

# К вопросу о долгосрочном сохранении цифровой информации

Шварцман М.Е.

Таруса 2007

# ЧТО СОХРАНЯТЬ

- материалы изначально созданные на компьютере и не имеющие печатного аналога
- он-лайновые публикации
- Web-сайты
- Блоги, форумы
- почтовые рассылки
- личные фото и видео архивы
- Интернет радио и TV
- CD и DVD
- и еще многое...

# принципиальное отличие ЦМ

- Задавая любую разумную периодичность копирования сайта, мы заранее понимаем, что можем что-то потерять возникающее и исчезающее в этих промежутках.
- Изменчивость ЦМ также должна отражаться и в их библиографическом описании

# принципиальное отличие ЦМ

- В процессе создания и бытования ЦМ они проходят множество этапов, на каждом из которых, они претерпевают изменения,
- В печатном документе, мы сохраняем одновременно и форму и содержание. Для ЦМ возможно существование в различных формах

# принципиальное отличие ЦМ

- . А смогут ли будущие поколения, а мы мыслим вечными категориями, прочесть то, что мы сохранили
- Могут ли они прочитать технически
- Могут ли они прочитать программно

# программы, проекты, рабочие группы

- NESTOR “Network of Expertise in Long-term Storage of Digital Resources”
- KORAL (Cooperative development of a long-term archive for digital information)
- Framework Programme 6 Call 5. Planets с бюджетом 14 000 000 €
- Digital preservation Europe
- CASPAR.

# программы, проекты, рабочие группы

- В королевской библиотеке Нидерландов создан отдел долговременной сохранности ЦМ. (Digital Preservation Department) В штате отдела состоят 8 человек, работающих полный день. При этом отдел занимается **только** теоретическими разработками и выдачей рекомендаций отделам библиотеки, которые занимаются оцифровкой или комплектованием ЦМ.

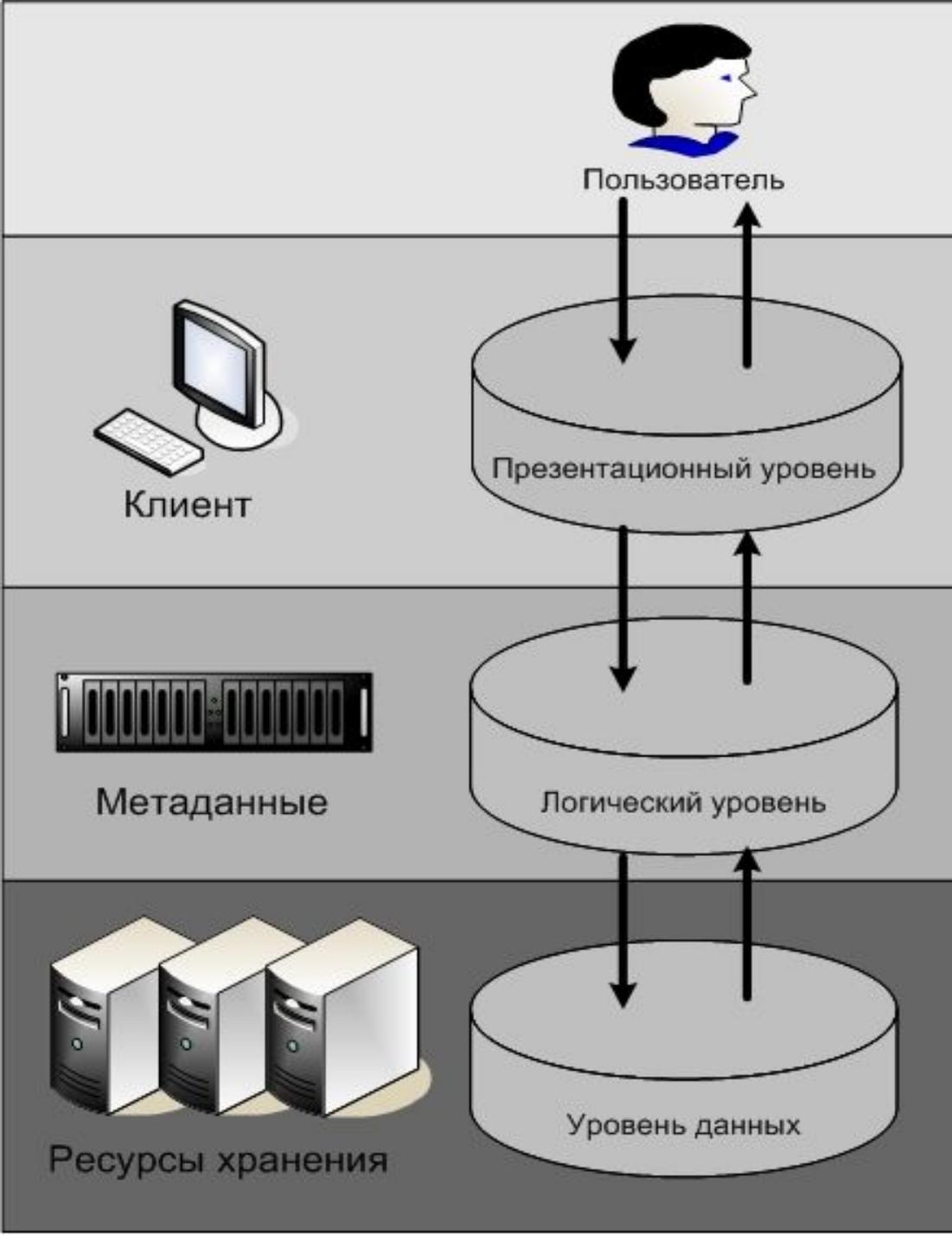
# основные рекомендации

- Проблему долговременного сохранения ЦМ нужно начинать решать на этапе их создания.
- более 65% стоимости технологического процесса жизненного цикла ЦМ составляет обеспечение сохранности,
- Нельзя решать проблему сохранности отдельно, она должна быть увязана в единый технологический цикл создания – обработки – хранения – использования ЦМ.

# Что у нас

Электронный депозитарий, обеспечивающей «вечную» сохранность электронных информационных ресурсов РГБ и создание объективных условий возможности воспроизведения всех ЭИР, с учетом замены (модернизации) программно-аппаратной среды.

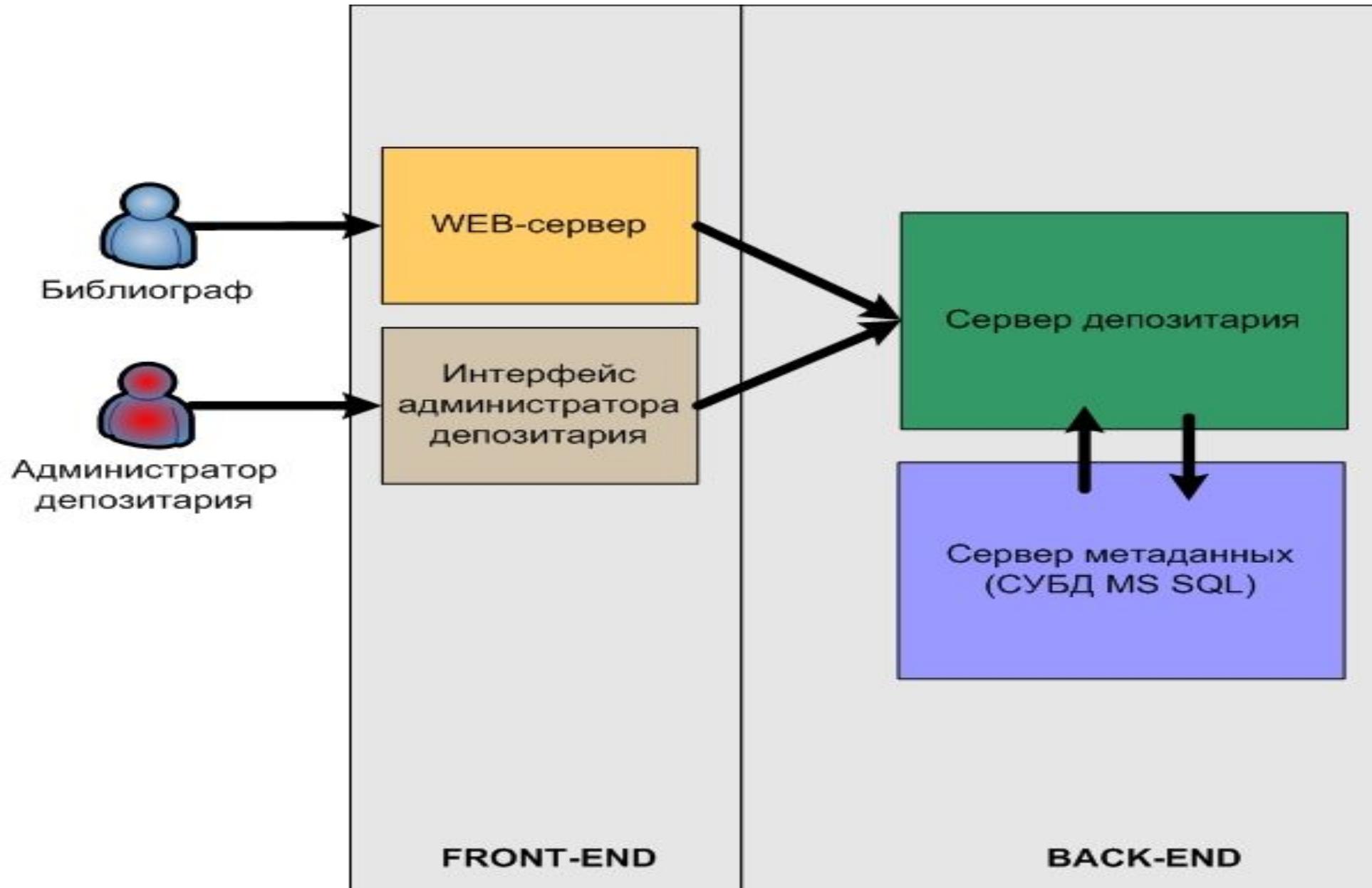
# Логическая архитектура



# Сервер метаданных

- Сервер метаданных представляет собой выделенный узел с установленной на нем реляционной базой данных (СУБД), в которой хранятся метаданные. Сервер метаданных может быть реплицирован с использованием утилит, входящих в дистрибутив базы данных

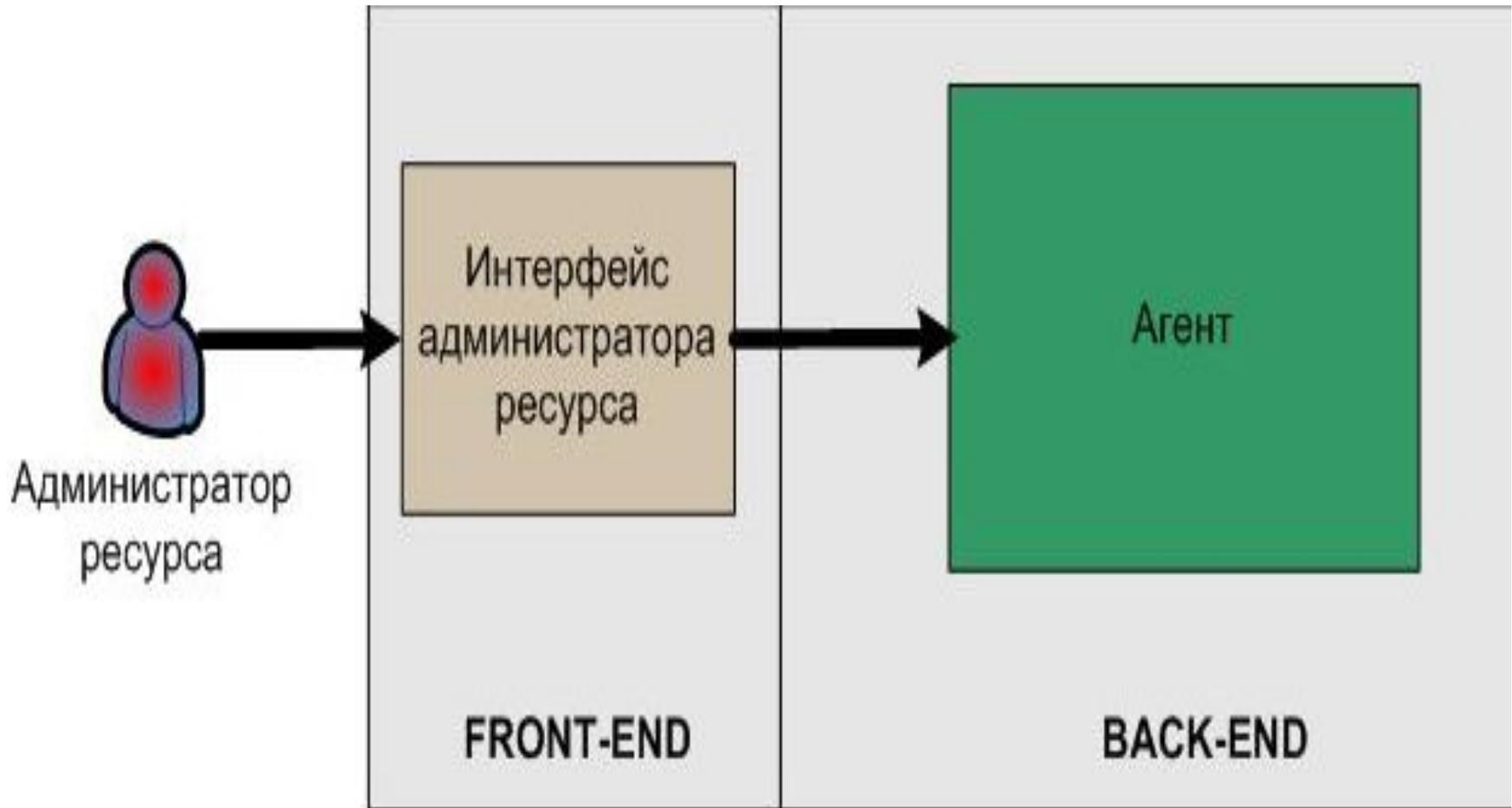
# Сервер депозитария



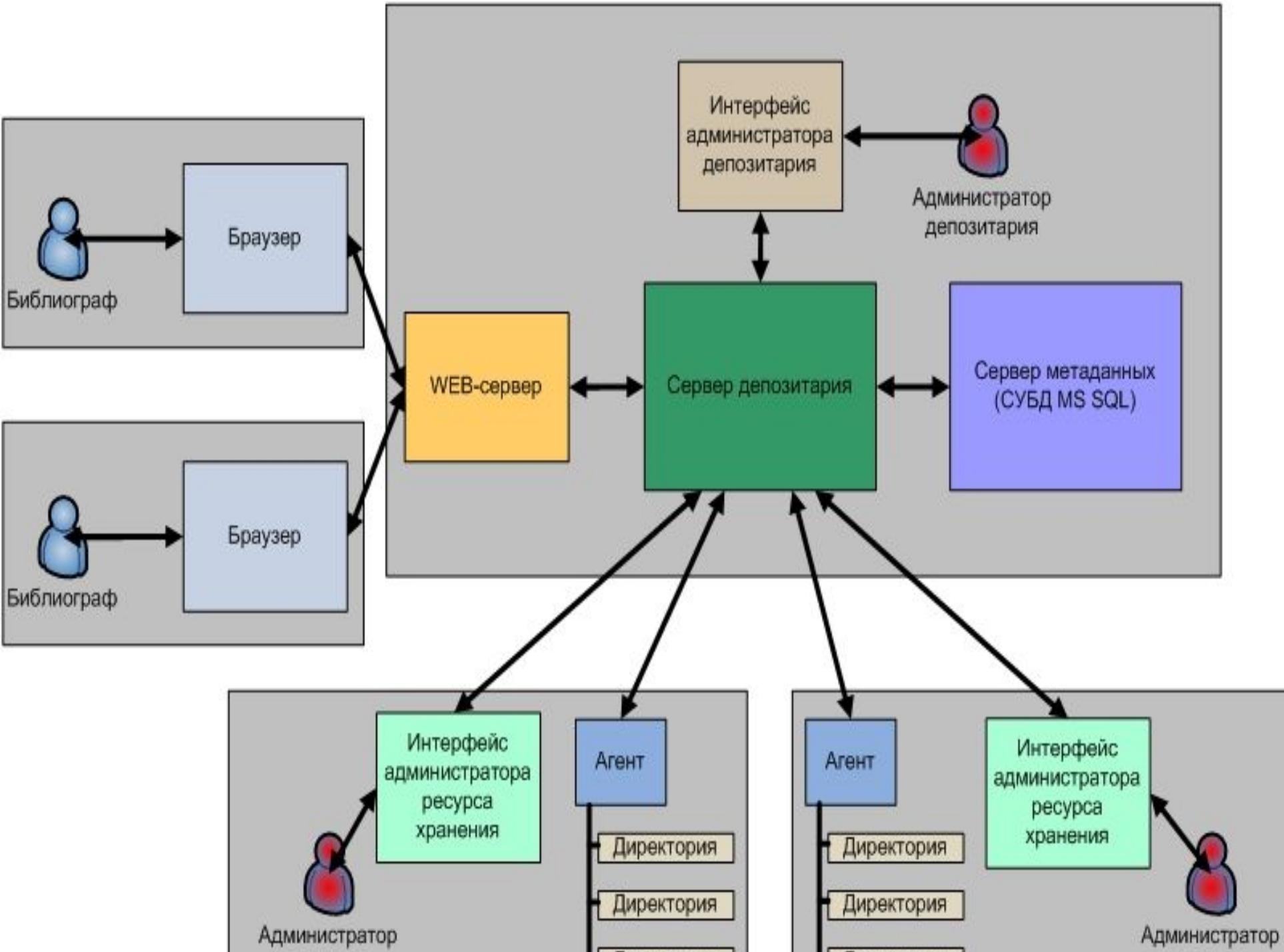
# Web-сервер

- Web-сервер используется для предоставления интерфейса взаимодействия пользователей с депозитарием. Пользователи с помощью браузера могут авторизоваться, осуществлять поиск по метаданным, добавлять новые и редактировать имеющиеся объекты данных.

# Агенты



Агенты устанавливаются на узлы хранения, входящие в состав депозитария. Они исполняют роль посредников между физическими устройствами хранения и остальной частью депозитария.





## Загрузка данных

файлы для загрузки

Файл (42)	Размер (4,169KB)
C:\RGB-DOC\0082p.jpg	106,034 byte
C:\RGB-DOC\0042.jpg	78,938 byte
C:\RGB-DOC\0043.jpg	81,877 byte
C:\RGB-DOC\0044.jpg	111,113 byte
C:\RGB-DOC\0045.jpg	79,686 byte
C:\RGB-DOC\0046.jpg	119,351 byte
C:\RGB-DOC\0046p.jpg	125,413 byte
C:\RGB-DOC\0047.jpg	126,679 byte
C:\RGB-DOC\0048.jpg	121,000 byte

Выбрать файлы

Выбрать папку

Удалить

Удалить все

объединить в многофайловый объект

расширение файлов описания MARC

.mrc

использовать описание MARC в качестве  
основного

Далее >>



## Загрузка данных

метаданные для загрузки

Коллекция \*

nat

Автор

Название

Аннотация

Добавить поле описания

Очистить

Загрузить

# Администрирование депозитария

File Edit View Help



- Depository administration
  - Depository management
  - Metadata
  - Data management
  - Resources**
- Users
  - Administrators
  - Librarians
  - Readers
    - Groups
    - Permissions
- Backup
- Audit
- Services
- Settings
- Logs
  - Local logs
  - Logs in depository
- Depository resource administration
- Depository librarian interface

Location	Name	Status	Owner	Shared folder	Quota	Free space
depo01(depo01:5625)		Enabled				
└─ raid		Enabled	depo01	F:\SP1\	Unlimited	100.00% (17179869183.59 T
test2(test2:5625)		Enabled				
└─ Storage		Enabled	test2	D:\Storage\	Unlimited	73.63% (44.08 Gb)
test-x,5625(test-x:5625)		Disabled				

Что дальше

# ССЫЛКИ

- Европейская конференция "The Challenge: Long-Term Preservation -Strategies and Practices of European Partnerships" on April, 2007 in Frankfurt/Main  
<http://www.langzeitarchivierung.de/eu2007>
- Сайт компании Шофт <http://www.shoft.net/>
- Проект библиотеки конгресса США «Digital Preservation»  
<http://www.digitalpreservation.gov/>