))>
<!ELEMENT albers (stdparll+, longcm, latprjo, feast, fnorth)>
<!ELEMENT azimequi (longcm, latprjo, feast, fnorth)>
<!ELEMENT equicon (stdparll+, longcm, latprjo, feast, fnorth)>
<!ELEMENT equirect (stdparll, longcm, feast, fnorth)>
<!ELEMENT gvnsp (heightpt, longpc, latprjc, feast, fnorth)>
<!ELEMENT gnomonic (longpc, latprjc, feast, fnorth)>
<!ELEMENT lamberta (longpc, latprjc, feast, fnorth)>
<!ELEMENT lambertc (stdparll+, longcm, latprjo, feast, fnorth)>
<!ELEMENT mercator ((stdparll | sfequat), longcm, feast, fnorth)>
<!ELEMENT modsak (feast, fnorth)>
<!ELEMENT miller (longcm, feast, fnorth)>
<!ELEMENT obqmerc (sfctrln, (obqlazim | obqlpt), latprjo, feast,
    fnorth)>
<!ELEMENT obqlazim (azimangl, azimptl)>
<!ELEMENT obqlpt ((obqlat, obqllong),(obqlat, obqllong))>
<!ELEMENT orthogr (longpc, latprjc, feast, fnorth)>
<!ELEMENT polarst (svlong, (stdparll | sfprjorg), feast, fnorth)>
<!ELEMENT polycon (longcm, latprjo, feast, fnorth)>
<!ELEMENT robinson (longpc, feast, fnorth)>
<!ELEMENT sinusoid (longcm, feast, fnorth)>
<!ELEMENT spaceobq (landsat, pathnum, feast, fnorth)>
<!ELEMENT stereo (longpc, latprjc, feast, fnorth)>
<!ELEMENT transmer (sfctrmer, longcm, latprjo, feast, fnorth)>
<!ELEMENT vdgrin (longcm, feast, fnorth)>
<!ELEMENT gridsys (gridsysn, (utm | ups | spcs | arcsys |
    othergrd))>
<!ELEMENT utm (utmzone, transmer)>
<!ELEMENT ups (upszone, polarst)>
<!ELEMENT spcs (spcszone, (lambertc | transmer | obqmerc |
    polycon))>
<!ELEMENT arcsys (arczone, (equirect | azimequi))>
<!ELEMENT localp (localpd, localpgi)>
<!ELEMENT planci (plance, (coordrep | distbrep), plandu)>
<!ELEMENT coordrep (absres, ordres)>
<!ELEMENT distbrep (distres, bearres, bearunit, bearrefd, bearrefm)>
<!ELEMENT local (localdes, localgeo)>
<!ELEMENT geodetic (horizdn?, ellips, semiaxis, denflat)>
<!ELEMENT vertdef (altsys?, depthsys?)>
<!ELEMENT altsys (altdatum, altres+, altunits, altenc)>
<!ELEMENT depthsys (depthdn, depthres+, depthdu, depthem)>
<!ELEMENT eainfo (detailed*, overview*)>
<!ELEMENT detailed ((enttyp, attr*)+)>
<!ELEMENT enttyp (enttyppl, enttypd, enttypds)>
<!ELEMENT attr (attrlabl, attrdef, attrdefs, attrdomv+,

```

```

        attrunit?, attrmres?, (begdatea, enddatea?)*, attrvai?,
        attrmfrq?)>
<!ELEMENT attrrdomv ((edom | rdom | codesetd | udom))>
<!ELEMENT edom      ((edomv, edomvd, edomvds, attr*)+)>
<!ELEMENT rdom      (rdomin, rdommmax, attr*)>
<!ELEMENT codesetd  (codesetn, codesets)>
<!ELEMENT attrvai   (attrva, attrvae)>
<!ELEMENT overview  ((eaover, eadetcit+)>
<!ELEMENT distinfo  ((distrib, resdesc?, distliab, stdorder*,
        custom?, techpreq?, availabl?)+)>
<!ELEMENT distrib   (cntinfo)>
<!ELEMENT stdorder  ((nondig | digform+), fees, ordering?,
        turnarnd?)>
<!ELEMENT digform   (digtinfo, digtopt)>
<!ELEMENT digtinfo  (formname, ((formvern | formverd), formspec?)*,
        formcont?, filedec?, transize?)>
<!ELEMENT digtopt   (((onlinopt | offoptn))+)>
<!ELEMENT onlinopt  (computer+, accinstr?, oncomp?)>
<!ELEMENT computer  ((networka | dialinst))>
<!ELEMENT networka (networkr+)>
<!ELEMENT dialinst  (lowbps, highbps?, numdata, numstop, parity,
        compress?, dialtel+, dialfile+)>
<!ELEMENT offoptn   (offmedia, reccap?, recfmt+, compat?)>
<!ELEMENT reccap    (recden+, recdenu)>
<!ELEMENT availabl  (timeinfo)>
<!ELEMENT metainfo  (metd, (metrd, metfrd?)*, metc, metstdn,
        metstdv, mettc?, metac?, metuc?, metsi?)>
<!ELEMENT metc      (cntinfo)>
<!ELEMENT metsi     (metscs, metsc, metshd)>
<!ELEMENT citeinfo  (origin+, pubdate, pubtime?, title, edition?,
        geoform?, serinfo?, pubinfo?, othercit?, onlink*, lworkcit?)>
<!ELEMENT serinfo   (sername, issue)>
<!ELEMENT pubinfo   (pubplace, publish)>
<!ELEMENT lworkcit  (citeinfo)>
<!ELEMENT timeinfo  ((sngdate | mdattim | rngdates))>
<!ELEMENT sngdate   (caldate, time?)>
<!ELEMENT mdattim   ((caldate, time?)+)>
<!ELEMENT rngdates  (begdate, begtime?, enddate, endtime?)>
<!ELEMENT cntinfo   ((cntperp | cntorgp), cntpos?, cntaddr+,
        cntvoice+, cnttdd*, cntfax*, cntemail*, hours?, cntinst?)>
<!ELEMENT cntperp   (cntper, cntorg?)>
<!ELEMENT cntorgp   (cntorg, cntper?)>
<!ELEMENT cntaddr   (addrtype, address*, city, state, postal,
        country?)>

<!--
<!-- Остальные элементы должны быть представлены в виде
<!--
<!--          <!ELEMENT xxxxxxxx (#PCDATA)>
<!--
<!-- где xxxxxxxx - каждый элемент из списка:

arczone, absres, abstract, accconst, accinstr, address, addrtype,
altdatum, altunits, altenc, altres, attracce,attraccr, attraccv,
attrdef, attrdefs, attrlabl, attrmfrq, attrmres, attrunit, attrva,
attrvae, azimptl, azimangl, bearrefd, bearrefm, bearres, bearunit,
begdate, begdatea, begtime, browsed, browsen, browsset, caldate,
city, cloud, codesetn, codesets, colcount, compat, complete,
compress, cntemail, cntfax, cntinst, cntorg, cntper, cntpos,

```

cnttdd, cntvoice, country, current, custom, datacred, denflat,
depthdn, depthdu, depthem, depthres, dialfile, dialtel, direct,
distres, distliab, eastbc, edition, ellips, enddate, enddatea,
endtime, enttypd, enttypds, enttypl, eadetcit, eaover, edomv,
edomvd, edomvds, extent, feast, fnorth, fees, filedec, formcont,
formname, formspec, formverd, formvern, gringlat, gringlon,
geogunit, geiform, gridsysn, heightpt, highbps, horizdn, horizpae,
horizpar, horizpav, hours, indspref, issue, landsat, latres,
latprjc, latprjo, localdes, localgeo, localpd, localpgi, logic,
longres, longcm, longpc, lowbps, update, mapprojn, metac, metd ,
metfrd, metrd, metsc, metscs, metshd, metstdn, metstdv, mettc,
metuc, native, networkr, nondig, northbc, numdata, numstop, obqlat,
obqlong, offmedia, oncomp, onlink, ordering, ordres, origin,
othercit, othergrd, otherprj, parity, pathnum, placekey, placekt,
plance, plandu, ptvctcnt, postal, procdat, procdesc, proctime,
progress, pubdate, pubplace, pubtime, publish, purpose, rdommax,
rdommin, rasttype, recden, recdenu, recfmt, resdesc, rowcount,
sdtstype, spcszone, sfctrln, sfctrmer, sfequat, sfprjorg,
secclass, secsys, sechandl, semiaxis, sername, srccitea, srcccontr,
srccurr, srcprod, srcscale, srcused, southbc, stdparll, state,
svlong, stratkey, stratkt, supplinf, techpreq, tempkey, tempkt,
themekey, themekt, time, title, transize, turnarnd, typesrc,
upszone, utmzone, udom, useconst, vpftype, vpflevel, vrtcount,
vertacce, vertaccr, vertaccv, westbc -->

<!-- End of FGDC Metadata DTD v 1.0 -->

Приложение D

Идентификаторы объектов сервера ZooPARK

В таблице представлены идентификаторы объектов, которые распознает сервер ZooPARK. Последняя колонка содержит текстовое представление объекта, которое используется в конфигурационных файлах.

Класс	OID	Текст
TRANSYN	2.1.1	BER
TRANSYN	1.0.2709.1.1	ISO2709
GENERAL	1.2.10161.2.1	ISOILL-1
ABSYN	Z.2.1	Z-APDU
APPCTX	Z.1.1	Z-BASIC
ATTSET	Z.3.1	Bib-1
ATTSET	Z.3.2	Exp-1
ATTSET	Z.3.3	Ext-1
ATTSET	Z.3.4	CCL-1
ATTSET	Z.3.5	GILS-
ATTSET	Z.3.6	STAS-attset
ATTSET	Z.3.7	Collections-attset
ATTSET	Z.3.8	CIMI-attset
ATTSET	Z.3.9	Geo-attset
ATTSET	Z.3.10	ZBIG-attset
ATTSET	Z.3.11	Util-attset
ATTSET	Z.3.12	XD-1
ATTSET	Z.3.13	ZTHES-attset
ATTSET	Z.3.14	FIN-1
ATTSET	Z.3.15	DAN-1
ATTSET	Z.3.16	Holdings

ATTSET	Z.3.17	MARC
ATTSET	Z.3.18	Bib-2
ATTSET	Z.3.19	ZeeRex
ATTSET	Z.3.1000.81.1	Thesaurus-attset
ATTSET	Z.3.1000.81.2	IDXPATH
DIAGSET	Z.4.1	Bib-1
DIAGSET	Z.4.2	Diag-1
DIAGSET	Z.4.3	Diag-ES
DIAGSET	Z.4.4	Diag-General
RECSYN	Z.5.1	Unimarc
RECSYN	Z.5.2	Intermarc
RECSYN	Z.5.3	CCF
RECSYN	Z.5.10	USmarc
RECSYN	Z.5.11	UKmarc
RECSYN	Z.5.12	Normarc
RECSYN	Z.5.13	Librismarc
RECSYN	Z.5.14	Danmarc
RECSYN	Z.5.15	Finmarc
RECSYN	Z.5.16	MAB
RECSYN	Z.5.17	Canmarc
RECSYN	Z.5.18	SBN
RECSYN	Z.5.19	Picamarc
RECSYN	Z.5.20	Ausmarc
RECSYN	Z.5.21	Ibermarc
RECSYN	Z.5.22	Carmarc
RECSYN	Z.5.23	Malmarc
RECSYN	Z.5.24	JPmarc
RECSYN	Z.5.25	SWEMarc
RECSYN	Z.5.26	SIGLEmarc
RECSYN	Z.5.27	ISSNmarc
RECSYN	Z.5.28	RUSmarc
RECSYN	Z.5.29	Hunmarc
RECSYN	Z.5.100	Explain
RECSYN	Z.5.101	SUTRS
RECSYN	Z.5.102	OPAC
RECSYN	Z.5.103	Summary
RECSYN	Z.5.104	GRS-0
RECSYN	Z.5.105	GRS-1

RECSYN	Z.5.106	Extended
RECSYN	Z.5.107	Fragment
RECSYN	Z.5.109.1	pdf
RECSYN	Z.5.109.2	postscript
RECSYN	Z.5.109.3	html
RECSYN	Z.5.109.4	tiff
RECSYN	Z.5.109.5	gif
RECSYN	Z.5.109.6	jpeg
RECSYN	Z.5.109.7	png
RECSYN	Z.5.109.8	mpeg
RECSYN	Z.5.109.9	sgml
RECSYN	Z.5.110.1	tiff-b
RECSYN	Z.5.110.2	wav
RECSYN	Z.5.111	SQL-RS
RECSYN	Z.5.1000.81.2	SOIF
RECSYN	Z.5.109.10	text-XML
RECSYN	Z.5.109.10	XML
RECSYN	Z.5.109.11	application-XML
RECSYN	Z.5.1000.155.1	rtf
RESFORM	Z.7.1	Resource-1
RESFORM	Z.7.2	Resource-2
RESFORM	Z.1000.81.1	UNInverse-Resource-Report
ACCFORM	Z.8.1	Prompt-1
ACCFORM	Z.8.2	Des-1
ACCFORM	Z.8.3	Krb-1
EXTSERV	Z.9.1	Pers. set
EXTSERV	Z.9.2	Pers. query
EXTSERV	Z.9.3	Per'd query
EXTSERV	Z.9.4	Item order
EXTSERV	Z.9.5	DB. Update (first version
EXTSERV	Z.9.5.1	DB. Update (second version
EXTSERV	Z.9.5.1.1	DB. Update
EXTSERV	Z.9.6	exp. spec
EXTSERV	Z.9.7	exp. inv
EXTSERV	Z.9.1000.81.1	Admin
USERINFO	Z.10.1	searchResult-1
USERINFO	Z.10.2	CharSetandLanguageNegotiation
USERINFO	Z.10.3	UserInfo-1

USERINFO	Z.10.4	MultipleSearchTerms-1
USERINFO	Z.10.5	MultipleSearchTerms-2
USERINFO	Z.10.6	DateTime
USERINFO	Z.10.1000.81.1	Proxy
USERINFO	Z.10.1000.81.2	Cookie
USERINFO	Z.10.1000.81.3	Client-IP
USERINFO	Z.10.1000.155.1	CodePage
USERINFO	Z.10.1000.155.2	ZGateway
USERINFO	Z.10.1000.155.3	SearchResultHits
USERINFO	Z.10.1000.155.4	ModulesInfo
USERINFO	Z.10.1000.155.5	RemoteInfo
ELEMSPEC	Z.11.1	Espec-1
VARSET	Z.12.1	Variant-1
SCHEMA	Z.13.1	WAIS-schema
SCHEMA	Z.13.2	GILS-schema
SCHEMA	Z.13.3	Collections-schema
SCHEMA	Z.13.4	Geo-schema
SCHEMA	Z.13.5	CIMI-schema
SCHEMA	Z.13.6	Update ES
SCHEMA	Z.13.7	Holdings-schema
SCHEMA	Z.13.8	Zthes-schema
SCHEMA	Z.13.1000.81.1	thesaurus-schema
SCHEMA	Z.13.1000.81.2	Explain-schema
SCHEMA	Z.13.1000.155.1	UIGGM-schema
SCHEMA	Z.13.1000.155.2	UIGGMe-schema
SCHEMA	Z.13.1000.155.4	PERSONS -schema
TAGSET	Z.14.1	TagsetM
TAGSET	Z.14.2	TagsetG
TAGSET	Z.14.3	STAS-tagset
TAGSET	Z.14.4	GILS-tagset
TAGSET	Z.14.5	Collections-tagset
TAGSET	Z.14.6	CIMI-tagset
TAGSET	Z.14.1000.81.1	thesaurus-tagset
TAGSET	Z.14.1000.81.2	Explain-tagset
TAGSET	Z.14.1000.155.1	UIGGM-tagset
TAGSET	Z.14.1000.155.2	UIGGMe-tagset
TAGSET	Z.14.1000.155.3	Geo-tagset
TAGSET	Z.14.1000.155.4	PERSONS -tagset

TAGSET	Z.14.1000.136.1	Zthes-tagset
NEGOT	Z.15.1	CharNegotiation-2
NEGOT	Z.15.2	NegotiateESSizes
NEGOT	Z.15.3	CharNegotiation-3
NEGOT	Z.15.1000.155.1	CP-Negotiation
QUERY	Z.16.1	QuerySQL
QUERY	Z.16.2	CQL
GENERAL	1.0.10646.1.0.2	UCS-2
GENERAL	1.0.10646.1.0.4	UCS-4
GENERAL	1.0.10646.1.0.5	UTF-16
GENERAL	1.0.10646.1.0.8	UTF-8