

Тема «Электронные (цифровые) библиотеки» в учебных курсах «Проектирование библиотечно-информационных систем» и «Программно-техническое обеспечение библиотечно-информационных систем» в Санкт-Петербургском государственном университете культуры и искусств

**А. Ц. Масевич**, ст. преподаватель каф.  
информационного менеджмента СПбГУКИ

**Л. А. Ходоровский**, к.т.н. доц. каф.  
информатики и математики СПбГУКИ

## Курсы «Проектирование АБИС» и «Программно-техническое обеспечение АБИС»

- Об курса преподаются по образовательным стандартам «Библиотечно-информационная деятельность»
- В настоящее время - 071201 Библиотечно-информационная деятельность 2003 года - специалисты
- Через год после завершения перехода на двухступенчатую систему по стандарту 071900, разработанному для двух ступеней обучения – бакалавры и магистры.

## Курсы «Проектирование АБИС» и «Программно-техническое обеспечение АБИС»

- Содержательно оба курса тесно связаны между собой
- При преподавании курса «Проектирование АБИС» использован опыт кафедры информатики и информационных технологий (в наст. время кафедра информатики и математики), где в течение многих лет преподавался курс «Проектирование информационных систем»

В настоящем сообщении мы представим следующие темы этих учебных курсов

- «Создание функциональной модели и модели потоков данных технологической линии оцифровки фондов библиотек» (курс проектирование АБИС – практическое занятие)
- «Программное обеспечение депозитария долговременного хранения цифровых ресурсов (курс программно-техническое обеспечение АБИС -семинар)»
- Программные комплексы для цифровых библиотек (курс программно-техническое обеспечение АБИС -семинар)»

Создание функциональной модели и модели потоков данных технологической линии оцифровка фондов библиотек

**ПРАКТИЧЕСКОЕ ЗАНЯТИЕ  
КУРС «ПРОЕКТИРОВАНИЕ АБИС»**

# Пример масштабного проекта оцифровки фонда Национальной библиотеки Франции: электронная библиотека Gallica

([http://www.bnf.fr/fr/collections\\_et\\_services/bibliotheques\\_numeriques\\_gallica.html](http://www.bnf.fr/fr/collections_et_services/bibliotheques_numeriques_gallica.html))

The image shows a screenshot of the Gallica website. The top left features a large graphic with the text "100 000 ouvrages numérisés" (100,000 digitized works) and the BnF logo. Below this, a navigation menu lists various services: "documentaire", "Périodiques et bases en ligne", "Banque d'images", "Reproduction des documents", "Services aux lecteurs", "Portails et guides thématiques", and "Les Signets de la BnF". The main content area includes a section titled "Gallica : référence numériques so" (Gallica: digital references so) with a sub-heading "Europeana". The text describes Gallica's mission since 2008 and mentions the Europeana project. A video player is overlaid on the bottom right, displaying the text "Mise en ligne de 2 500 documents par semaine" (2,500 documents online per week). The video player interface includes a progress bar and a volume icon.

# Проектирование технологической линии «Оцифровка фондов библиотеки»

- Для раскрытия темы была использована статья директора Норвежской национальной библиотеки Вигдис Моэ Скарстейн, которая содержит детальное описание технологии оцифровки фондов национальной библиотеки Норвегии

**Скарстейн Вигдис Моэ** Статья национальной цифровой библиотекой: стратегия оцифровки для национальной библиотеки Норвегии / пер. с англ. А.Ц.Масевича // Национальные библиотеки в XXI веке: сб. статей. – СПб : Изд-во РНБ, 2012. – С.104 – 122



# Основные технологические операции производственной линии оцифровки по материалам Королевской библиотеки Норвегии и национальной библиотеки Франции

1. Отбор для оцифровки и установка очередности отобранных книг;
2. Извлечение книг из фондов;
3. Транспортировка;
4. Извлечение метаданных (библиографических описаний и др.);
5. Сканирование печатных материалов;
6. Оптическое распознавание текста и его структурный анализ;
7. Генерация объектов хранения, контроль качества;
8. Ввод их в хранилище для долговременного хранения;
9. Ввод в каталог указания о наличии оцифрованной копии;
10. Индексирование распознанного текста и метаданных для поисковой системы.



# CASE - технологии

- CASE – Computer Aided Software/System Engineering (В русскоязычной литературе нет соответствующего общепринятого термина).

*Перевод: Анализ, разработка и проектирование ( = Engineering) программных продуктов / сложных систем ( = Software/System) с использованием компьютерных технологий (Computer Aided).*

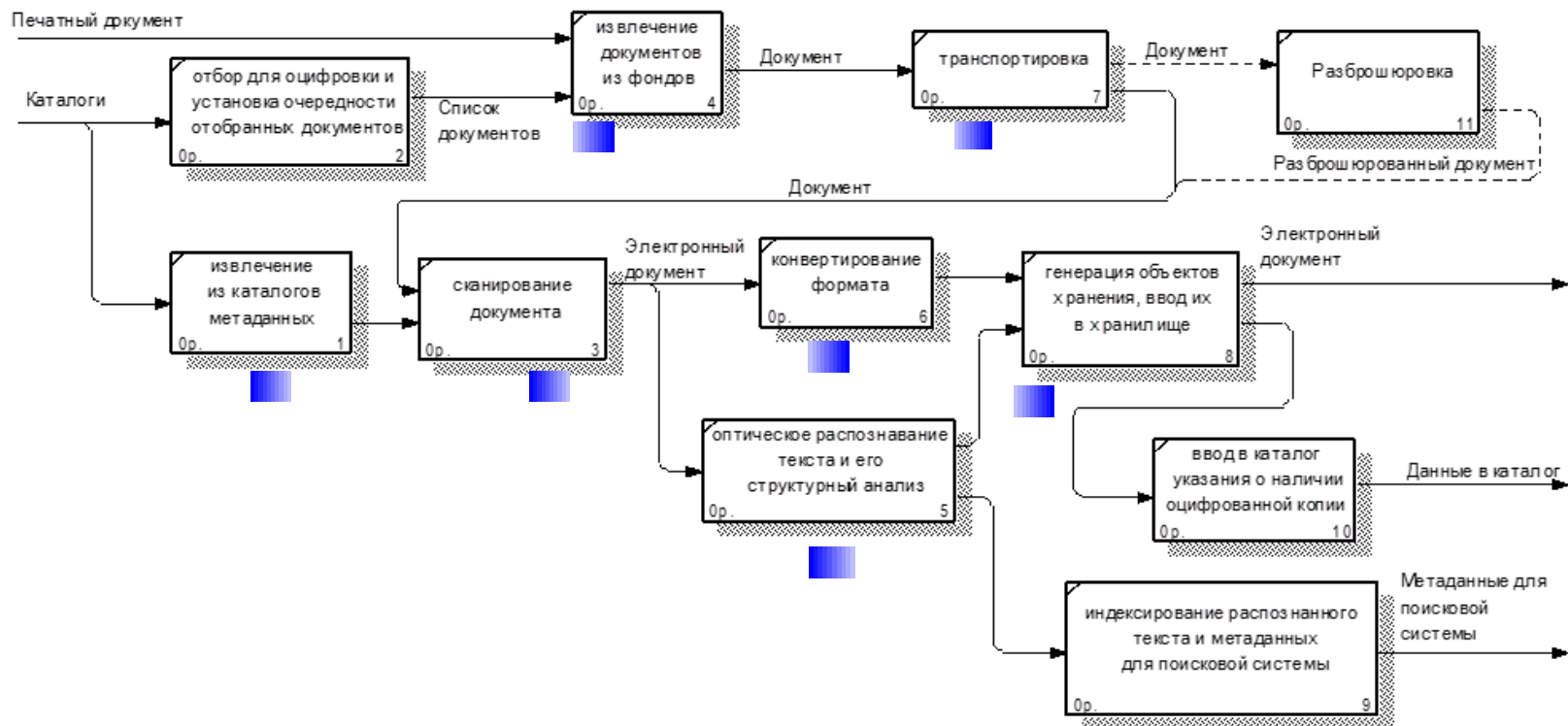
Определение: Набор инструментов и методов для проектирования программного обеспечения, который помогает обеспечить высокое качество программ, отсутствие ошибок и простоту в обслуживании программных продуктов.

# CA ERwin Process Modeler 7.3 (ранее BPwin)

- CA ERwin Process Modeler (ранее BPwin) - инструмент для моделирования, анализа, документирования и оптимизации технологических и бизнес процессов.
- CA ERwin Process Modeler обычно используется для графического представления процессов.
- Графически представленная схема выполнения работ, обмена информацией, документооборота визуализирует модель процесса.
- Графическое изложение этой информации позволяет перевести задачи управления организацией из области сложного ремесла в сферу инженерных технологий.

# Модульно-функциональная схема технологической линии оцифровки фонда библиотеки

USED AT:	AUTHOR: Ходоровский Л.А., Масевич А.Ц.	DATE: 05.02.2013	WORKING	READER	DATE	CONTEXT:
	PROJECT: Технологическая линия оцифровки документов	REV: 07.02.2013	DRAFT			
			RECOMMENDED			
	NOTES: 1 2 3 4 5 6 7 8 9 10		PUBLICATION			A-0

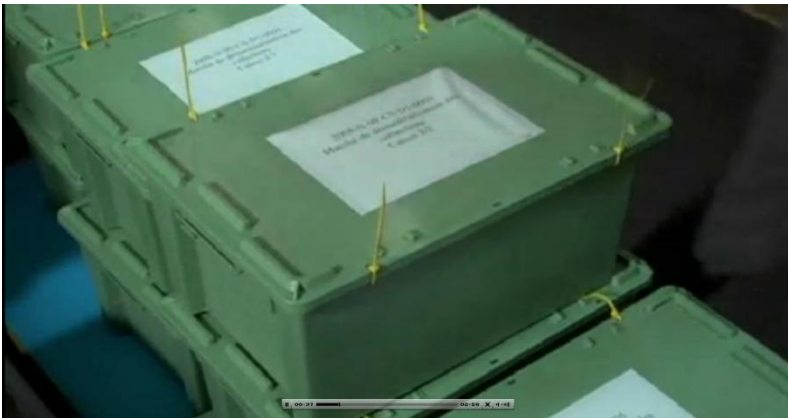


NO DE:	TIT LE:	NUMBER :
A0	Технологическая линия оцифровки документов	

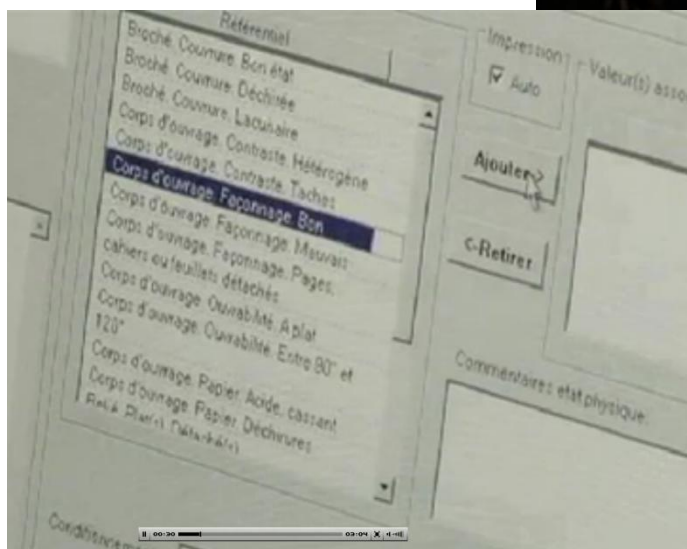
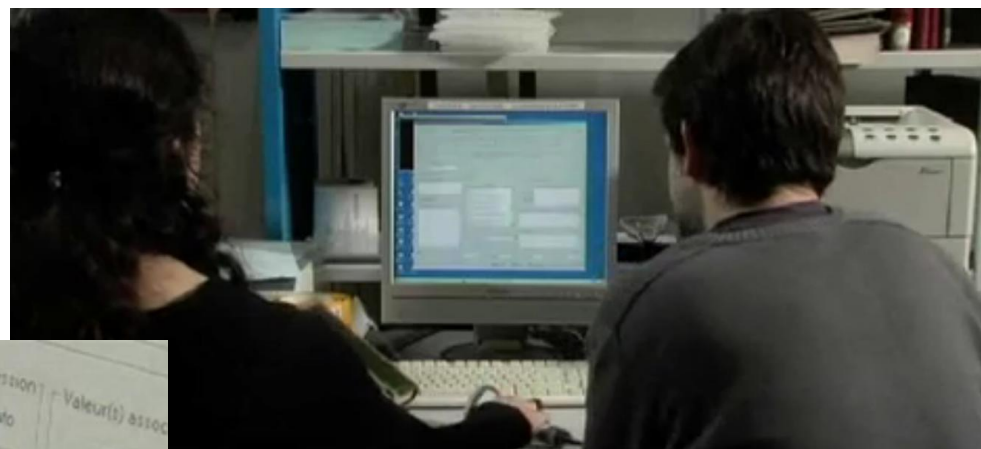
# Отбор для оцифровки и установка очередности отображенных книг



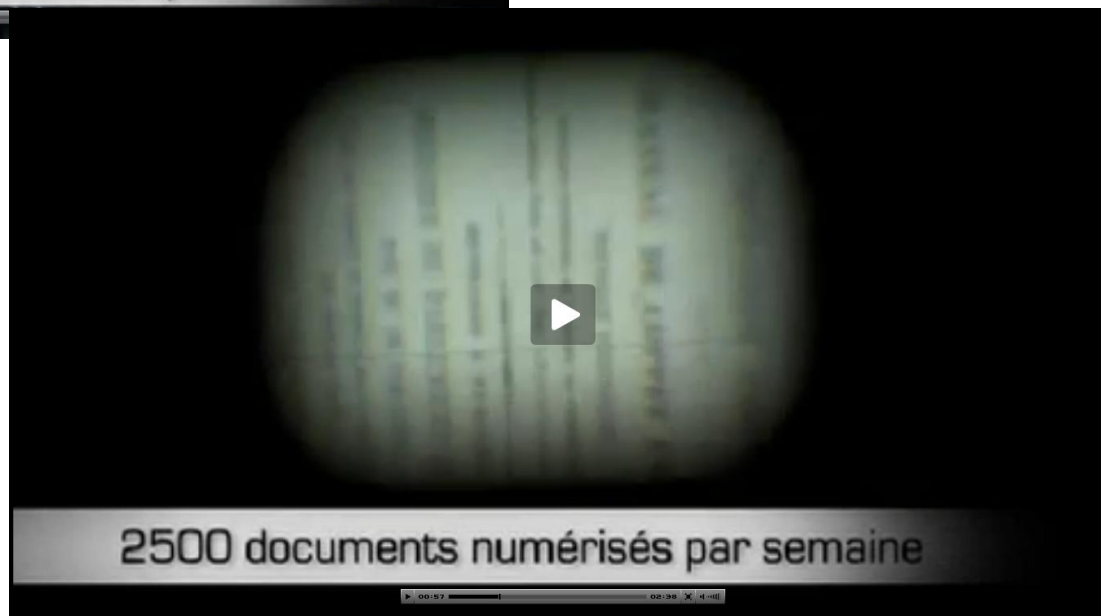
# Транспортировка книг



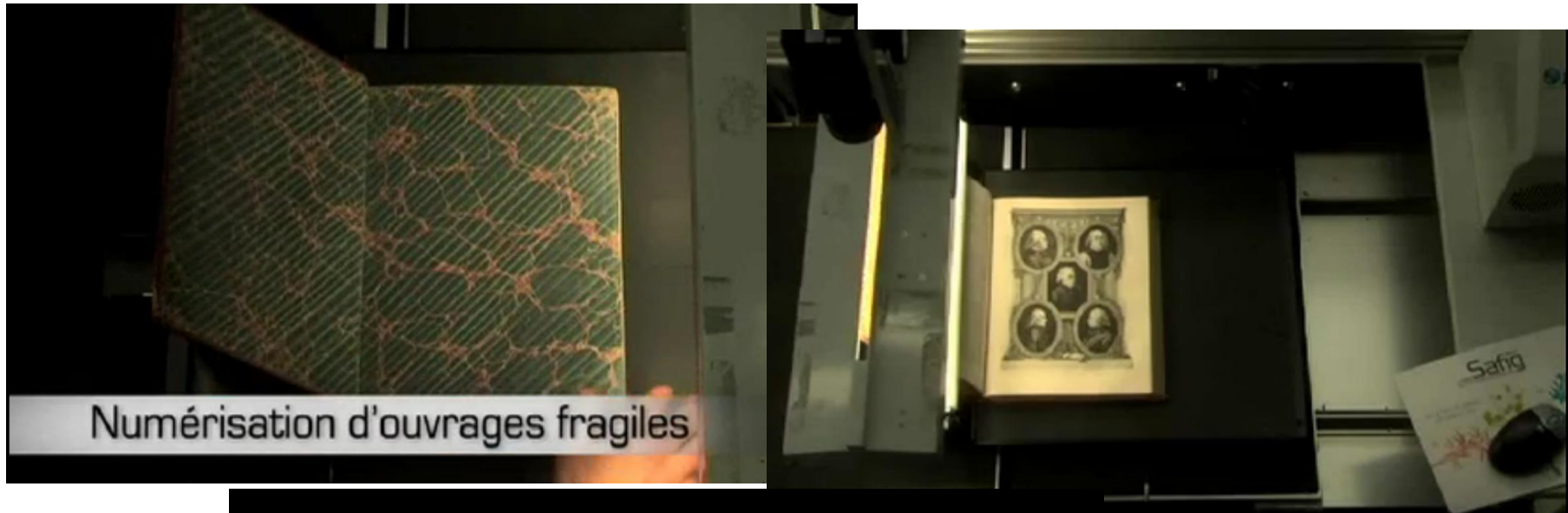
# Извлечение метаданных (библиографических описаний и др.)



# Сканирование книг (в автоматическом режиме )

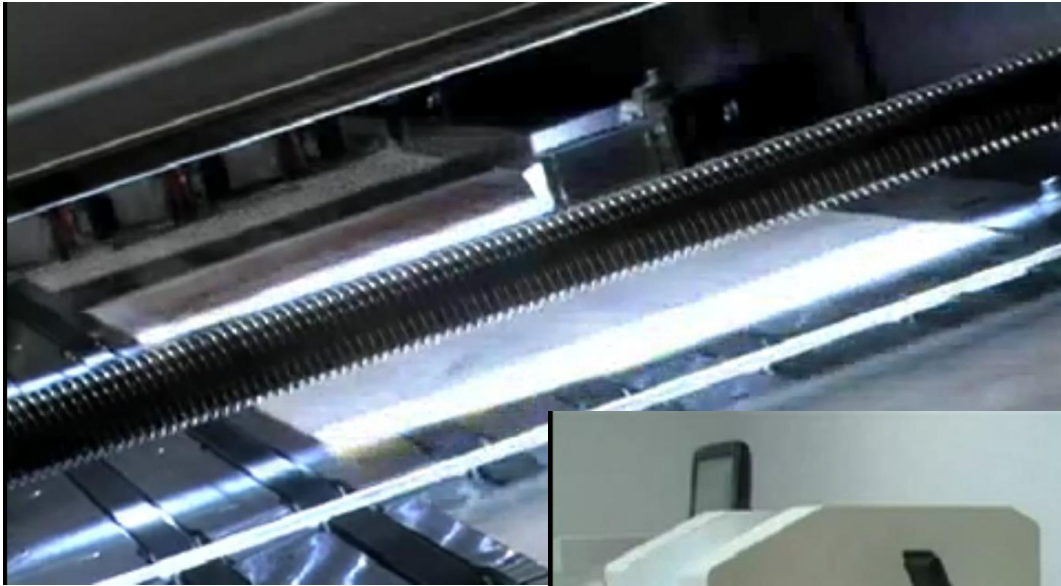


# Сканирование книг (в ручном режиме )

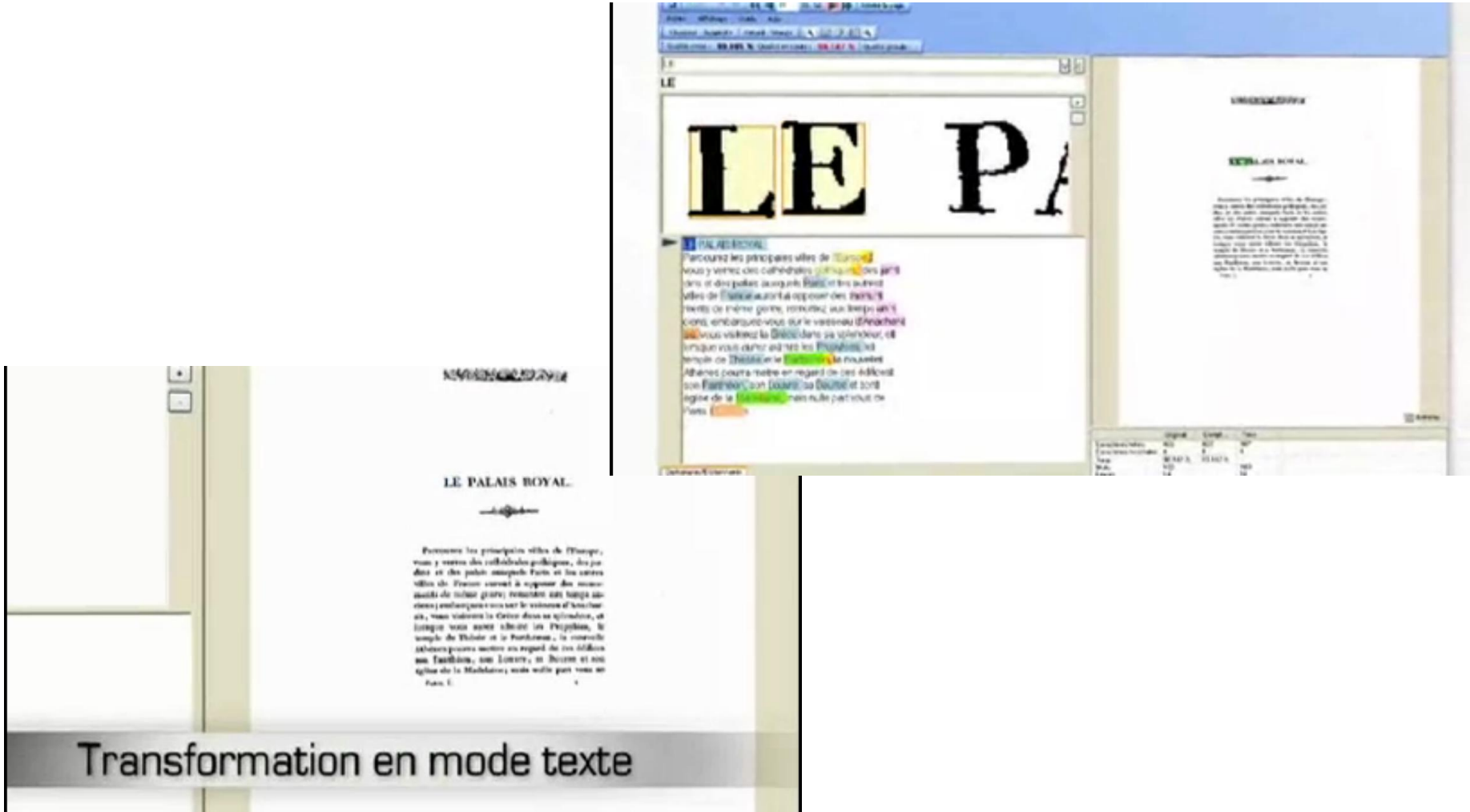




# Сканирование листового материала



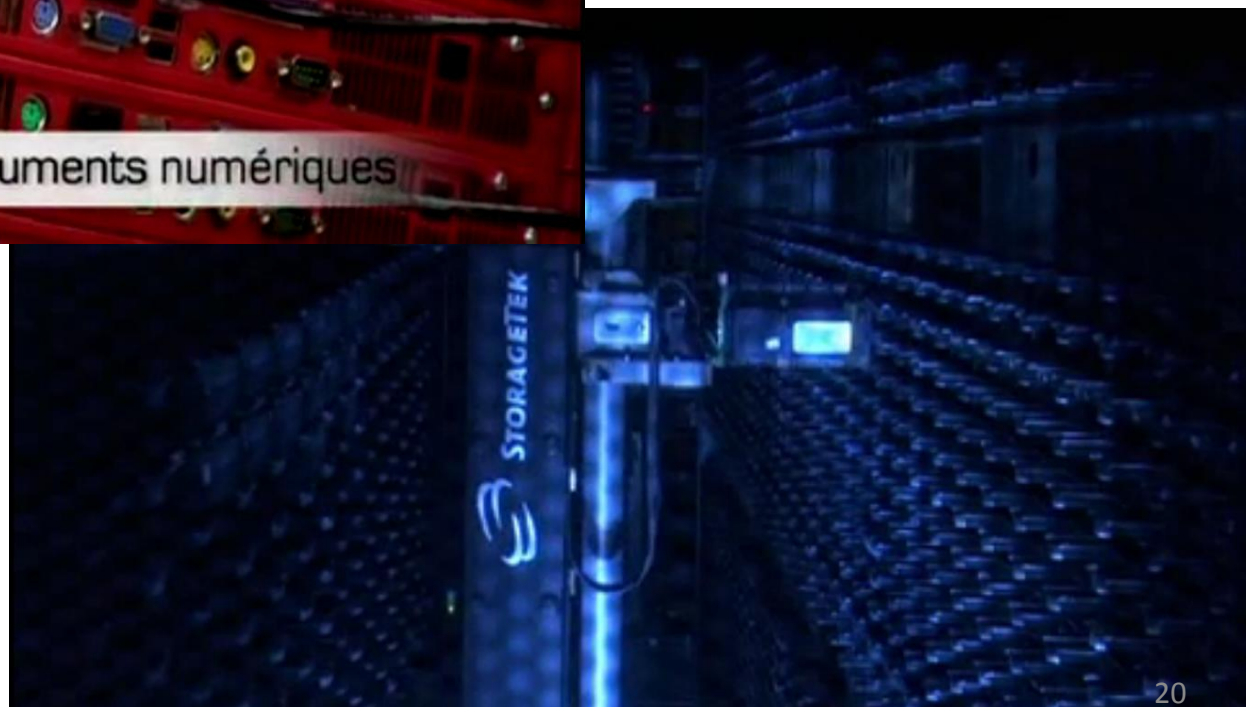
# Оптическое распознавание текста и его структурный анализ



# Контроль качества ■

```
<editeur>[s.n.?]</editeur>
<dateEdition>1912/10/26</dateEdition>
HYPERLINK "file:///Z:\\TestsClients\\TEST1\\03DocView\\>
<type>Année</type>
<valeur>73</valeur>
</tomaison>
HYPERLINK "file:///Z:\\TestsClients\\TEST1\\03DocView\\>
<type>Fascicule</type>
<valeur>250</valeur>
</tomaison>
<nombrePages>12</nombrePages>
.../12148/cb32
Contrôle qualité
<reference type="CODEBARREPROVENANCE">3751100248871
```

# Ввод материала в хранилище для долговременного хранения







Программное обеспечение  
депозитария  
долговременного хранения  
цифровых ресурсов

**СЕМИНАР**  
**КУРС «ПРОГРАММНО-ТЕХНИЧЕСКОЕ**  
**ОБЕСПЕЧЕНИЕ АБИС»**

В чем состоит проблема сохранности цифровых материалов? (из доклада Барбары Сирмен на конференции CENL 2006 г. , Санкт-Петербург)

## □ Элементарные проекты с перспективой □ Европейские проекты по долгосрочному хранению

- *DPE (Digital Preservation Europe)*
  - Проект, цель которого – поддержка сотрудничества в области сохранности документов, учет и координация разработок в этой области
- *Caspar (Cultural, Artistic, and Scientific Knowledge for Preservation, Access and Retrieval)*
  - Проект касается в основном предметов искусства, науки и культуры
- *Planets, Digital Preservation Research and Technology*
  - Организация, планирование, разработка технологий хранения цифровых объектов

# Система длительного хранения цифровых материалов

- Эталонная модель Open Archival Information System (OAIS)
- Разработка осуществлена консультативным комитетом космических информационных систем (CCSDS)
- Издана как рекомендация CCSDS (Blue Book) 650.0-B-1 (January 2002)
- Периодически пересматривается: ISO 14721:2003
- <http://public.ccsds.org/publications/archive/650x0b1.pdf>

Определение OAIS:  
*Открытая стандартная  
схема архива длительного  
хранения цифровых данных*



# SPAR (SYSTEME DE PRESERVATION ET D'ARCHIVAGE REPARTI DE LA BIBLIOTHEQUE NATIONALE DE FRANCE)

- ❖ В качестве примера реализации стандарта OAIS используется схема Системы хранения и распределенного архива национальной библиотеки Франции (SPAR).
- ❖ [http://www.bnf.fr/fr/professionnels/conserver\\_spar/s.conserver\\_spar\\_presentation.html](http://www.bnf.fr/fr/professionnels/conserver_spar/s.conserver_spar_presentation.html)

The screenshot shows a web browser window displaying the BnF website. The page title is "Bibliothèque nationale de France". The main content area is titled "Conserver : le projet SPAR et l'archivage numérique". The text on the page describes the system's purpose and its evolution, mentioning the acquisition of IPRES 2012 and the implementation of the OAIS standard. The page also includes a navigation menu, a search bar, and a sidebar with various links and social media icons.

→

**General Catalogue** | Historical newspapers | Periodicals 1850-1940 | Books 1700-1870 | Memory of the Netherlands

**Zoeken**

Go to: [Renew loans and user data](#)  
Try now: [WorldCat Local](#)

🕒 Catalogues & Digital Collections: access to catalogues, e-journals & databases

⚠️ **Please note** This Saturday the General Catalogue temporarily unavailable.

● ● ● ● ●  
**Treasures of the KB**  
🕒 Illuminated Medieval manuscripts

**News**

- New KB-fellow: Marc van Oostendorp
- Dutch-American Newspapers online available
- 🕒 News archive



**The Library**

- Address and opening hours
- Ask a librarian
- How to use the KB
- Request a KB pass
- Rules for borrowing books and publications
- Staff Profiles & AOP

**Exhibitions & collections**

- Web exhibitions
- Digitized books
- Special collections
- KB fellowship
- Manuscripts Exhibition
- Related KB projects

**About the KB**

- Organization and policy
- Vacancies
- Expertise
- Research
- For publishers
- For the press

**DIAS – Digital Information Archiving System**  
**LTP – Long Time Preservation**

## Семинар по теме программное обеспечение систем длительного хранения цифровых ресурсов

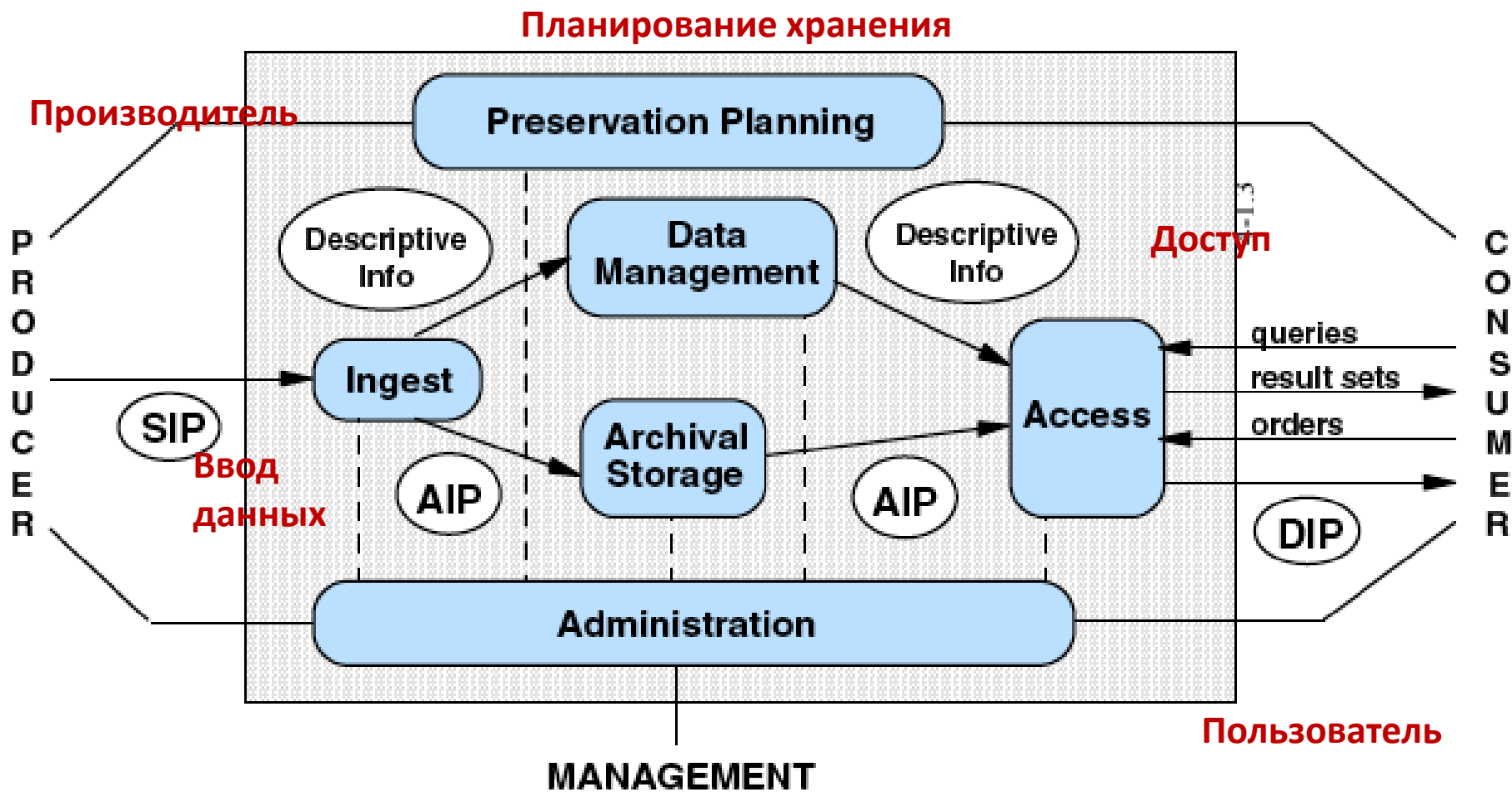
При проведении занятия используется схема системы, с сайта BNF. На ней обозначены названия стандартных бесплатных программных пакетов, которые могут использоваться в модулях системы

Студенты получают задания самостоятельно найти материалы по этим пакетам и составить два описания – одно детальное, другое предельно краткое.

Подробное описания составляется в формате MS WORD или в формате презентации PowerPoint (по выбору студента)

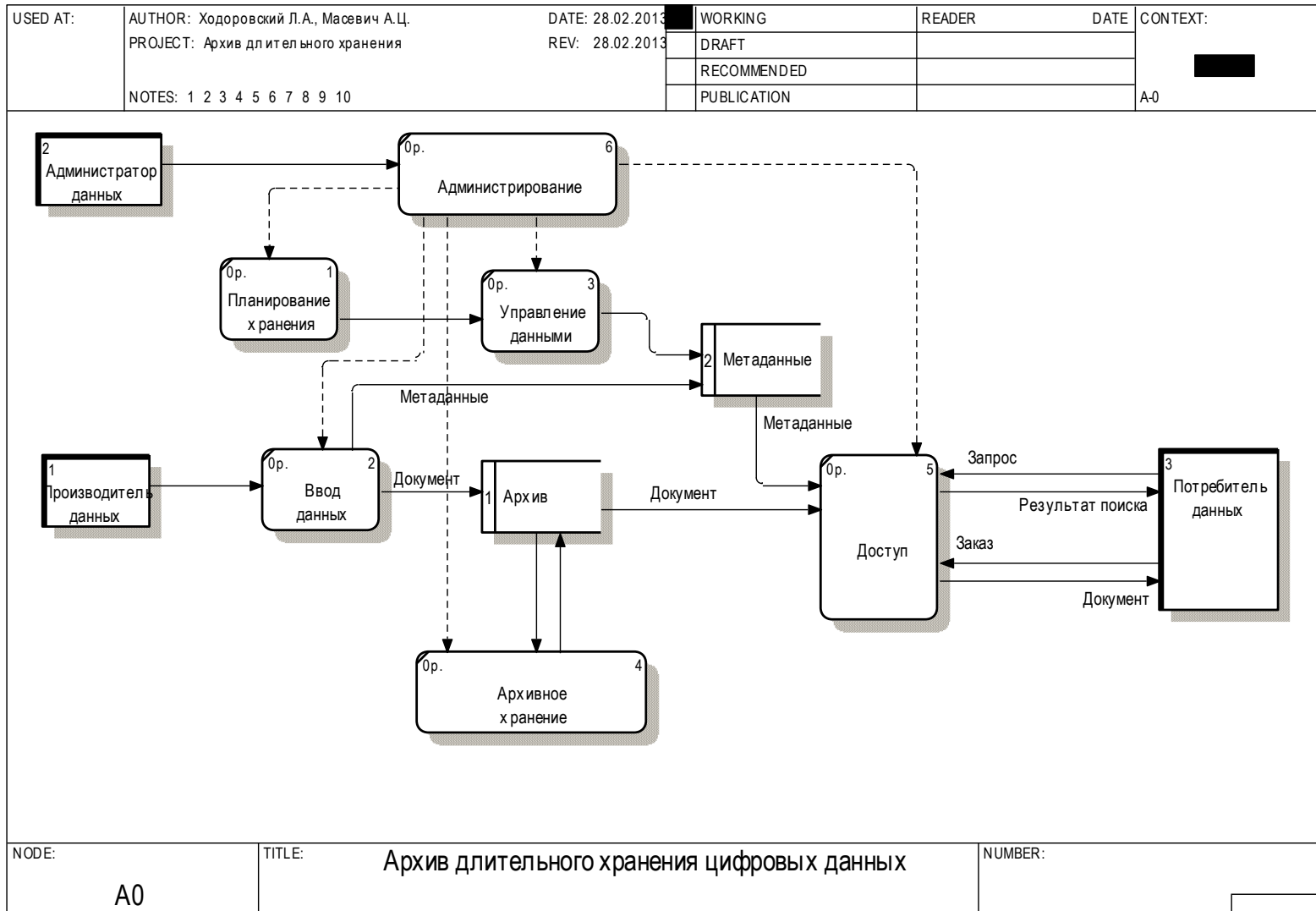
Краткие описания средствами PowerPoint присоединяются к схеме, взятой из национальной библиотеки Франции.

# Схема OAIS



Управление данными, архивное хранение, администрирование

# Схема потоков данных для депозитария долговременного хранения, построенная с помощью СА Erwin Process Modeler (Bpwin)



NODE:

A0

TITLE:

Архив длительного хранения цифровых данных

NUMBER:

# Описание стандартных пакетов программ в

## системе длительного хранения

**QUARTZ Scheduler** представляет собой систему планирования заданий, которые могут быть интегрированы практически с любым программным обеспечением или использоваться самостоятельно. Quartz Scheduler включает в себя многие функции корпоративного класса, такие как транзакции JTA и кластеризации.  
<http://www.quartz-scheduler.org/>

**STRUTS** был создан для того, чтобы чётко разделить модель (бизнес-логику), представление (HTML-страницы) и контроллер (отвечающий за передачу данных от модели к представлению и обратно).  
<http://struts.apache.org/>

**Open LDAP** (англ. Lightweight Directory Access Protocol — «облегченный протокол доступа к каталогам») — протокол прикладного уровня для доступа к службе каталогов. Относительно простой протокол, использующий TCP/IP и позволяющий выполнять операции авторизации (bind), поиска (search) и сравнения (comparasion), а также операции добавления, изменения или удаления записей.  
<http://www.openldap.org/>

**JHOVE** - совместный проект JSTOR и Гарвардского Университета, выполняет формат-специфическую идентификацию, проверяет формат и описывает цифровые объекты.

**QUARTZ Scheduler** представляет собой систему автоматизированного составления графиков исполнения заданий. Она может быть интегрирована практически с любым программным обеспечением или использоваться самостоятельно.  
<http://www.quartz-scheduler.org/>

**Virtuoso Universal Server** является программным средством класса мульти-модель данных. Оно обеспечивает эффективное решение для управления данными, доступа и интеграции данных.

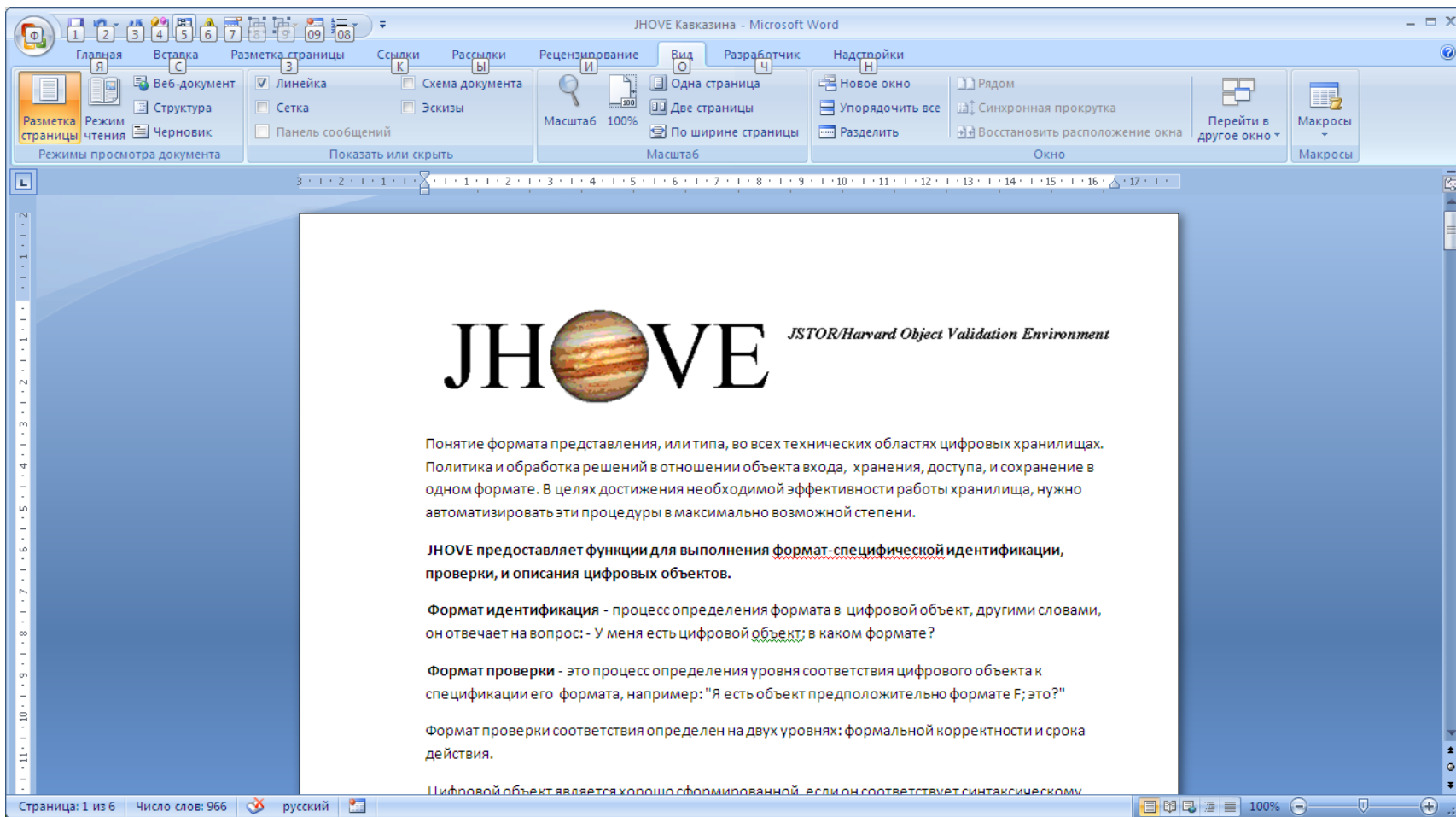
**Jersey** — язык программирования для создания интерфейса для доступа к документам.  
<http://jersey.java.net/>

**iRODS (Rule-Oriented Data-management System)** — интегрированная система управления данными с открытым исходным кодом. RODS помогает исследователям, архивистам и другим сформировать (организовать, защитить и сохранить) большие объёмы файлов на компьютере — от сравнительно небольшого числа до сотен миллионов файлов, содержащих петабайты данных.  
[https://www.iods.org/index.php/What\\_is\\_iRODS%3F](https://www.iods.org/index.php/What_is_iRODS%3F)

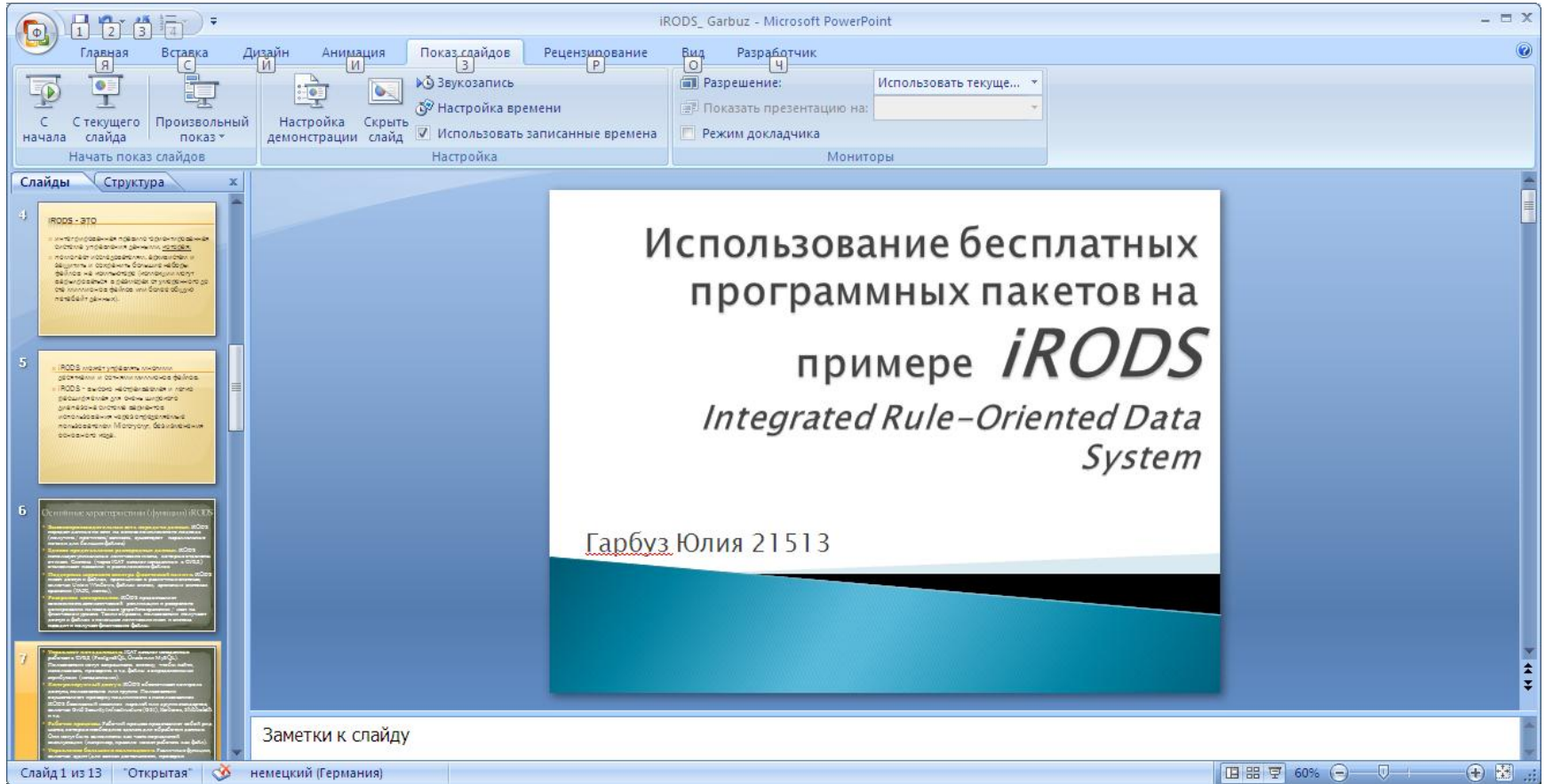
**SCHEMATRON** - язык и инструмент для получения данных о моделях (образцах), найденных в XML -документах

**OpenSolaris** — операционная система с открытым исходным кодом, созданная корпорацией Sun Microsystems на базе Solaris. Также термин OpenSolaris может использоваться по отношению к открытой кодовой базе Solaris и сообществу, которое её разрабатывает.  
<http://hub.opensolaris.org/bin/view/Main/>

# Пример детального описания, выполненного студенткой в Word



# Пример детального описания, выполненного студенткой в PowerPoint







# Программные комплексы для цифровых библиотек

**СЕМИНАР  
КУРС «ПРОГРАММНО-ТЕХНИЧЕСКОЕ  
ОБЕСПЕЧЕНИЕ АБИС»**

По состоянию на октябрь 2013 на запрос «digital library software» (программное обеспечение) ИПС Google представляет 217 млн (!) ссылок.

В истории программного обеспечения АБИС можно выделить три этапа: индивидуальные системы, т.е. разработанные для конкретной, как правило, национальной библиотеки (середина 1960-х – конец 1980-х), создание и распространение программных комплексов (1990-е годы до середины 2000-х), создание отдельных модулей (приблизительно после 2010).

Настоящий раздел курса «Программно-техническое обеспечение АБИС» дается в виде традиционного семинарского занятия с докладами студентов, которому предшествует инструктивная лекция.

Студентам предлагается перечень программных комплексов, из которого они на свое усмотрение выбирают программный комплекс для описания, и примерный список вопросов, которые они должны отразить в своих докладах.

# Слайд из презентации студента, посвященной программному комплексу Greenstone (общая характеристика комплекса)

greenstone digital  
library software



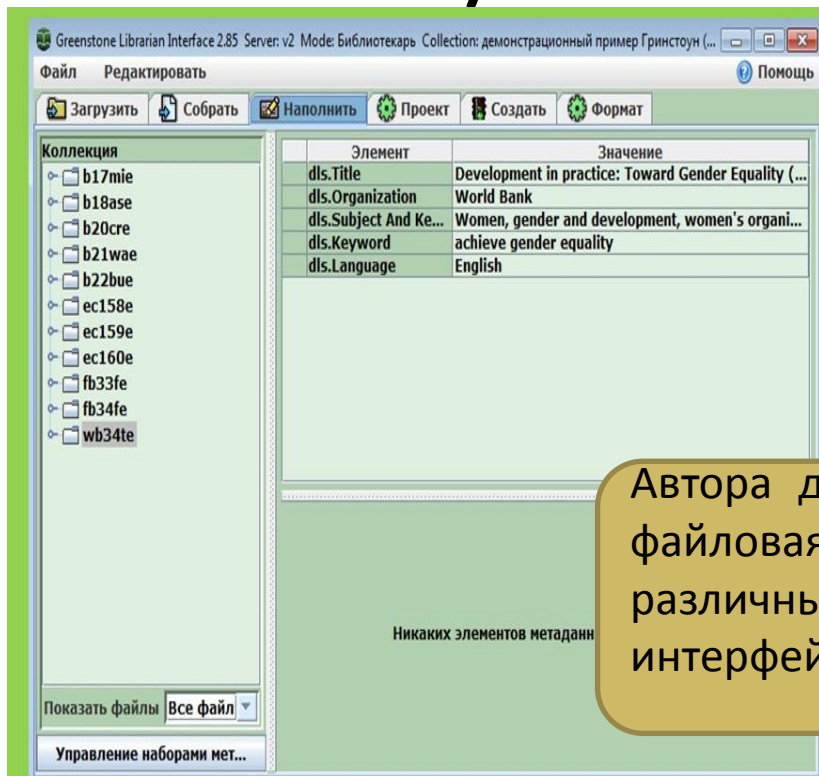
Greenstone – это бесплатный программный комплекс для создания электронных библиотек.

Предназначен для использования главным образом в образовательных (библиотеках, университетах, школах) и других общественных учреждениях.

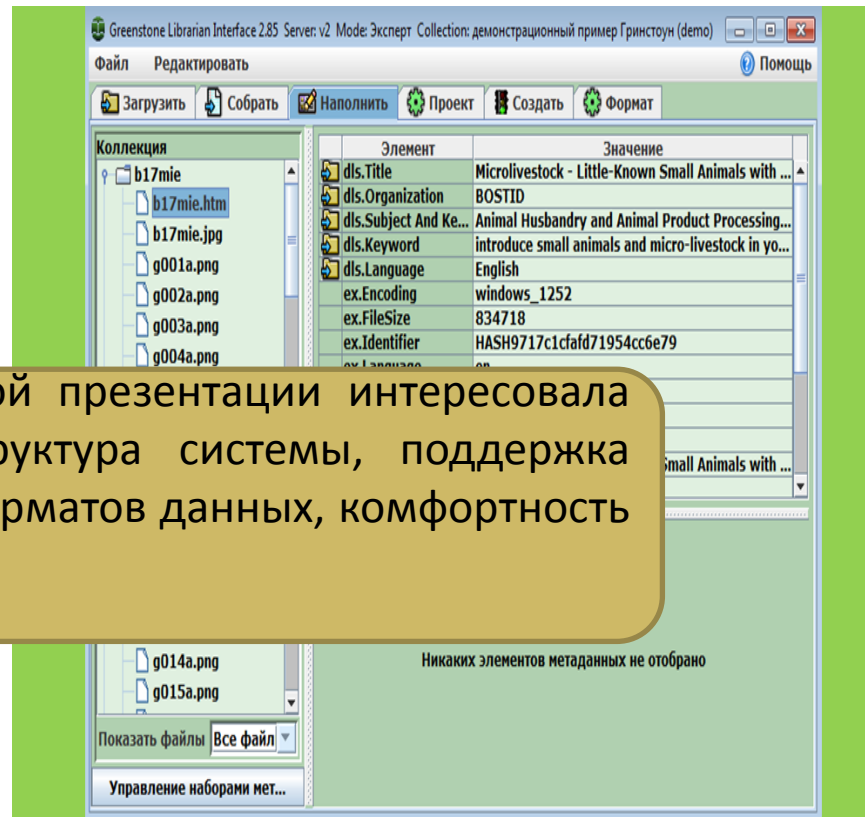
Разработчик – университет Вайкато (Гамильтон, Новая Зеландия) в сотрудничестве с ЮНЕСКО.

Начало Greenstone было положено в 1995 году с созданием в университете группы разработчиков New Zealand Digital Library Project.

# Слайды из презентации студента на тему комплекс Greenstone



Работает с четырьмя форматами метаданных, в том числе с Dublin Core.



Greenstone воспринимает все основные форматы текстовых (TXT, DOC, RTF, HTML и т.д.), графических (GIF, PNG, JPG и т.д.), звуковых (MP3, OGG и т.д.) файлов, а также PDF-документы и ZIP-архивы.

Автора данной презентации интересовала файловая структура системы, поддержка различных форматов данных, комфортность интерфейса.



**LOGOS**  
Bible Software

ТЫЧИНСКАЯ НАТАЛЬЯ  
ГР. 21413



# РЕЗУЛЬТАТ ПОИСКА ПРОГРАММЫ НА NGRAM VIEWER

## ОСНОВНАЯ ИНФОРМАЦИЯ

- ✓ **Разработчик:** Logos Research Systems
- ✓ **Выпуск рабочей версии:** 5.0 / Ноябрь 2012
- ✓ **Операционная система:** MS Windows Vista/7/8 / Apple Mac OS X / Apple iPhone / Google Android Linux / Amazon Kindle Fire / Android
- ✓ **Тип системы:** e-books, digital library, Bible study
- ✓ **Веб-сайт:** [www.logos.com](http://www.logos.com)

## ФУНКЦИИ

- ✓ Logos Bible Software это приложение для электронных библиотек, разработанное специально для изучения Библии.
- ✓ Узкая тематика позволяет программе реализовать широкий спектр возможностей.
- ✓ В дополнение к базовой функциональности электронных книг, она включает в себя:
  - ✓ ссылки, соединяющие ресурсы;
  - ✓ функцию заметок;
  - ✓ функцию лингвистического анализа для изучения Библии и переводе и на оригинальных языках.
- ✓ Logos Bible Software обеспечивает инструменты и ресурсы для изучения Библии на самых разнообразных платформах на Mac, ПК и мобильных устройствах.

Google books Ngram

Graph these case-sensitive comma-separated

between 1980 and 2008 from the

Search lots of books

Logos Bible Software



Search in Google Books:

1980 - 1993

1994 - 2003

2004 - 2007

2007

2008

Logos Bible Software

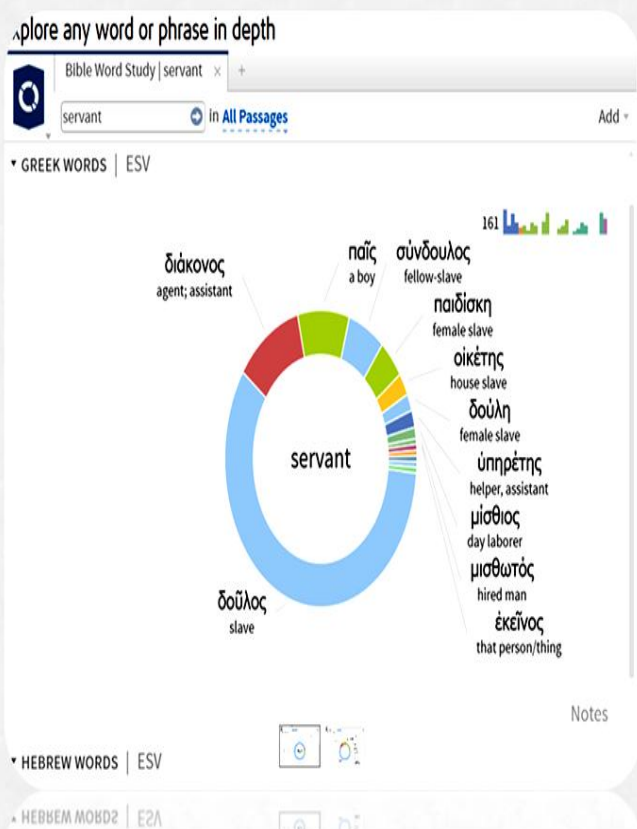
English

# Слайды из презентации студентки по программному комплексу LOGOS

В данной презентации сделан акцент на узкую направленность разработки и связанный с этим набор функций, позволяющих осуществлять историко-лингвистический анализ текста.

ВОЗМОЖНОСТЬ ПОИСКА ОПРЕДЕЛЕНИЙ, СИНОНИМОВ, ПЕРЕВОДОВ И ПР.

СИСТЕМА ФАКТИЧЕСКИХ СПРАВОК ПО БИБЛЕЙСКИМ ПЕРСОНАЖАМ И СОБЫТИЯМ



Bible Facts | Barnabas

Barnabas

Barnabas & Paul in Cyprus (Ac 13:4-12)

Barnabas — a Levite and apostle who sought Paul's help as his associate in the mission to the Gentiles.

Barnabas sells a field Ac 4:36-37

The Jerusalem Council Ac 15:6-21

Barnabas and Paul bring assistance to Judea Ac 11:30, 12:25

Paul visits Jerusalem after his conversion Ac 9:26-29

Paul preaches in the synagogue at Pisdian Antioch Ac 13:14-43

ROLE

DICTIONARIES

RELATIVES

Mark Cousin

LEMMAS

Barnabas, Βαρναβᾶς Barnabas, Acts 4:36, 9:27, 11:19-30, 12:20-13:3, 7, 13-52

Joseph, Joses, Ἰωσήφ Ióséph, Acts 4:36

Paul

Barnabas

Mark

Reasoned with the Jews

Assisted Paul & Barnabas

Sergius Paulus

Elymas

Summoned Paul & Barnabas

False prophet

Barnabas & Paul





DSpace™

<http://www.dspace.org/>

DSpace™

## О DSpace

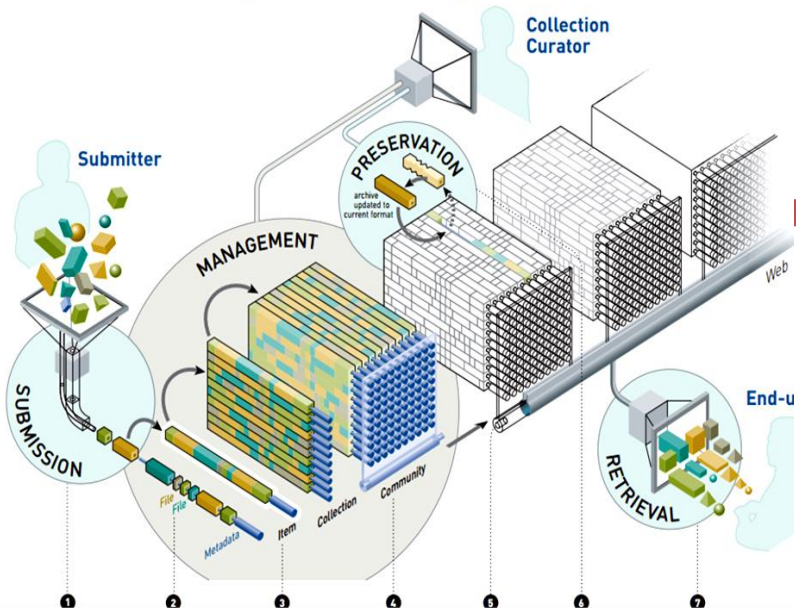
- Dspace – это программное обеспечение для академических, некоммерческих и коммерческих организаций, которое дает возможность строить открытые цифровые хранилища.
- Dspace бесплатен и прост в установке, полностью настраивается с целью соответствия потребностям любой организации
- Dspace обеспечивает сохранность и простой доступ ко всем типам цифрового контента, включая текст, изображения, видео, т.д.
- Постоянно растущее сообщество разработчиков стремятся постоянно расширять и совершенствовать программное обеспечения, добиваясь улучшения DSpace для пользователей.

ина Н. 21413гр

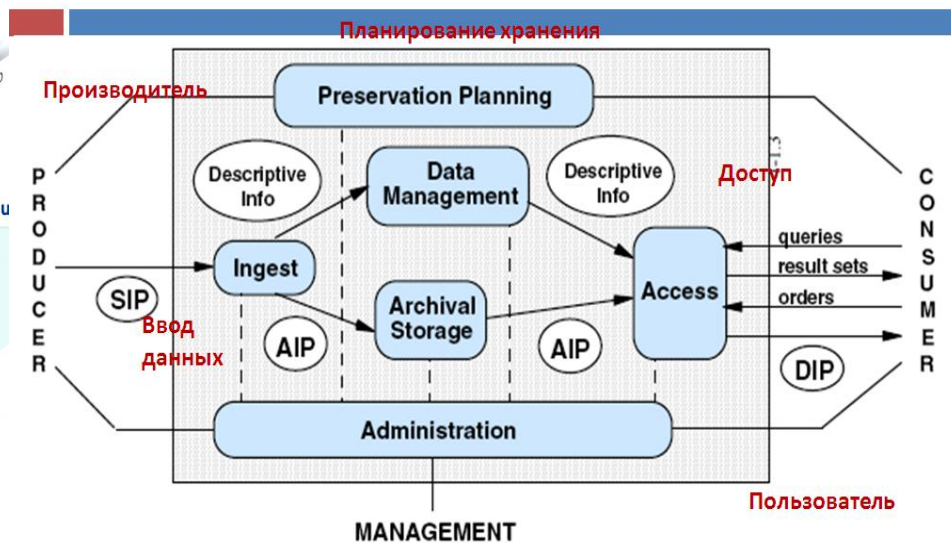
# Слайды из презентации студентки о комплексе DSpace

DSpace™

## Принцип работы



Автор презентации использовала полученное ранее представление о стандартной схеме хранилища цифровых ресурсов OAIS и сопоставила эту схему со схемой описываемого комплекса.



Управление данными, архивное хранение, администрирование

# Таким образом:

Выбор студентами материалов для докладов, то, какой стороне они предпочитают уделить больше внимание, показывает их интересы, видение вопроса.

Существует идея, поддержанная на кафедре информационного менеджмента, предложить студентам подготовить коллективный доклад на тему

«Программные комплексы цифровых библиотек» для ежегодной научной конференции студентов и аспирантов. Курс «Проектирование АБИС» завершается

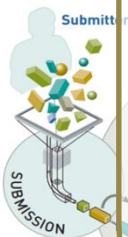
составлением учебного технического задания на проектирование цифровой электронной библиотеки, при выполнении этой работы студенты используют знания,

полученные в ходе изучения трех рассмотренных тем.

- ✓ Авторская организация CERN
- ✓ Разработка Server Software Consortium
- ✓ Постоянный выпуск 1.0.0 (до декабря 2009 года)
- ✓ Development language - Python

DSpace™

Принцип работы

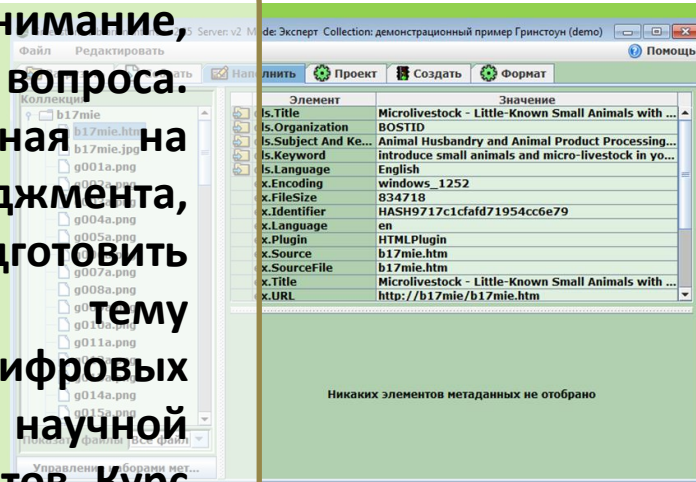


MANAGEMENT

RESERVATION

End-user

RETRIEVAL




Greenstone воспринимает все основные форматы текстовых (TXT, DOC, RTF, PDF), графических (GIF, PNG, JPG и т.д.), звуковых (MP3, OGG и т.д.) файлов, а также PDF-документы и ZIP-архивы.

LOGOS

ТЫЧИНСКАЯ НАТАЛЬЯ  
ГР. 2

# Заключение:

- Актуальность темы «Проектирование электронных библиотек» и необходимость включения её в программу подготовки специалиста по специальности «Библиотечно-информационная деятельность» очевидна.
- На кафедре информационного менеджмента СПбГУКИ преподавание этой темы осуществляется в курсах «Проектирование Автоматизированных библиотечно-информационных систем» и «Программно-техническое обеспечение информационных систем»
- Такие темы как оцифровка фондов библиотек, долговременное хранение цифровых ресурсов и программные комплексы для электронных библиотек сегодня наиболее практически актуальны, и потому при обучении им уделяется особое внимание
- Используется также опыт создания и поддержки собраний цифровых ресурсов трех крупных национальных библиотек, что позволяет студентам познакомиться с последними достижениями библиотечной науки и практики
- При преподавании курса используется опыт кафедры информатики и математики по преподаванию курса «Проектирование информационных систем» с использованием CASE-систем для проектирования информационных систем.
- В преподавании отдельных тем курса применен разработанный нами прием сочетания аудиторного обучения и самостоятельной работы студентов, что, как нам представляется, вызывает интерес и способствует усвоению материала.



**Благодарим за  
внимание!**