

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

Жижимов О.Л., Мазов Н.А.

# **Применение протокола Z39.50: проекты и решения**

## Краткий обзор проектов по построению информационных систем, в том числе и распределенных, с применением протокола Z39.50:

- **библиотечные проекты**  
*UNlverse, ZLOT, LibWeb, АРБИКОН* и др.;
- **проекты GILS (FirstGOV)**
- **электронные библиотеки**  
*LAURIN, DNER, EULER* и др.;
- **проекты в области культурного наследия**  
*Aquarelle, Zavier, AHDS, ARTISTE* и др.;
- **проекты в области естественных наук**  
*Species Analyst, FishNet, MaNIS, HerpNET* и др.;
- **геоинформационные проекты**  
*FGDC Clearinghouse, SEDAC, IAI-DIS* и др.

# Модели интеграции информационных ресурсов

**Сводный каталог** - представляет собой некую базу данных, в которую включаются записи из разных каталогов. Дублируемые записи при этом убираются, 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. Для обеспечения прозрачного доступа к распределенным каталогам принято использовать протокол Z39.50.

# *UNiverse*

- Проект *UNiverse* финансировался Европейской комиссией в четвертой 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.
- Проект ориентировался на открытые стандарты включая *ANSI Z39.50* для поиска и извлечения информации из баз данных, *ILL* для обслуживания заказов по межбиблиотечному абонементу (МБА) и *GEDI* (*Group on Electronic Document Interchange*) рекомендации для электронной доставки документов.
- Консорциум *UNiverse* 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.

# UNiverse

- Большое количество географически распределенных библиотечных каталогов, которые в рамках проекта **UNiverse** должны вести себя как единый виртуальный каталог, комбинируется с множеством высококачественных библиотечных услуг и конечному пользователю и библиотекарю.



# UNiverse

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 "межобластные" 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 МБА и доставки документов;
- объединение электронной доставки документов на стандартной основе.

# ***UNiverse***

- Проект ***UNiverse*** явился стартовой площадкой для многих более поздних проектов. В частности, можно отметить проект ***UNEX (UNiverse Expansion)***, в котором в качестве партнера участвовала Библиотека Академии Наук России (БАН).
- Проект ***UNEX*** 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 расширения списка партнеров, это был первый международный проект по интеграции ресурсов на основе *Z39.50*, в котором участвовала Россия.

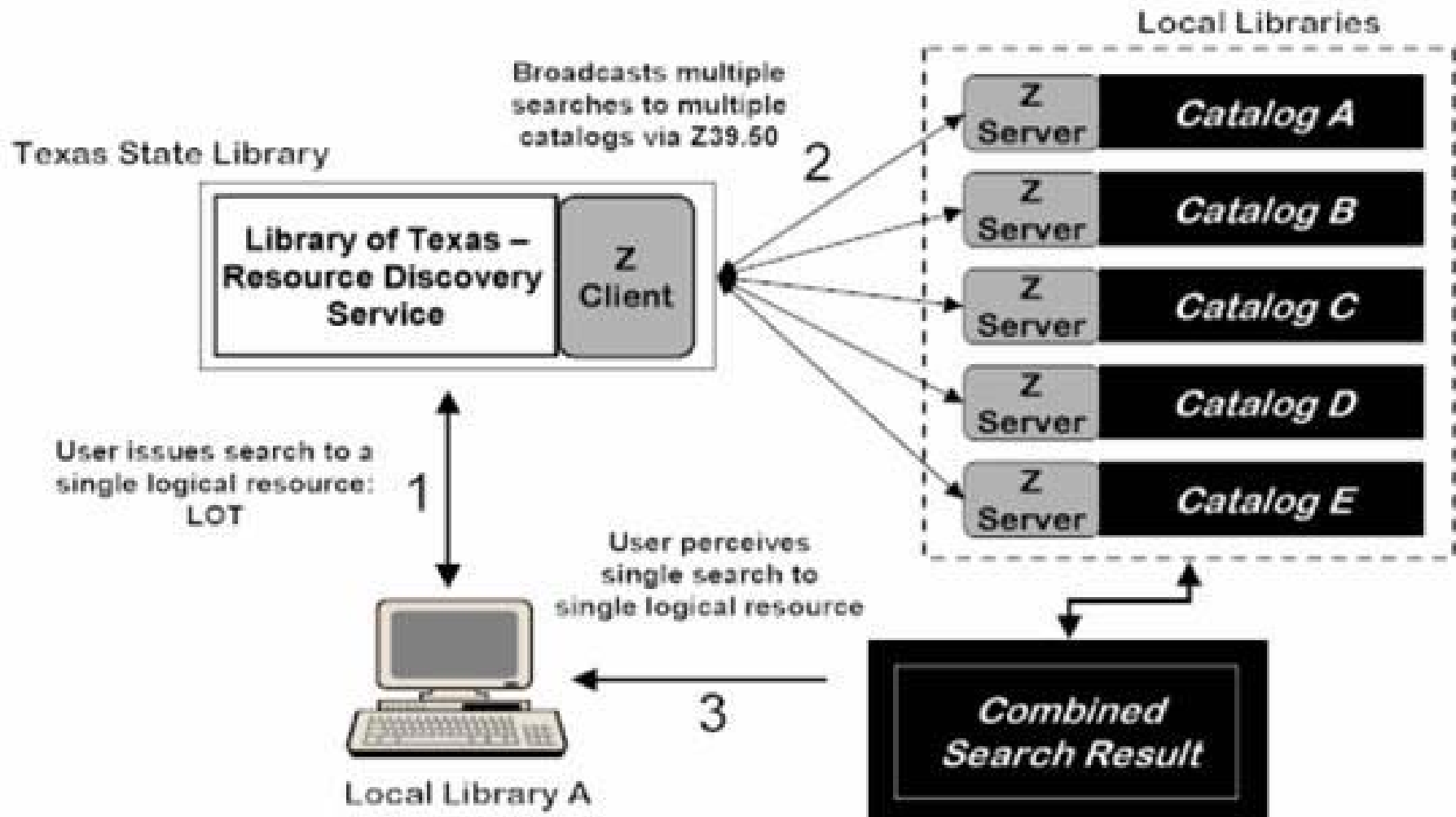
# ZLOT (*LOT - The Library of Texas*)

Проект выполняется с декабря 2001 года с первоначальным бюджетом 500 тыс. USD и призван в конечном счете вывести сервис доступа к информационным ресурсам более 600 публичных и академических библиотек штата на новый уровень. Проект охватывает многие компоненты:

- 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.



# Архитектура ZLOT



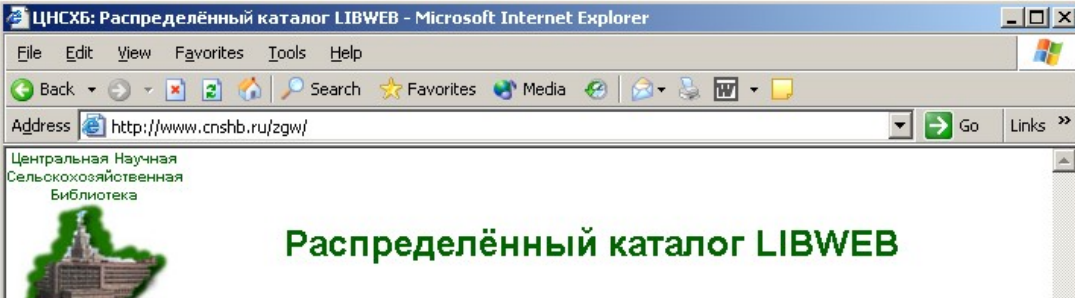
# *LibWeb*

- 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).
- В ходе выполнения проекта была осуществлена пилотная стадия создания распределенной сети ***LibWeb*** в том числе на основе протокола Z39.50.
- 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
- В соответствии с принципами проекта ***LibWeb***, все типовые решения бесплатно передавались всем участникам проекта.

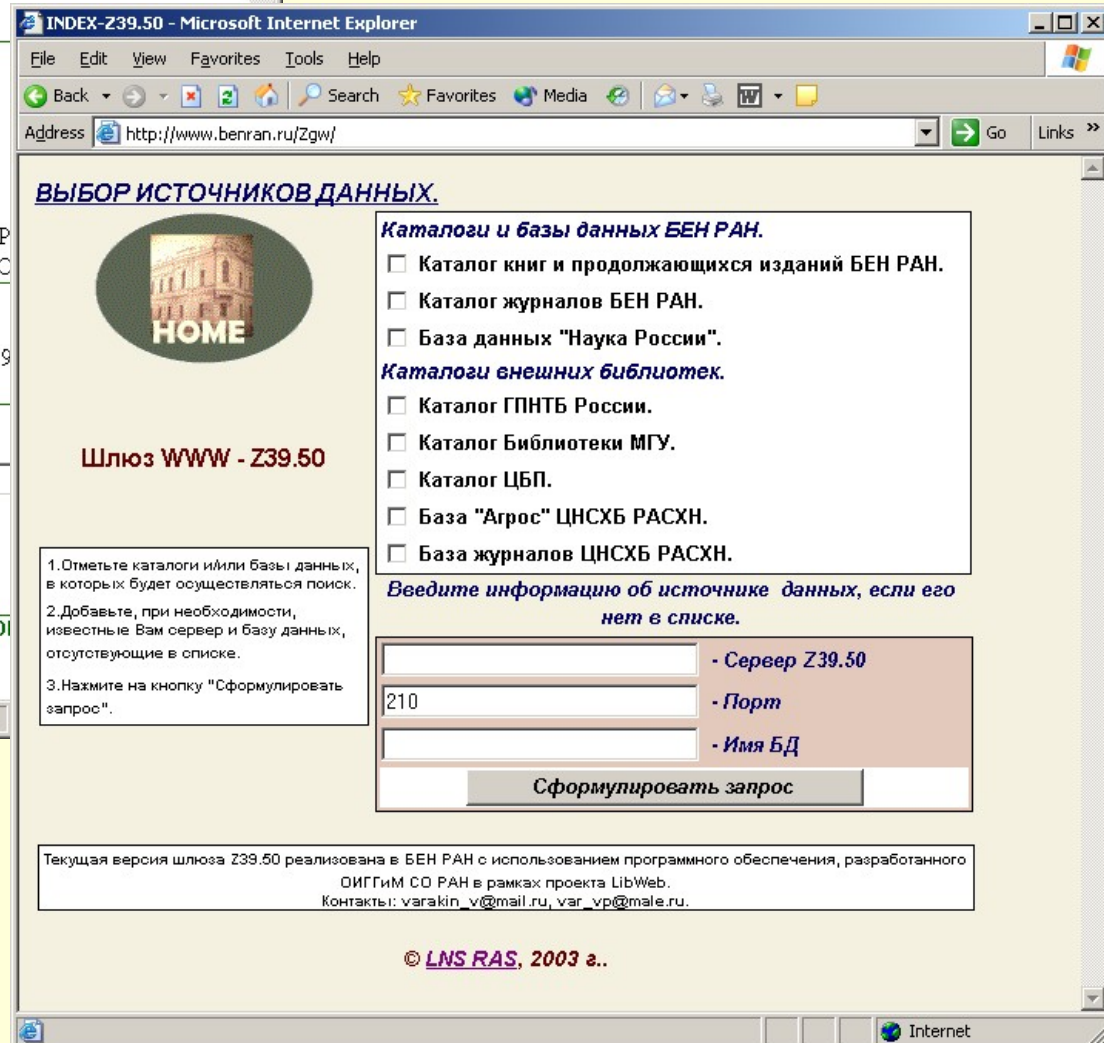
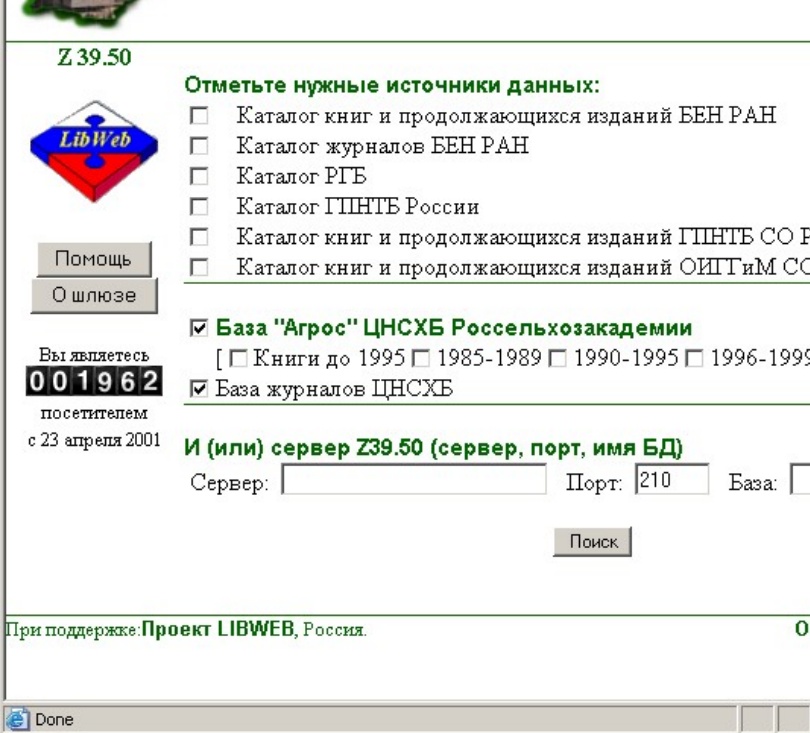
# ***LIBWEB***

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)



# LibWeb



# АРБИКОН

АРБИКОН :: Корпоративный поиск - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address <http://www.arbikon.ru/services/index.asp> Go Links

Новости Документы О проекте Обучение Сервисы Гостевая книга Авторизация

Поиск по ресурсам АРБИКОН

Поиск в аналитике РКП

Поиск в аналитике АРБИКОН

Медицинский портал АРБИКОН

Удалённый доступ к MeSH

Текущая библиография РКП

Портал библиотек АРБИКОН для слабовидящих

Поиск по ресурсам РНБ

Статистика

Справка

Стандарты

Корпоративный профиль АРБИКОН

Поиск по сайту

Ссылки

## Сервисы

При просмотре полных записей в режиме **расширенного поиска** участники АРБИКОНА имеют возможность **выгрузки записей** в форматах семейства MARC. Также участникам АРБИКОНА предоставляется доступ к **аналитике РКП**. Чтобы использовать эти возможности, необходима [авторизация](#)

[ [Расширенный поиск](#) ] [ [Справка](#) ]

### Ресурсы российских корпоративных библиотечных систем

Материалы	Базы данных
Все	RUSLANet
Монографии	Consensus Omnium
Сериальные	ОКБС Челябинского региона
Статьи, доклады	Ярославская КБС

Автор

И

Заглавие

И

Тематика

Количество извлекаемых записей

[ [Расширенный поиск](#) ] [ [Справка](#) ]

Done Internet

# ***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.

# ***GILS***

- Одной из первых попыток решения проблемы каталогизации электронных ресурсов было создание профиля ***GILS*** (***G***overnment ***I***nformation ***L***ocator ***S***ervice) для государственных учреждений США в 90-х годах прошлого века.
- Профиль ***GILS*** создавался вместе со стандартом *Z39.50* и является его частью. Профиль содержит описание схемы данных (метаданных) ***GILS***, включающей наборы поисковых атрибутов, набор меток, наборы элементов и структуру записи ***GILS***.
- Сегодня информационная система на основе ***GILS*** в США является составной частью, например, портала правительственной информации США (*FirstGOV*). В качестве примера точки доступа можно указать ***GPO*** (*U.S. ***G***overnment ***P***rinting ***O***ffice*) и многие другие.

Government Information Locator Service (GILS) - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media Go Links

Address [http://www.access.gpo.gov/su\\_docs/gils/index.html](http://www.access.gpo.gov/su_docs/gils/index.html)

Home | Privacy | Site Search | Help | Comments  
Site Contents | What's New | Online Bookstore | Finding Aids | Library Services

**GPO Access**  
Keeping America Informed  
Superintendent of Documents

## Government Information Locator Service

**Make your selection(s):**

- All Records on GPO Access GILS Site
- Pointer Records to other GILS sites
- Individual Agency GILS databases on GPO Access
  - ...Central Intelligence Agency
  - ...Consumer Product Safety Commission
  - ...Department of Commerce

**Enter search term(s):**

**Maximum records returned:**  Default is 40. Maximum is 200.

**Related Database:**

- Privacy Act Notices

The default searches all GILS records located on GPO Access as well as another WAIS-based Federal GILS sites: the National Aeronautical and Space Administration at NASA.

Phrases must be in quotation marks (" "). The operators ADJ (adjacent), AND, OR and NOT can be used, but must be capital letters. For example: "foreign affairs" AND "cd-rom". Word roots can be searched using an asterisk (\*) following the word stem. For example: legislat\* will retrieve both legislation and legislative.

Internet



# Электронные библиотеки

Ключевым моментом проектов по созданию и организации доступа к электронным библиотекам является разнородность интегрируемых информационных ресурсов.

# **LAURIN**

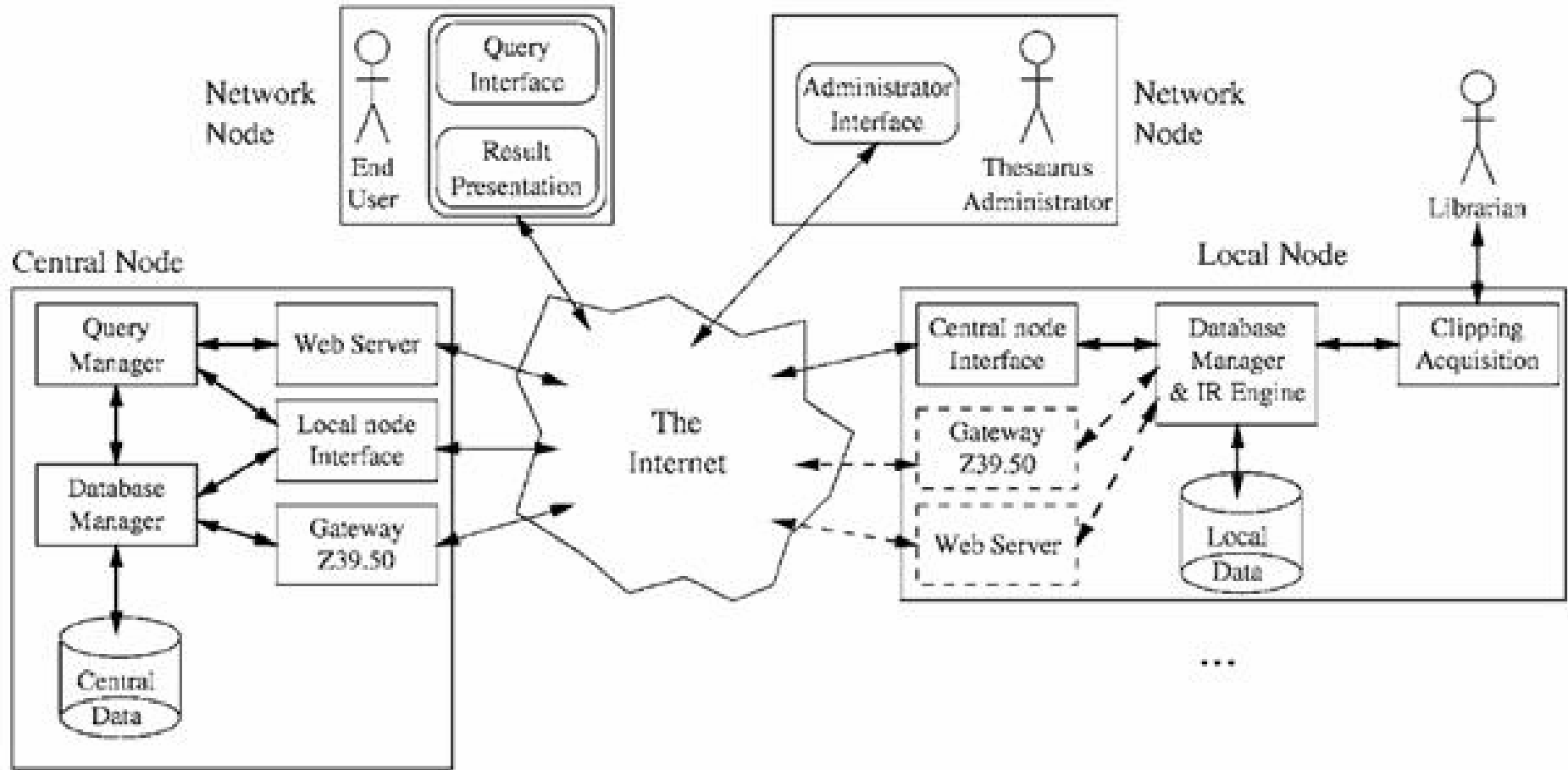
**LAURIN** (*Libraries and Archives Collecting Newspaper Clippings Unified for their Integration into Networks*) – проект по созданию информационной системы по газетным публикациям. Проект финансировался ЕС с мая 1998 года по август 2000 года в четвертой рамочной программе (*Telematics for Libraries*) с привлечением семнадцати участников от несколько стран.

Главные цели проекта:

- Создание прототипа распределенной информационной системы оцифрованных газет и архива вырезок, предоставляющей централизованный доступ через Internet для поиска и извлечения информации.
- Создание универсальной модели, которая использовалась бы библиотеками для сканирования, оцифровки, сохранения, и индексирования газетных вырезок, делая их доступными через **LAURIN** сеть.

# LAURIN

*архитектура информационной системы*



# ***DNER***

- ***DNER*** (***D****istributed* ***N****ational* ***E****lectronic* ***R****esource*) – 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* комитета *JISC* (***J****oint* ***I****nformation* ***S****ystems* ***C****ommittee*) содержит четыре проекта (*Xgrain*, *ZBLSA*, *ZETOC* и *DocuSend*), три из которых основаны на *Z39.50*.

# ZETOC

- Проект **ZETOC** заключался в создании сервисов для *MIMAS* (*Manchester Information & Associated Services*) 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 свыше 8000 статей. Количество журналов – 20 тыс., материалов конференций – 16 тыс.

**Address: zetoc.mimas.ac.uk**

**Port: 2121**

**Database name: zetoc**

**Record syntax: SUTRS, GRS-1 or XML**

# **EULER**

## **European Libraries and Electronic Resources** *in Mathematical Sciences*

Цель проекта - создание единой точки доступа в информационную систему, интегрирующую разнородные ресурсы (научные библиографические базы данных, библиотеки OPAC с услугами доставки документов, электронные журналы от академических издателей, архивы предпечатных изданий и серой литературы, предметно-ориентированные системы Интернет - порталы, различные поисковые системы Интернет), и предоставление пользователю единых интерфейсов. Пользовательские интерфейсы разрабатывались в тесном сотрудничестве с математическим сообществом.

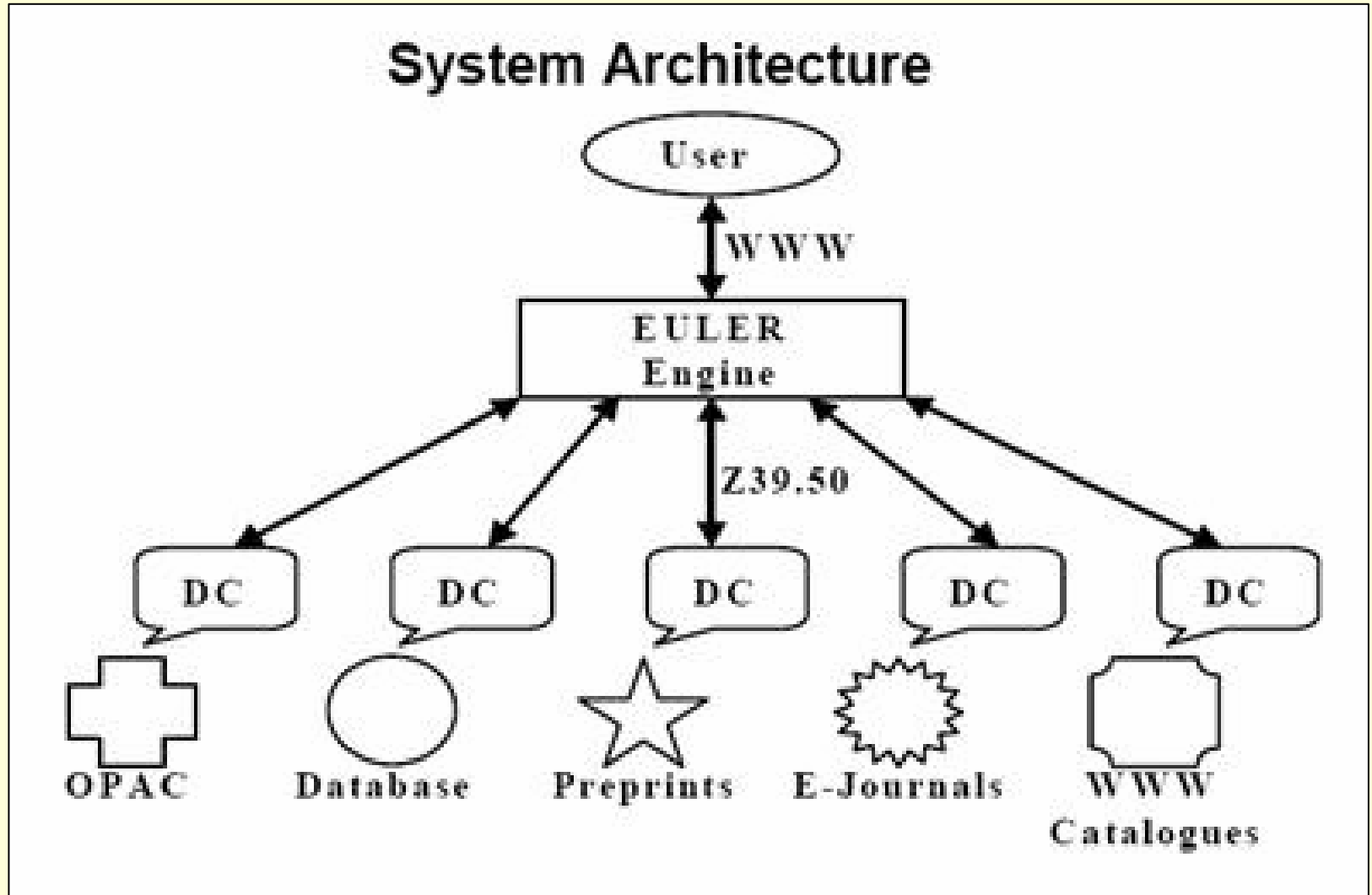
Проект финансировался с апреля 1998 года по ноябрь 2002 года Европейской Комиссией в рамках программы Telematics Applications подпрограммы Telematics for Libraries.

# ***EULER***

- Для интеграции информационных ресурсов использовались общие описания (метаданные), основанные на *Dublin Core*, доступ к которым предоставляется по протоколу *Z39.50*.
- Базы метаданных в проекте являются распределенными, каждая из них поддерживается своим поставщиком информации, интеграция ресурсов происходит в реальном времени. Фактически в проекте реализована модель виртуального каталога.

# Архитектура *EULER*

*EULER* основан на Dublin Core (DC)





# EULER - расширенный поисковый интерфейс пользователя

The screenshot shows a Microsoft Internet Explorer browser window displaying the EULER Advanced Search page. The browser's address bar shows the URL <http://www.emis.de/projects/EULER/advanced.html>. The page features a header with the Euler formula  $e^{ix} = \cos x + i \sin x$ , a portrait of Leonhard Euler, and the word "EULER" in large blue letters. A blue navigation bar contains the text "Advanced Search" and a link "All About EULER". Below this, there are two identical search forms. Each form has a "EULER Search" button on the left and a "Clear Form" button on the right. The search criteria are: Author, Title, Anywhere, Classification, and Date, each with a dropdown menu and an adjacent text input field. A link "See also: Simple Search" is positioned between the two forms. At the bottom of the page, there is a footer with copyright information: ©2001, 2002 EULER Consortium; supported by the IST Programme of the European Community: Project EULER-TAKEUP (IST-2000-29446); and EDBM indexing and retrieval software: © 2001 Cellule MathDoc, UJF & CNRS.

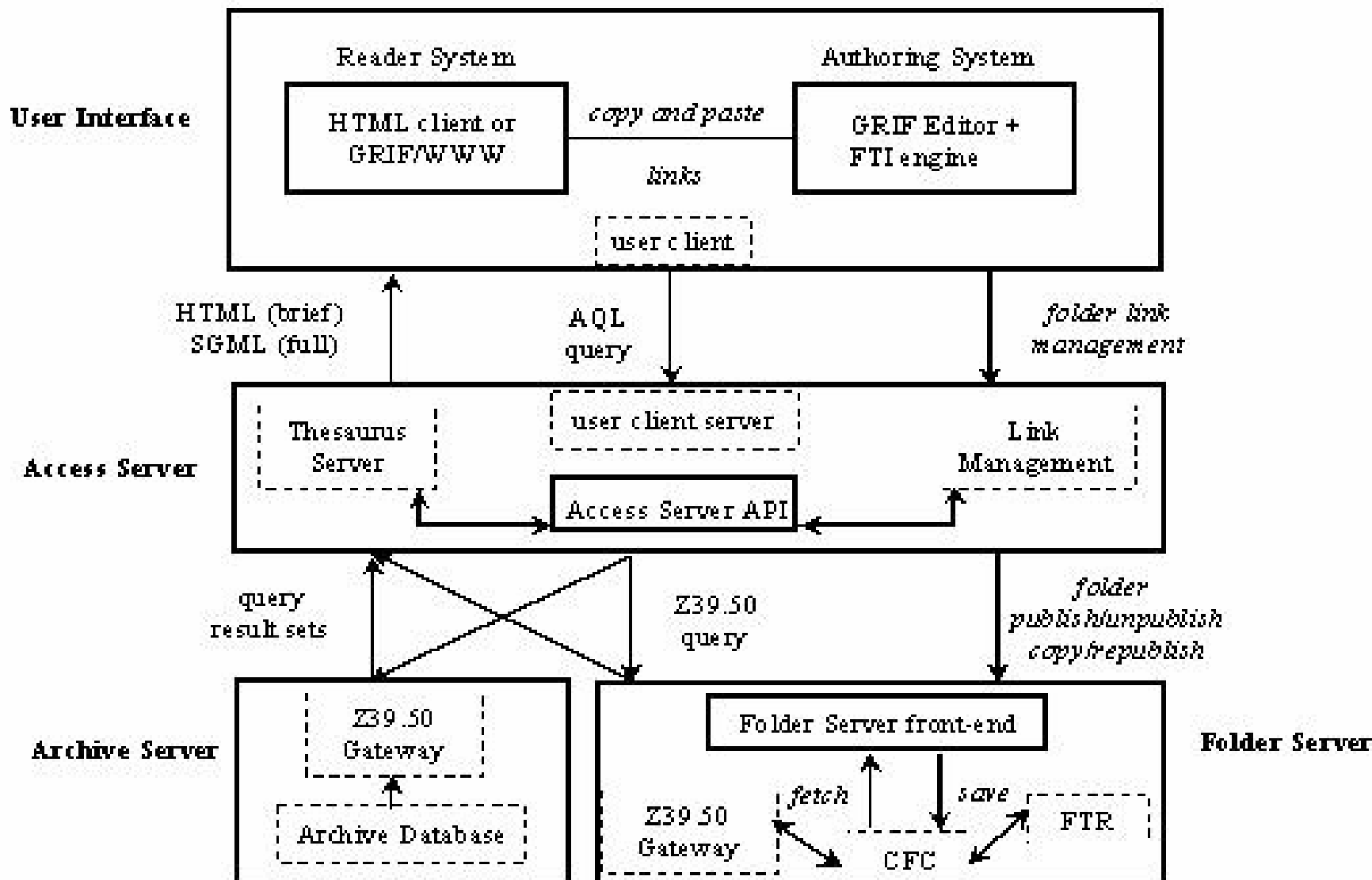
# Проекты по культурному наследию

Если для библиотечных проектов создания распределенных систем «знаковым» явился проект ***UNiverse***, то для проектов по культурному наследию «знаковым» явился проект ***Aquarelle***.

# *Aquarelle*

- ***Aquarelle*** – трехлетний проект в конце прошлого века, поддержанный *Telematics Applications Programme* Европейского союза, ориентирован на проектирование системы доставки ресурсов через Интернет, относящихся к культурному наследию.
- Информационная система ***Aquarelle*** основана на протоколе *Z39.50* и должна поддерживать доступ к гетерогенным базам данных, включая архивы *SGML*-документов.
- Метаданные в системе ***Aquarelle*** основаны на профиле *CIMI* (**C**onsortium for the **I**nterchange of **M**useum **I**nformation). Более того, сам профиль *CIMI* во многом обязан именно этому проекту

# Aquarelle



# Zavier

## Z39.50 Arts Victoria Interoperability Project

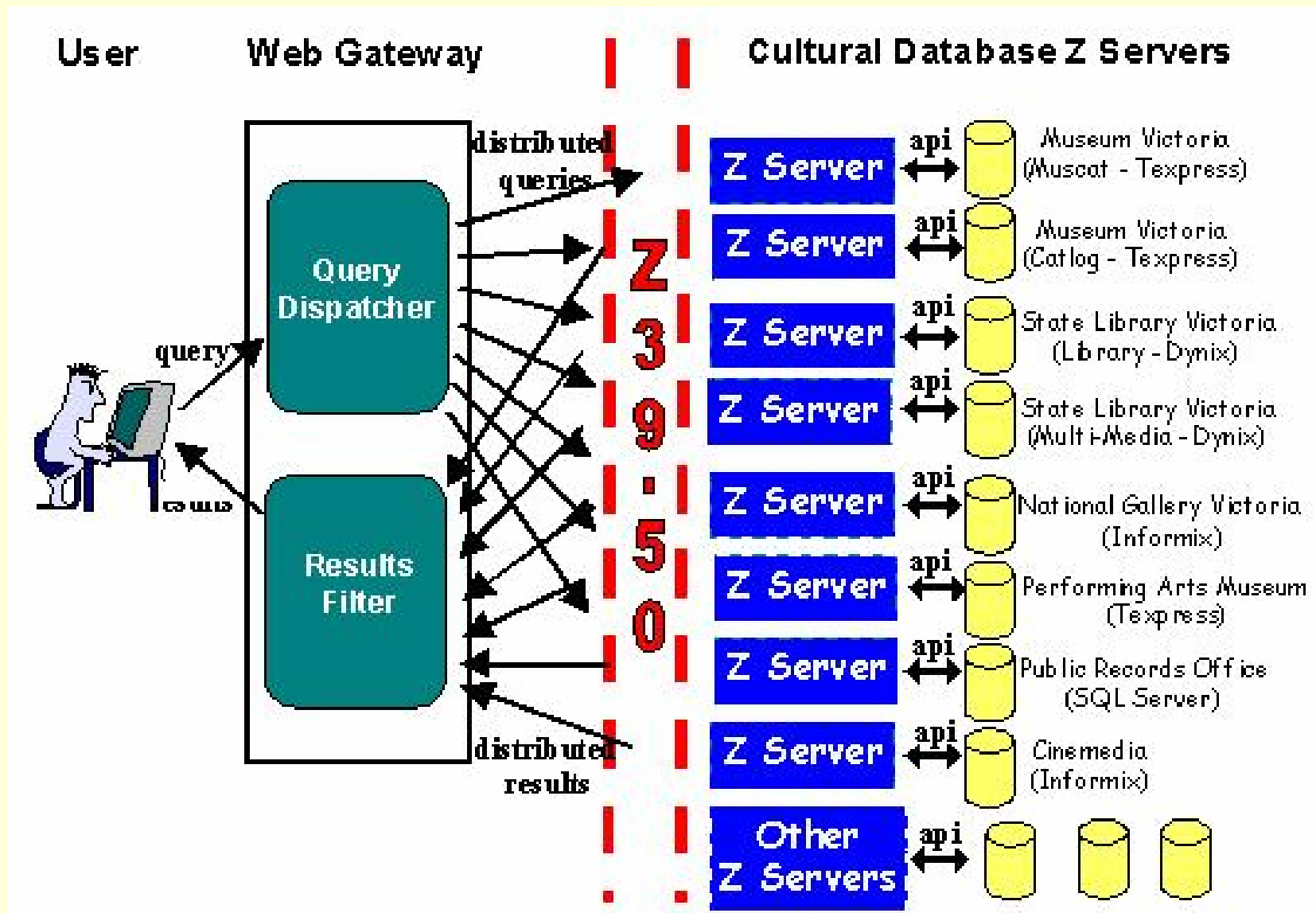
*Австралийский проект Zavier - экспериментальный проект создания точки доступа через Интернет к распределенным гетерогенным базам данных по викторианскому культурному наследию.*

*Главная цель проекта* - определение через экспериментальную систему возможности использования Z39.50 и профиля CIMI (**C**onsortium for the **I**nterchange of **M**useum **I**nformation) для сквозного поиска и извлечения информации из баз данных по культурному наследию.

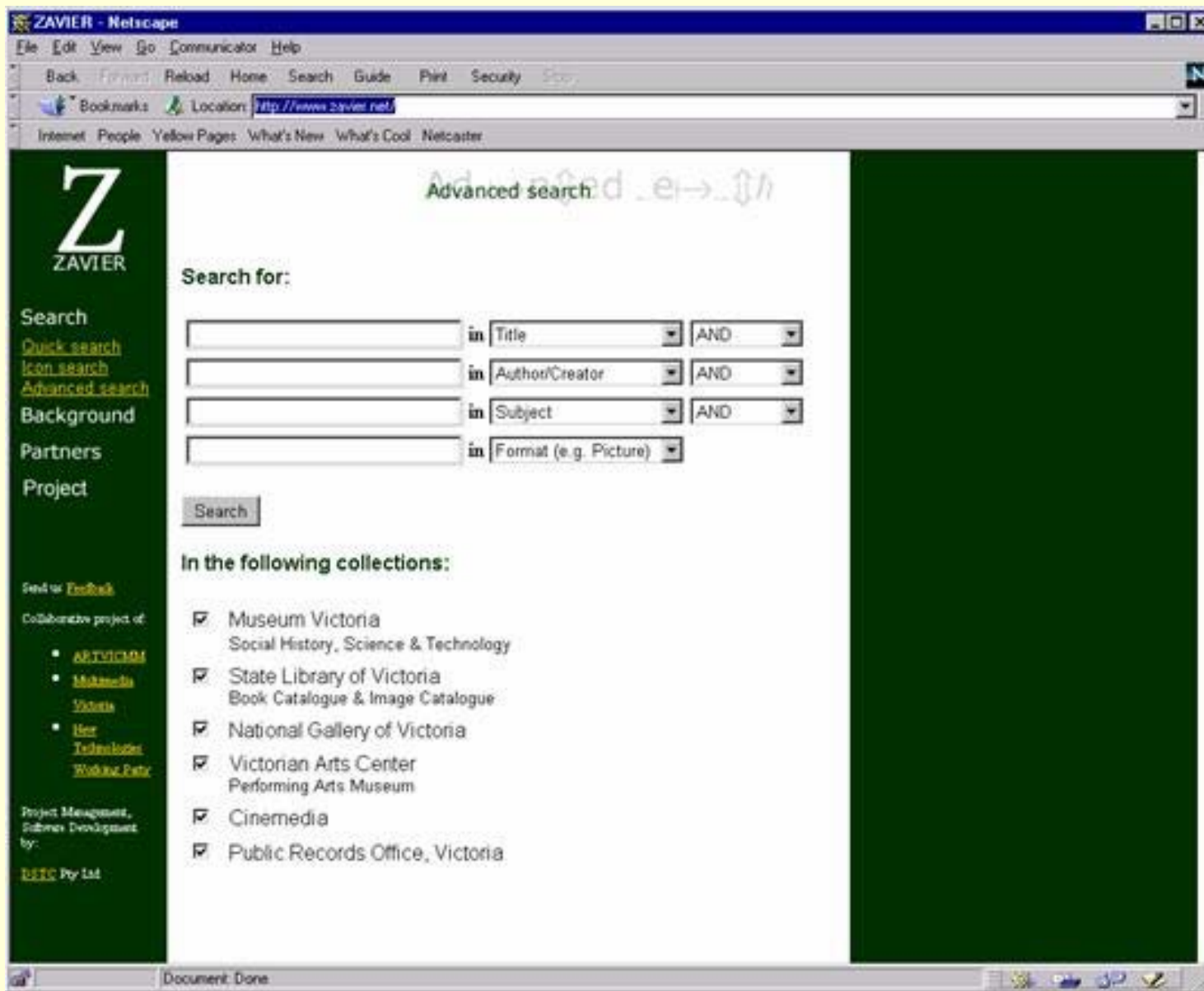
*В пилотном проекте были интегрированы ресурсы шести Викторианских агентств.*

# Архитектура *Zavier*

*Zavier* основан на CIMI



# Zavier - расширенный поисковый интерфейс пользователя



# **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** и **AHRD** (**Arts and Humanities Research Board**).
- В проекте отрабатываются технологии и стандарты, создаются методические рекомендации по следующим разделам: археология, история, прикладное искусство, языкознание и изобразительное искусство.



# **AHDS**

- Создаваемые в проекте базы данных описывают различные типы информации в соответствии с различными стандартами каталогизации. Доступ к данным обеспечивает различное программное обеспечение и 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** (**A**rchaeology **D**ata **S**ervice), **OTA** (**T**he **O**xford **T**ext **A**rchive), **VADS** (**V**isual **A**rts **D**ata **S**ervice) и другие.

# AHDS

## 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)

Refresh

**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"/>		

Use this drop-down box to determine how many search results are presented on each results page.

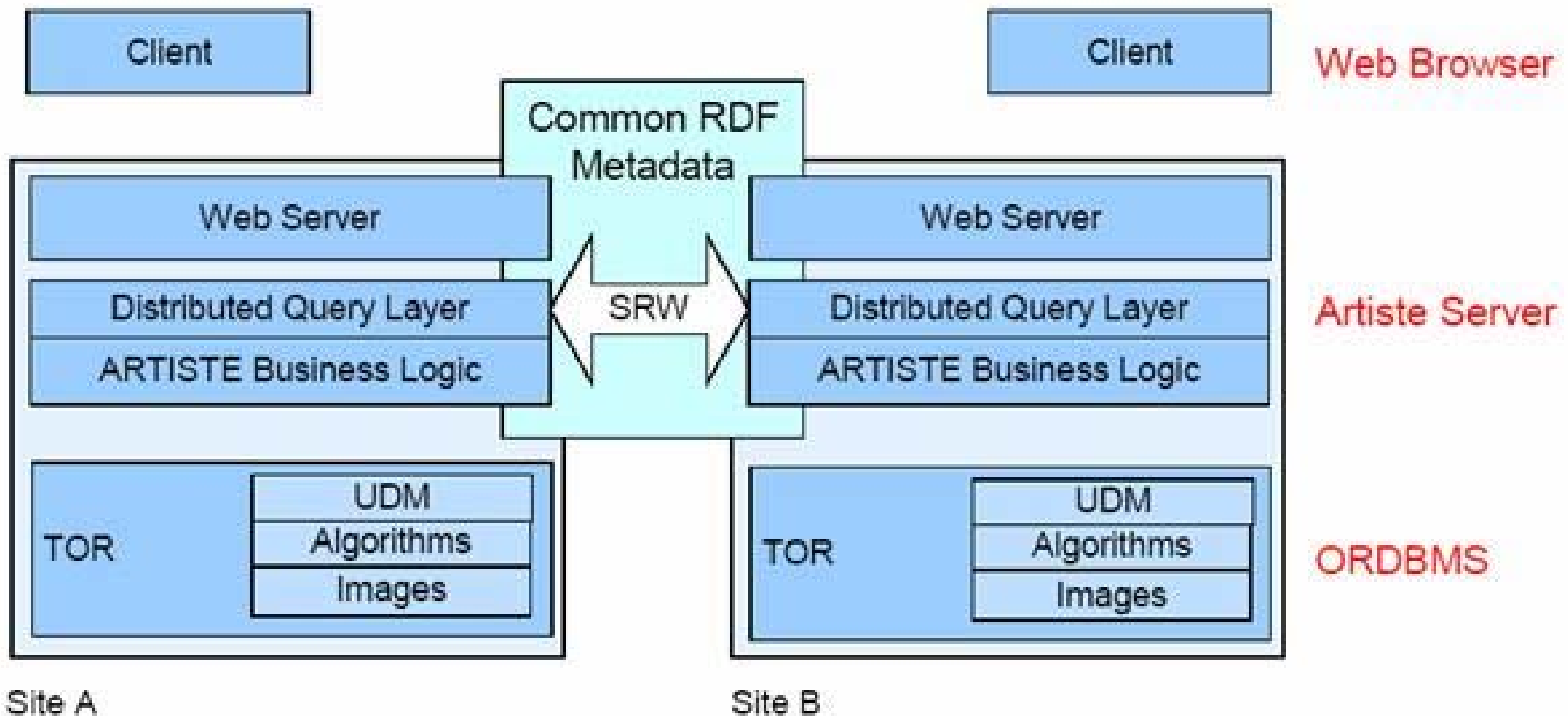
Records to display per page

# **ARTISTE**

- Проект **ARTISTE** (*An integrated Art Analysis and Navigation Environment*) с бюджетом 2,7 млн. евро, частично финансируемый Европейским Союзом через пятую 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 музеи в Лондоне, Лувр во Франции.
- Суммарно информационный ресурс проекта **ARTISTE** превышает 160 тыс. изображений и 5 млн. текстовых метаописаний.

# ARTISTE

Многоуровневая архитектура информационной системы **ARTISTE** основана на новом протоколе **SRW**, который является развитием протокола **Z39.50** (представление **Z39.50** в терминах **XML** и **SOAP** через **HTTP**).



# Геоинформационные проекты

*Несмотря на достаточно большой успех технологий, основанных на протоколе Z39.50, в области построения распределенных информационных систем, интегрирующих метадаанные библиотек, архивов, музеев и других, на наш взгляд, истинным бенефисом технологий Z39.50 следует считать внедрение его как основы в распределенные геоинформационные системы.*

# **FGDC Clearinghouse**

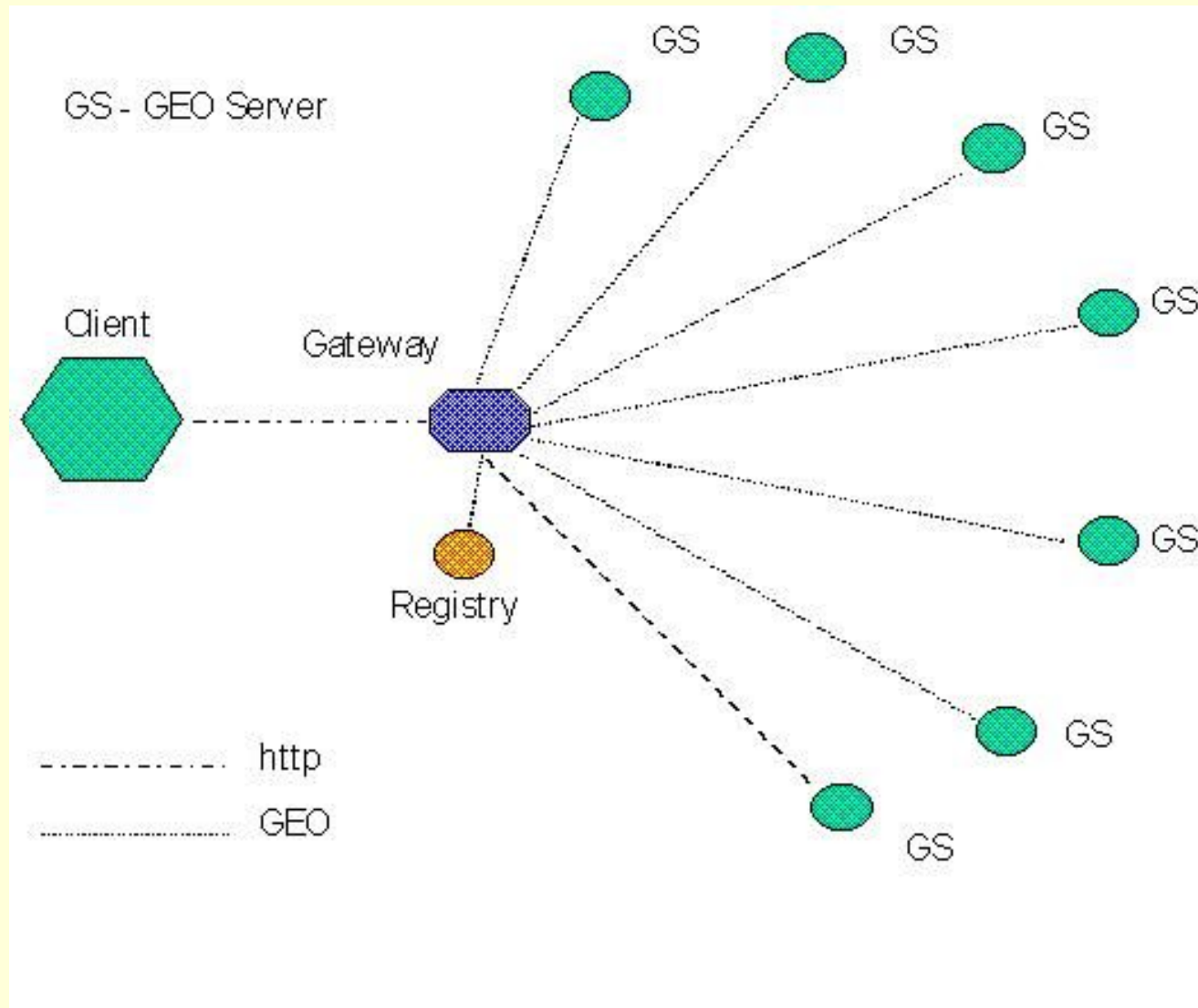
*Под эгидой Федерального Комитета по Географическим Данным США (FGDC) на основе профиля GEO и протокола Z39.50 была создана распределенная информационная система по пространственным метаданным **Clearinghouse**, включающая сотни серверов Z39.50.*

*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.*

# Архитектура *FGDC Clearinghouse*

*FGDC Clearinghouse* основан на GEO



# ***FGDC Clearinghouse***

По технологии ***FGDC Clearinghouse*** сегодня функционируют подобные системы для других предметных областей:

- ***SEDAC (Socioeconomic Data and Applications Center)*** – 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 (Content Standard for Digital Geospatial Metadata)***.
- ***GCMD (Global Change Master Directory)*** – в отличие ***Clearinghouse*** предоставляет доступ к централизованному каталогу (более чем 1500 различных организаций-участников). Предназначен в основном для научного сообщества в области наук о Земле.



# ***FGDC Clearinghouse***

- ***NBII Metadata Clearinghouse*** (*NBII – National Biological In-fo-r-ma-tion Infrastructure*) является инициативой **USGS (U.S. Geological Survey)** по созданию распределенной информационной системы, которая содержит метаданные, описывающие биологические данные и информационные продукты, и основана на подмножестве *CSDGM – NBII Biological Profile*.
- ***IAI-DIS (Inter-American Institute for Global Change Research – Data and Information System)*** – 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.

## **Выводы:**

- 1. Несмотря на достаточно "зрелый" возраст протокола Z39.50, интерес к нему в мировом информационном сообществе существует, т.к. продолжают появляться проекты построения информационных систем на основе этого протокола.*
- 2. Чаще всего протокол Z39.50 применяется в случаях*
  - по чисто историческим причинам (библиотечные системы);*
  - при необходимости интеграции гетерогенных информационных ресурсов на основе открытых международных стандартов;*
  - при построения надежных и интероперабельных информационных систем, находящихся в режиме промышленной эксплуатации.*
- 3. Практически во всех проектах используется уникальное серверное программное обеспечение Z39.50, применение которого связано с особенностями конкретных проектов.*
- 4. В качестве точки доступа в информационные системы чаще всего выступают шлюзы Z39.50 - HTTP.*
- 5. Просматриваются тенденции интеграции технологий Z39.50 с другими технологиями (XML, RDF, SOAP) в том числе и совсем новыми (SRW).*

**Применение протокола Z39.50:  
проекты и решения**

**Благодарю за внимание!**