

Особенности научных электронных библиотек

М.Р. Когаловский
ИПР РАН

kogalov@cemi.rssi.ru

Что такое электронная библиотека?

- ЭБ - сложная наукоемкая система
- Широкий круг заинтересованных специалистов (ИТ, библиотечное дело, музейное дело, правоведение, образование, информационная безопасность, лингвистика, научные исследования во многих областях знаний)
- Разные точки зрения на назначение и функции ЭБ
- Спектр систем: от вырожденных до продвинутого типа

ЭБ - новый класс информационных систем

- Глобальный доступ к системе в среде Веб
- Поддержка метаданных для системы и пользователей
- Неоднородные информационные ресурсы
- Распределенные информационные ресурсы
- Материализованные и виртуальные коллекции
- Интеграция информационных ресурсов на различных уровнях
- Развитые средства каталогизации информационных ресурсов
- Пользовательские интерфейсы с повышенным уровнем семантики, с возможностями их персонализации
- Мультиязыковые текстовые информационные ресурсы и интерфейсы системы
- Встроенные сервисы: поиск, визуализация данных, статистический аппарат и др.

Цели создания научных ЭБ

- Обеспечение научным работникам возможности быстрого доступа к необходимым информационным ресурсам
- Предоставление результатов фундаментальных научных исследований широкому исследовательскому сообществу
- Создание новых технологий научных исследований, эффективного инструментария для их проведения
- Предотвращение утраты ценных научных коллекций для будущих поколений ученых
- Обеспечение возможностей для научного сотрудничества не только в региональном, ведомственном, национальном, но и в международном масштабе.

Состояние разработок научных ЭБ

- Охватываются практически все области знаний
- Имеется полный спектр систем по их масштабу:
 - ЭБ конкретного журнала или группы журналов
 - ЭБ отдельного исследовательского коллектива
 - ЭБ научного учреждения, региона, ведомства
 - национальная/международная ЭБ
- По архитектуре:
 - от сосредоточенных до распределенных
 - от материализованных до виртуальных
- По тематике коллекций информационных ресурсов:
 - универсальные
 - для данной области знаний или научного направления
 - на основе личных архивов крупных ученых
- По технологиям:
 - отдельные Веб-сайты, системы баз данных, текстовые системы
 - системы, основанные на интеграции технологий.

Имеются ли особенности у научных ЭБ ?

Критерии оценки:

- Область применения
- Функциональные возможности
- Характер и содержание информационных ресурсов
- Используемые методы и технологии.

Рассматриваются невырожденные случаи.

Область применения научных ЭБ

- Продвинутое и изощренное пользователи
- Нерегламентируемые информационные потребности
- Динамичный характер информационных потребностей
- Стремление пользователей к сотрудничеству
- Необходимость обмена информацией
- Многоязыковая среда научного сообщества
- Разнообразие предметов исследования
- Разнообразие технологий научного исследования
- Неоднородность используемой информации
- Оперирование данными на разных уровнях абстракции
- Демократичность научного сообщества.

Функции научных электронных библиотек

- Хранение и актуализация коллекций информационных ресурсов
- Поддержка взаимосвязей информационных ресурсов
- Обеспечение доступа пользователей к коллекциям
- Поддержка метаданных, в том числе каталогов коллекций
- Формирование коллекций (оцифровка, регистрация результатов наблюдений и экспериментов непосредственно в процессе их проведения)
- Интеграция издательских технологий и технологий формирования коллекций
- Поддержка различных сведений о пользователях
- Предоставление различных встроенных или настроенных сервисов - приложений, превращающих ЭБ в исследовательский полигон (виртуальная обсерватория, виртуальная химическая или биологическая лаборатория и т.п.).

Содержание информационных ресурсов научных ЭБ

- Научные публикации в различных формах (статьи, доклады, монографии, диссертации, авторефераты, ...)
- Библиографическая информация
- Персоналия
- Событийная информация (календарь конференций и т.п.)
- Результаты различного рода экспериментов, наблюдений, измерений, моделирования исследуемой реальности
- Модели исследуемых процессов, явлений, феноменов, представленные в разнообразных формах
- Метаданные, описывающие такие ресурсы
- Разнообразные научные коллекции и их элементы
- Каталоги коллекций и описания их элементов, классификаторы и другие средства систематизации.

Свойства информационных ресурсов научных ЭБ

- Неоднородность информационных ресурсов (в различных аспектах)
- Разнообразии сред представления - текст, числовые данные, статические изображения, видео, аудио, мультимедиа ...
- Сверхбольшие объемы информационных ресурсов
- Свойства информационных ресурсов описываются метаданными для системы и для пользователей
- Виртуальные коллекции информационных ресурсов
- Часто используются не реальные, а гипотетические данные
- Существенное значение имеет фактор старения информационных ресурсов
- Представление информационных ресурсов предметной области исследования в форме, позволяющей непосредственно проводить исследование (не информационная поддержка исследований, а *исследовательский полигон*).

Технологии, используемые в ЭБ

- Библиотечные технологии и стандарты каталогизации
- Традиционные технологии баз данных, базы данных XML
- Технологии текстового поиска, современные модели поиска, индексирование текста, изображений, аудио, видео
- Веб-технологии, платформа XML
- Объектные технологии
- Технологии ГИС
- Технологии интероперабельности
- Методы, модели и архитектуры интеграции неоднородных информационных ресурсов
- Технологии Java
- Компонентные технологии
- Различные подходы к представлению метаданных (стандарты XML, DC, стандарты геоданных ...)
- Протоколы глобального доступа к информационным ресурсам (в частности, к распределенным ресурсам)

Сводка свойств научных ЭБ - 1

- Специфическая предметная область
- Продвинутое, изощренное пользователи
- Динамичность и нерегламентированный характер потребностей пользователей
- Важна роль метаданных, ориентированных не только на систему, но и на пользователей
- Повышенные требования к пользовательским интерфейсам
- Способность поддержки виртуальных коллекций
- Мультиязыковый характер интерфейсов и информационных ресурсов
- Неоднородные коллекции информационных ресурсов
- Разнообразие содержания информационных ресурсов

Сводка свойств научных ЭБ - 2

- Сложные взаимосвязи между ресурсами различных видов
- Очень крупные объемы информационных ресурсов
- Нетрадиционные технологии формирования коллекций информационных ресурсов
- Потребность в интеграции издательских технологий и технологий формирования коллекций
- Демократичный характер доступа к ресурсам
- Стремление к использованию свободно распространяемого программного обеспечения
- Использование технологий открытых систем.

Перспективные направления развития ЭБ

- Разработка содержания метаданных для конкретных классов ЭБ
- Семантическая интеграция неоднородных коллекций
- Методы индексирования и поиска нетекстовых информационных ресурсов
- Технологии создания и поддержки электронных коллекций
- Использование стандартов и технологий XML для представления и описания информационных ресурсов
- Методология и инструментальные средства разработки ЭБ
-

Выводы

- Спектр используемых в научных ЭБ архитектурных решений и информационных технологий не специфичен для систем этого класса
- Научные ЭБ обладают специфической предметной областью, и с этим связаны особые требования к ним
- Специфичны по содержанию и неоднородны их коллекции информационных ресурсов
- В ряде случаев используются специфические методы формирования коллекций информационных ресурсов
- Научные ЭБ оснащаются специфическими сервисами, благодаря которым они могут использоваться не только как источник информационных ресурсов, но и как полигон для непосредственных научных исследований.

Спасибо за внимание