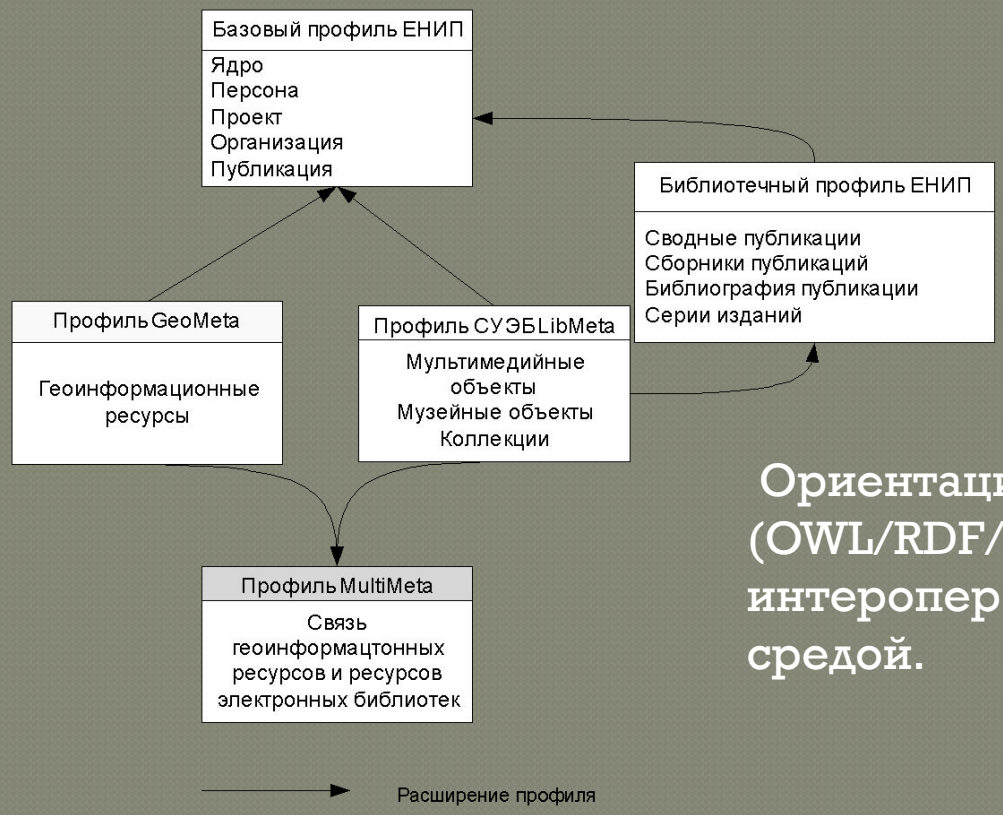
MultiMeta - Система интеграции пространственных данных и ресурсов электронных библиотек

О.М. Атаева, А.А. Каленкова, В.А. Серебряков
ВЦ РАН (Москва)

Постановка задачи

- Системы управления пространственными данными позволяют интегрировать данные и знания о территории, строить и использовать модели природных и социально-экономических явлений и процессов, их взаимодействия в системе «общество – природная среда», использовать методы пространственного анализа, обеспечивать территориальное планирование и управление.
- В силу специфики своей предметной области участники процесса интеграции пространственных данных обладают данными и метаданными, дополняющими пространственную информацию: это могут быть публикации, включающие набор метаданных и полный текст, музейные экспонаты, представленные метаданными и медиа-объектами, самостоятельные медиа-объекты такие как фотографии, видео и аудиозаписи.
- Интеграция разнородной информации, предоставляемой несколькими источниками, позволит максимально составить описание территории, провести анализ, выявить взаимосвязи, получить принципиально новые возможности поиска: ресурсы, созданные в разных системах, представляются связанными друг с другом единой системой навигации.

Схема метаданных MultiMeta



Ориентация на семантический веб (OWL/RDF/RDFS) , возможности интероперабельности с внешней средой.

Портал «ГеоМета»

- Портал «ГеоМета» – это стандартизированная и децентрализованная **среда управления пространственной информацией**, разработанная для доступа к базам геоданных, картографическим продуктам и связанным с ними метаданным из различных источников, облегчающая обмен пространственной информацией между организациями и ее совместное использование посредством интернета.

- Портал «ГеоМета» представляет собой платформу для создания распределенной среды интеграции неоднородных источников геоинформационных данных и предоставления к этой среде единой точки входа (веб-портала), которая позволит ученым в сфере наук о Земле легко находить специализированные данные и приложения, производить вычислительные эксперименты, визуализировать результаты деятельности.

Основные особенности портала «ГеоМета»

- Система поддерживает следующие основные типы ресурсов: *Пространственные данные* (картографические данные и их метаданные) и дополнительные типы ресурсов, такие, как *Организация, Персона, Публикация, Проект* и различные рубрикаторы и классификаторы.
- К функциональностям ГИС-части системы относятся:
  - каталогизация, сбор, поиск геопространственных метаданных;
  - размещение геоданных в собственном или удаленном хранилище и предоставление к ним доступа;
  - предоставление доступа к распределенным геопространственным данным по стандартизованным протоколам;
  - визуализация карт, редактирование элементов.

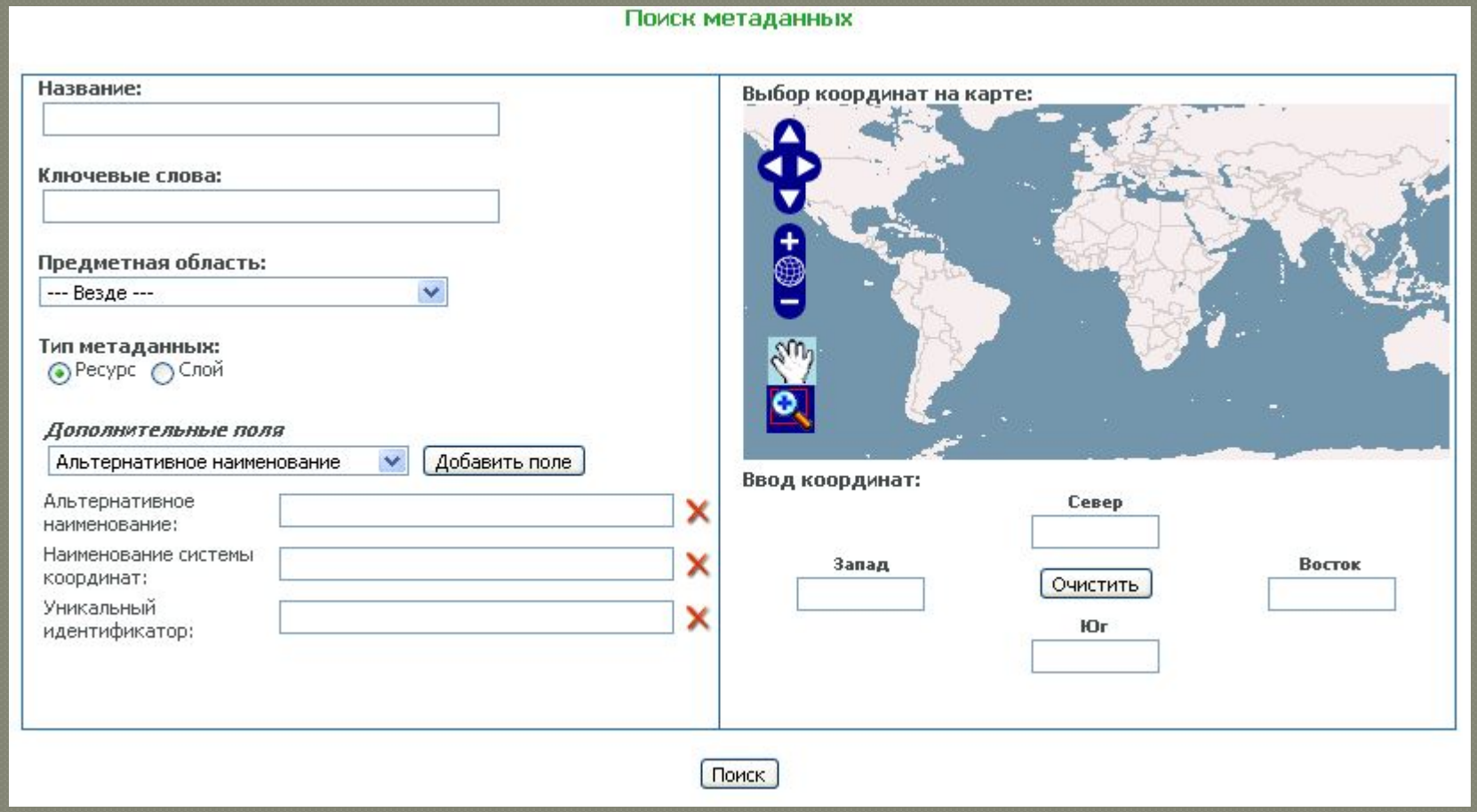
Пространственные метаданные

- Задача интеграции распределенных неоднородных источников пространственных данных и сервисов решается путем выделения базовой схемы метаданных, реализованной в виде OWL-онтологии на основе стандартов ISO 19115:2003 «Географическая информация. Метаданные» и ISO 19139:2007 «Географическая информация. Метаданные – XML схема» и полностью включающей российский стандарт ГОСТ Р 52573-2006 «Географическая информация».

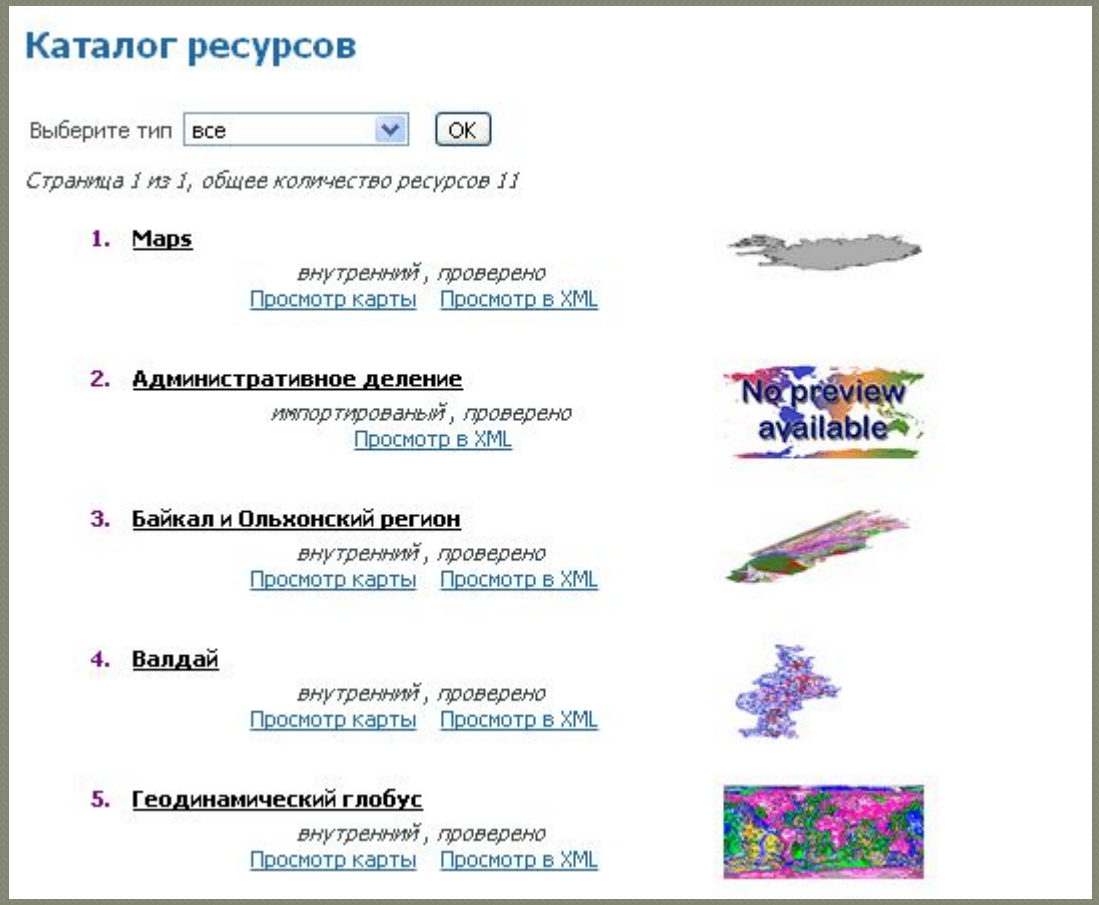
Функции системы «ГеоМета»

- **Поиск основных геоинформационных ресурсов**
- **Просмотр метаданных и пространственных данных**
- **Регистрация пространственных геоинформационных ресурсов**
- **Периодический сбор метаданных (harvesting)**

Поиск геоинформационных ресурсов



Страница Каталога ресурсов



Просмотр метаданных

Идентификационная информация о данных (Identification Information):

Citation (Общие сведения):
- **Title** (Наименование ресурса): *Байкал и Ольхонский регион*
- **Alternative title** (Альтернативное наименование):
- **Date** (Дата создания ресурса):
- **Presentation Form**(Форма представления данных): *Аналоговая модель*

- **Abstract** (Краткое содержание):
- **Purpose** (Назначение набора данных):
- **Status**(Состояние):

Data Identification (Идентификационные данные):
- **Spatial Representation Type** (Способ представления данных):
- **Spatial Resolution** (Пространственное разрешение):
- **Dataset language** (Язык описания данных):
- **Character set** (Кодировка):
- **Topic category** (Предметная область данных):
- **Spatial Domain extent:**
  - **Bounding Coordinates** (Географические координаты данных):
    - **West Bounding Coordinate** (Долгота крайней западной точки): *106.239*
    - **East Bounding Coordinate** (Долгота крайней восточной точки): *106.979*
    - **North Bounding Coordinate** (Широта крайней северной точки): *53.0741*
    - **South Bounding Coordinate** (Широта крайней южной точки): *52.6389*

- **Supplemental Information** (Дополнительные сведения):
- **Status of Date of Capture** (Описание создания):
- **Date of Capture (DD.MM.YYYY)** (Дата создания):

Browse Graphic (Пример изображения для просмотра):
- **File Name** (Файл графического представления): *Baykal.JPG*
- **File Description** (Описание графического представления):

Информация о слоях

⊞ **Слой 1** (tecton) **Просмотр в XML** (Просмотреть в ГИС-клиенте)

⊞ **Слой 2** (rework) **Просмотр в XML** (Просмотреть в ГИС-клиенте)

⊞ **Слой 3** (metzon) **Просмотр в XML** (Просмотреть в ГИС-клиенте)

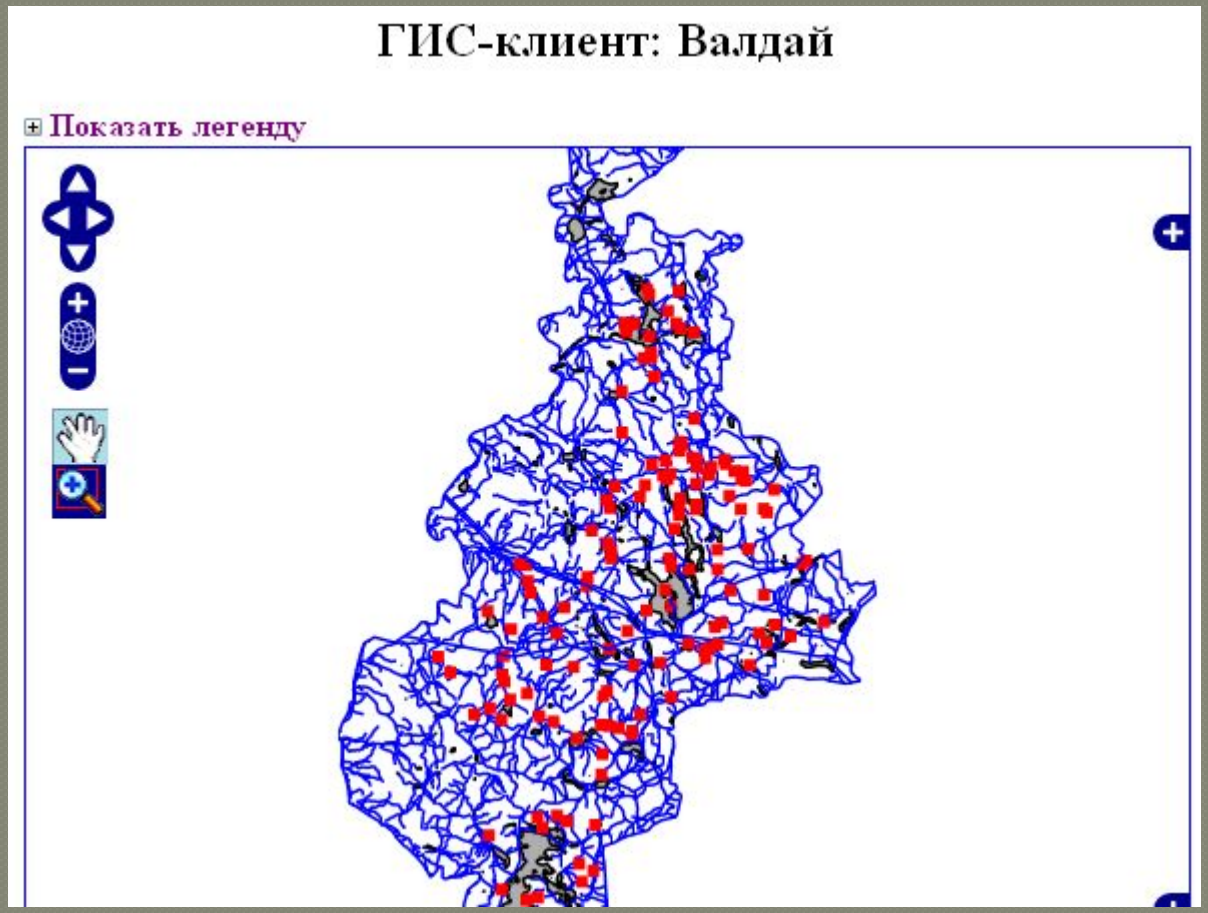
⊞ **Слой 4** (lines) **Просмотр в XML** (Просмотреть в ГИС-клиенте)

⊞ **Слой 5** (dep) **Просмотр в XML** (Просмотреть в ГИС-клиенте)

Просмотр метаданных в формате XML


```
  <?xml version="1.0" encoding="Windows-1251" ?>
- <gmd:MD_Metadatauri="ea250dfb-6bec-49f4-a9ba-
    c23339c44b34"xsi:SchemaLocation="http://www.isotc211.org/2005/gco ../gco/gco.xsd
    http://www.isotc211.org/2005/gmd ../gmd/gmd.xsd"xmlns="http://www.isotc211.org/2005/gmd"xmlns:xsi="http://ww
    -
    instance"xmlns:gco="http://www.isotc211.org/2005/gco"xmlns:gmd="http://www.isotc211.org/2005/gmd"xmlns:gm="htt
    <gmd:contact />
  - <gmd:fileIdentifier>
      <gco:CharacterString>ea250dfb-6bec-49f4-a9ba-c23339c44b34</gco:CharacterString>
    </gmd:fileIdentifier>
  - <gmd:referenceSystemInfo>
    - <gmd:MD_ReferenceSystem>
      - <gmd:referenceSystemIdentifier>
        - <gmd:RS_Identifier>
          - <gmd:code>
              <gco:CharacterString>EPSG:4326</gco:CharacterString>
            </gmd:code>
          - <gmd:codeSpace>
              <gco:CharacterString>EPSG:4326</gco:CharacterString>
            </gmd:codeSpace>
          </gmd:RS_Identifier>
        </gmd:referenceSystemIdentifier>
      </gmd:MD_ReferenceSystem>
    </gmd:referenceSystemInfo>
```

Просмотр изображения с помощью интерфейса портала ГеоМета



Просмотр изображения с помощью интерфейса портала ГеоМета

ГИС-клиент содержит следующие элементы:

1. Кнопки управления смещением карты
2. Кнопки управления масштабом
3. Режим управления картой мышью
4. Режим выбора области просмотра
5. Панель выбора слоев
6. Мини-карта
7. Масштабная линейка
8. Панель отображения легенды карты
9. При одиночном клике мышью на карте выводится всплывающее окно с результатами соответствующего GetFeatureInfo-запроса к WMS-серверу.

Регистрация геопространственного ресурса

Идентификационная информация о данных | Информация о координатной основе | Информация о распространении | Информация о качестве | Информация о метаданных | Слои

Общие сведения | Идентификационные данные | Пример изображения для просмотра | Ключевые слова | Связанные публикации и проекты

Регистрация ресурса

Идентификационная информация/Общие сведения

Наименование ресурса*:

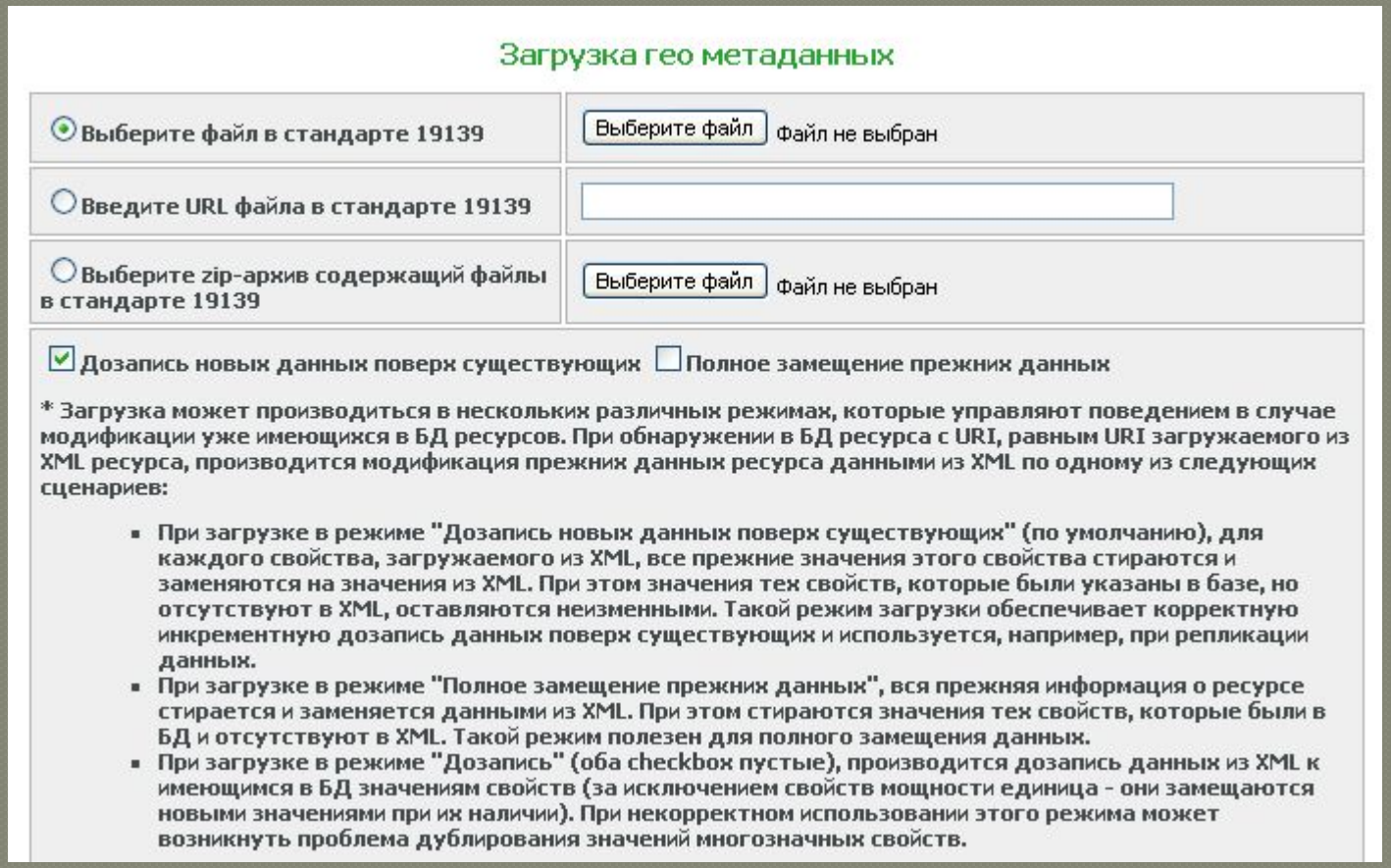
Альтернативное наименование:

Дата создания ресурса*: Выбрать

Форма представления данных: Аналоговая модель

Краткое содержание*:

Загрузка метаданных в формате XML



Загрузка гео метаданных

Выберите файл в стандарте 19139	Выберите файл Файл не выбран
Введите URL файла в стандарте 19139	
Выберите zip-архив содержащий файлы в стандарте 19139	Выберите файл Файл не выбран

Дозапись новых данных поверх существующих Полное замещение прежних данных

* Загрузка может производиться в нескольких различных режимах, которые управляют поведением в случае модификации уже имеющихся в БД ресурсов. При обнаружении в БД ресурса с URI, равным URI загружаемого из XML ресурса, производится модификация прежних данных ресурса данными из XML по одному из следующих сценариев:

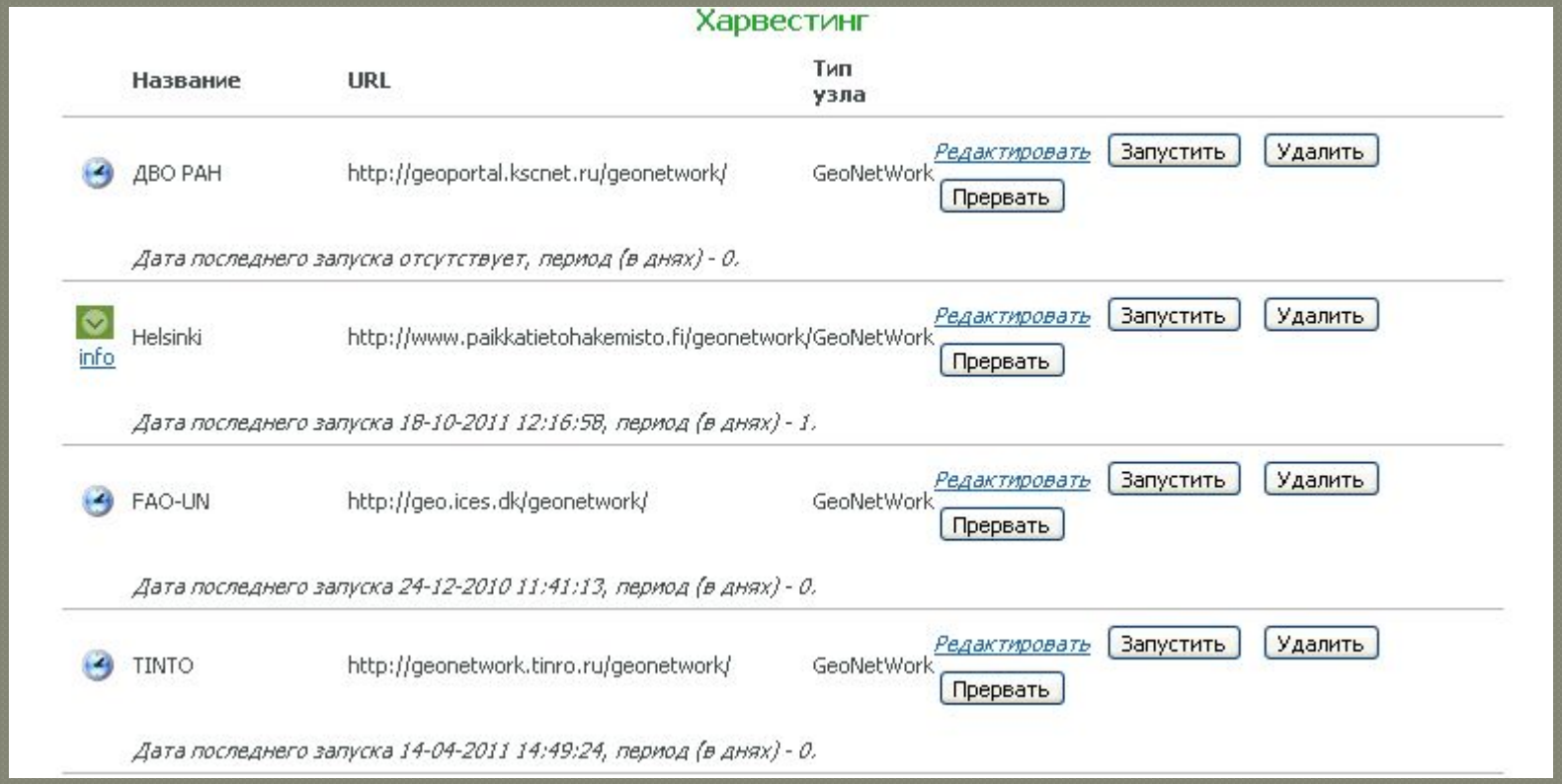
- При загрузке в режиме "Дозапись новых данных поверх существующих" (по умолчанию), для каждого свойства, загружаемого из XML, все прежние значения этого свойства стираются и заменяются на значения из XML. При этом значения тех свойств, которые были указаны в базе, но отсутствуют в XML, оставляются неизменными. Такой режим загрузки обеспечивает корректную инкрементную дозапись данных поверх существующих и используется, например, при репликации данных.
- При загрузке в режиме "Полное замещение прежних данных", вся прежняя информация о ресурсе стирается и заменяется данными из XML. При этом стираются значения тех свойств, которые были в БД и отсутствуют в XML. Такой режим полезен для полного замещения данных.
- При загрузке в режиме "Дозапись" (оба checkbox пустые), производится дозапись данных из XML к имеющимся в БД значениям свойств (за исключением свойств мощности единица - они замещаются новыми значениями при их наличии). При некорректном использовании этого режима может возникнуть проблема дублирования значений многозначных свойств.

Периодический сбор метаданных (harvesting)

В настоящее время реализованы два протокола обмена:

- GeoNetwork – порталы на базе ресурсов системы «Geonetwork opensource» и
- Web DAV – каталоги на базе сервера WebDAV, используются, в частности, для обмена между системами ГеоМета.
- Обмен метаданными между системами ГеоМета и «GeoNetwork» выполняется по стандартному протоколу «GeoNetwork 2.1» для сбора данных с узлов типа «GeoNetwork». Для «GeoNetwork» метаданные с узлов ГеоМета доступны по протоколу WebDav.

Список узлов опроса



Харвестинг

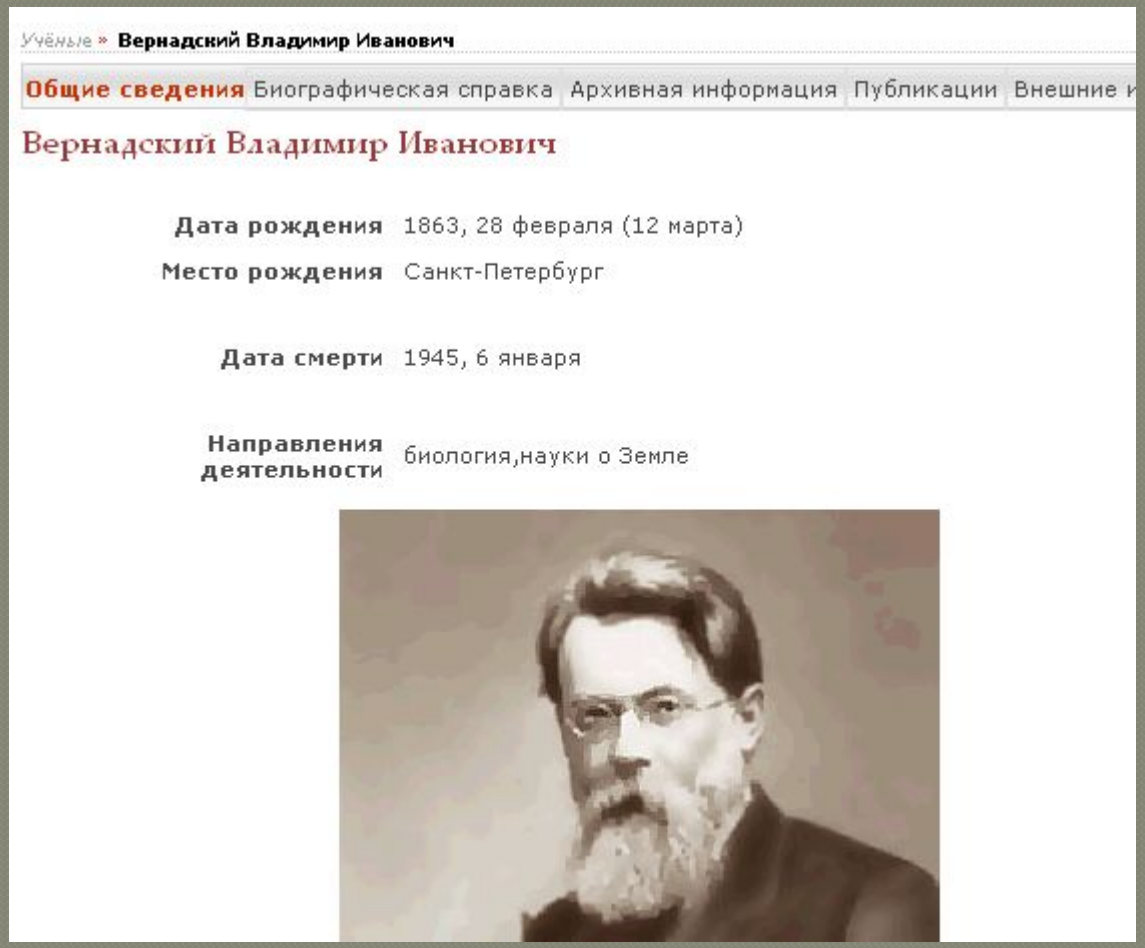
	Название	URL	Тип узла			
	ДВО РАН	http://geoportal.kscnet.ru/geonetwork/	GeoNetWork	Редактировать	Запустить	Удалить
				Прервать		
	Дата последнего запуска отсутствует, период (в днях) - 0.					
info	Helsinki	http://www.paikkatietohakemisto.fi/geonetwork/	GeoNetWork	Редактировать	Запустить	Удалить
				Прервать		
	Дата последнего запуска 18-10-2011 12:16:58, период (в днях) - 1.					
	FAO-UN	http://geo.ices.dk/geonetwork/	GeoNetWork	Редактировать	Запустить	Удалить
				Прервать		
	Дата последнего запуска 24-12-2010 11:41:13, период (в днях) - 0.					
	TINTO	http://geonetwork.tinro.ru/geonetwork/	GeoNetWork	Редактировать	Запустить	Удалить
				Прервать		
	Дата последнего запуска 14-04-2011 14:49:24, период (в днях) - 0.					

Профиль метаданных СУЭБ LibMeta

- Разработаны дополнительные прикладные профили поддержки музейной деятельности и мультимедийных представлений
- К основным типам данных, представленных в СУЭБ LibMeta, относятся Публикации, Персоны (авторы) и Предметы

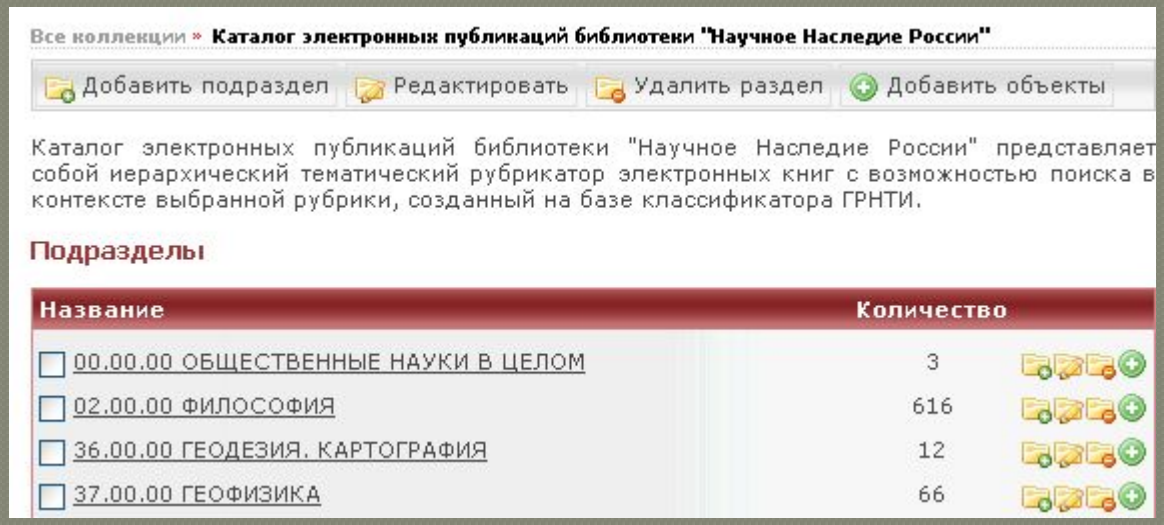
Общая архитектура СУЭБ LibMeta

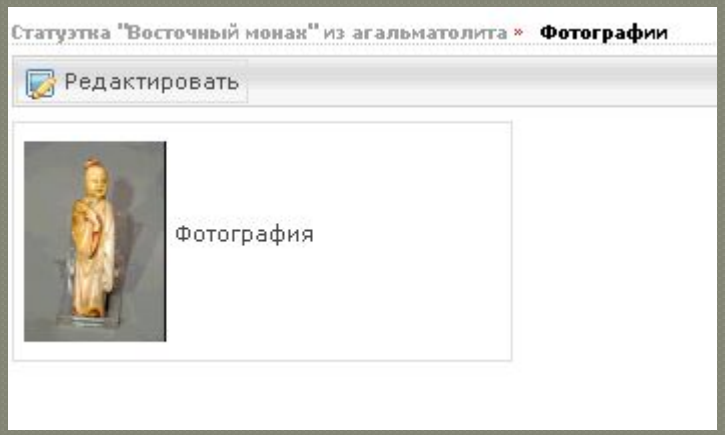
- Подсистема работы с метаданными об ученых, публикациях, музейных объектах позволяет просматривать, редактировать, а также производить поиск информации об ученом, публикации, музейном объекте.



Общая архитектура СУЭБ LibMeta

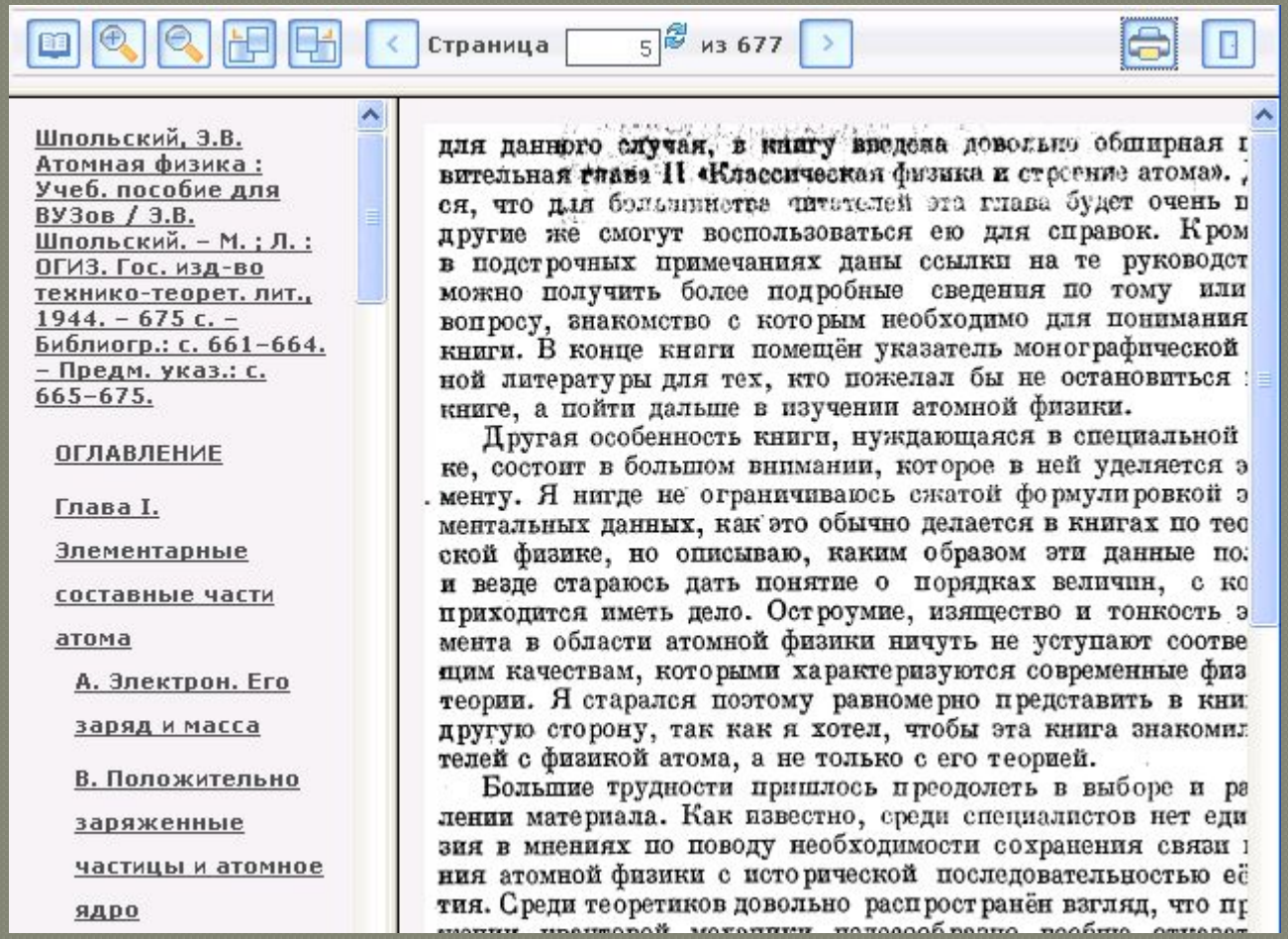
- Подсистема работы с коллекциями позволяет просматривать, редактировать и выполнять поиск по коллекции.
- Подсистема работы с наборами дополнительных атрибутов дает возможность создавать наборы атрибутов, назначать их некоторому музейному предмету.
- Подсистема работы с медиа-объектами позволяет просматривать и редактировать медиа-объекты.





Общая архитектура СУЭБ LibMeta

- Подсистема хранения и просмотра отсканированн ых текстов дает возможность просматривать подряд страницы издания, переходить на любую заданную страницу (в том числе на предыдущую, на последующую, на страницу с заданным номером).



Интеграция СУЭБ LibMeta с другими информационными системами

- Реализован OAI-PMH протокол, СУЭБ LibMeta может выступать как в качестве клиента, так и в качестве сервера.
- Схема представления музейной информации реализована в соответствии с концептуальной моделью CIDOC-CRM.

просмотра геоинформационного ресурса

Связанные объекты

Публикации

Почвы бассейнов Урюмкана и Газимура (Забайкальской области)

Музейные объекты

Друза длиннопризматических кристаллов гипса

Коллекции

Коллекция музейных предметов

Мультимедийные материалы

Фотографии

Цветок

Озеро

Голоса птиц

Черный коршун

MultiMeta – план работ

- Определение координат для музейных, мультимедийных и прочих непространственных ресурсов.
- Загрузка словарей и классификаторов территорий, привязка ресурсов к термам словарей и классификаторов.
- Показ ссылок на непространственные ресурсы на карте.

Спасибо за внимание!