

О методах сравнения информационных систем, их классификации и специальных сервисах

Жижимов О.Л.

Институт вычислительных технологий СО РАН, Новосибирск

Аннотация

В докладе рассматриваются вопросы, связанные со сравнительными характеристиками информационных систем как пакета программного обеспечения, как инсталлированного пакета программного обеспечения, как информационного хранилища. Различные характеристики информационных систем могут быть сгруппированы в некоторый классификатор, на основании которого могут быть описаны характеристики произвольной информационной системы. Предлагаемый классификатор и, соответственно, схема данных могут служить основой для сравнения различных информационных систем и, как частное, так называемых электронных библиотек. Обсуждается необходимость создания специальных сервисов, информирующих клиента о возможностях той или иной информационной системы, необходимых для создания адаптивных информационных систем.

Обилие программного обеспечения для построения электронных библиотек, большое количество инсталляций этого программного обеспечения и наполнение этих инсталляций разнообразным информационным содержанием принуждает рассматривать мир электронных библиотек как множество слабосвязанных сущностей, объединяемых на первый взгляд только неудачным общим названием «электронные библиотеки». Каждый элемент этого множества имеет как некоторые общие характеристики, совокупность которых позволяет относить его к множеству электронных библиотек, так и уникальные характеристики, позволяющие различать элементы между собой.

Цель настоящей работы – определение набора характеристик, позволяющего идентифицировать любую электронную библиотеку с точностью, достаточной для описания основных ее свойств. Этот набор характеристик может называться паспортом электронной библиотеки или, другими словами, схемой данных описания электронных библиотек. Именно последний термин будет использоваться ниже для обозначения указанного набора характеристик. При этом термин «электронная библиотека» здесь будет употребляться вместе с термином «информационная система». Ниже оба термина будут равнозначны, т.к. любая электронная библиотека представляет собой, прежде всего, информационную систему, в которой усилены те или иные ее качества. При этом основное качество ЭБ – поглощение, каталогизация и хранение электронных документов, предоставление доступа к этим документам. Поэтому можно говорить, что электронная библиотека – это специализированная информационная система, ориентированная на работу с электронными документами [1]. Суть и глубина этой работы могут быть различны и зависят от совокупности требований, которые предъявляют к ЭБ на этапе ее разработки. Реализация этих требований порождает список технических характеристик той или иной ЭБ. Совокупность этих характеристик отличает одну ЭБ от другой, т.е. является уникальной характеристикой ЭБ.

Следует заметить, что в настоящее время термин «электронная библиотека» используется в литературе в различных смыслах. Этим термином могут обозначаться

- Набор специального программного обеспечения, предназначенного для создания информационных систем и хранилищ электронных документов.
- Набор реализованных сервисов на базе определенного программно-аппаратного комплекса – конкретная инсталляция и конфигурация специального программного обеспечения.
- Совокупность информационного контента конкретной информационной системы и сервисов доступа к этому контенту.

Исходя из этого, набор параметров, обеспечивающий исчерпывающую характеристику ЭБ, можно рассматривать как схему метаданных для описания ЭБ как

- Пакета специального программного обеспечения
- Локальной инсталляции набора специализированных пакетов
- Информационного контента текущей инсталляции

Следовательно, описание объекта «электронная библиотека» должно содержать блоки, описывающие каждую из вышеуказанных ипостасей. Здесь и ниже будет использоваться представление схематичной записи в формате XML без строгого определения элементов.

Итак, исходя из вышесказанного, на верхнем уровне иерархии элементов описание объекта «электронная библиотека» должно выглядеть так

```
<elibInfo>
  <softwarInfo> . . . . </softwarInfo>
  <installInfo> . . . . </installInfo>
  <contentInfo> . . . . </contentInfo>
</elibInfo>
```

Справедливости ради следует отметить, что вопросам определения набора параметров для описания характеристик электронных библиотек и информационных систем посвящено много работ. Дело в том, что вопрос о списке этих параметров возникает всегда, когда появляется желание сравнить между собой некоторые специализированные программные пакеты для построения электронных библиотек или существующие их инсталляции с соответствующим информационным наполнением (см., например, [2-4]). Тем не менее, следует отметить несколько работ, результаты которых частично могут использоваться при формировании детальной структуры <elibInfo>.

Для формирования детальной структуры <softwarInfo> безусловно полезны таблицы сравнительных характеристик, приведенные в работах, например, [5-7]. Ценность этих работ здесь определяется не столько сравнительными характеристиками различного программного обеспечения (DSpace [<http://dspace.org>], OPUS, Eprints [<http://software.eprints.org>], Fedora, CDSware [<http://cdsware.cern.ch/>] и др.), сколько фиксацией списка параметров, по которым это сравнение проводилось.

Для формирования детальной структуры элементов <installInfo> и <contentInfo> безусловно следует обратиться к унифицированным описаниям баз данных и сетевых сервисов, структура которых зафиксирована в стандарте ISO-23950 (ANSI/NISO Z39.50) [8] в так называемых категориях сервиса Explain. Этим стандартом определены правила доступа и структура базы данных IR-Explain-1, записи которой содержат описания всех текущих

сервисов и структур, относящихся к текущему серверу, предоставляющему доступ к информационным ресурсам. Конечно, соответствующая схема данных (схема данных Explain) изначально ориентирована на описание сервисов, доступных по протоколу Z39.50. Однако, многие элементы этой схемы вполне могут быть использованы для формирования структур <installInfo> и <contentInfo>, т.к. изначально ориентированы на описание тех же самых сущностей.

Более простой подход может быть основан не на схеме данных Explain, а на схеме данных ZeeRex (<http://explain.z3950.org>), включающей только некоторые элементы из схемы Explain, упрощая тем самым представление описания информационной системы и всех ее компонент.

Исходя из вышесказанного, структура <installInfo> должна включать следующие информационные блоки

- Общее описание сервера
- Описание сервисов
- Описание интерфейсов доступа к сервисам
- Фиксация правообладателя
- Фиксация административной контактной информации

В то время как структура <contentInfo> должна включать

- Описание информационных ресурсов
- Описание интерфейсов доступа к ресурсам

При этом информационными ресурсами могут являться

- Базы данных
- Коллекции
- Отдельные документы

Видно, что в схеме описания любой информационной системы, в том числе и в описании электронной библиотеки, должны присутствовать элементы, содержащие информацию, аналогичную информации из категории targetInfo схемы данных Explain. При этом некоторые элементы targetInfo должны быть включены в структуру <installInfo>, описывающую конфигурацию конкретной информационной системы.

```
<installInfo>
  <name> . . . </name>
  <description> . . . </description>
  <accessInfo>
    <host> . . . </host>
    <port> . . . </port>
    <authentication> . . . </authentication>
  </accessInfo>
  <administrator>
    <name> . . . </name>
    <mail> . . . </mail>
    <phone> . . . </phone>
  </administrator>
</installInfo>
```

Технические характеристики специализированных программных комплексов для построения электронных библиотек

Как и любые информационные системы, электронные библиотеки могут быть классифицированы по характеристикам, которые можно сгруппировать

1. По типам хранимых ресурсов
2. По типам обрабатываемых метаданных
3. По возможностям поиска информации
4. По интерфейсам доступа
5. По интерфейсам ввода информации
6. По архитектуре
7. По аппаратной среде исполнения
8. По программной среде исполнения
9. По способу хранения и обработки данных
10. По возможностям ограничения доступа к ресурсам
11. По возможностям обеспечения интероперабельности
12. По возможностям сбора статистики
13. По способу реализации, лицензирования и поддержки
14. По используемым стандартам и рекомендациям

Ниже будут рассмотрены все эти группы характеристик более подробно. Полное описание параметров приведено в Приложении.

1. Хранимые ресурсы

Одна из основных характеристик ЭБ – перечень типов поддерживаемых электронных документов и файлов. При этом эти типы должны определять

1. Смысловой тип документа – монография, статья, презентация, изображение, видеозапись и пр.
2. Физический формат представления документа – документы PDF, MS Word, MS PowerPoint, MS Excel, TeX, текстовые документы, файлы в графических растровых или векторных форматах, файлы в аудио- и видеоформатах, и т.п.
3. Поддерживаемые типы текстового кодирования документов (языки и кодовые таблицы)
4. Поддерживаемые типы сжатия документов, работа с архивами (zip, gz, rar и пр.) «на лету».
5. Максимальный размер обрабатываемых документов с привязкой к формату представления.
6. Возможность распознавания внутренней структуры документов.
7. Возможность преобразования документов из одного физического формата в другой (pdf – txt, doc - txt, doc - rtf, jpg – txt, jpg – gif и т.п.) по требованию пользователя.
8. Максимальное количество документов, сопоставляемых с единицей учета и хранения в ЭБ.
9. Максимальное количество документов, обработка которого допустима в ЭБ.
10. Максимальный суммарный размер всех документов, т.е. максимально допустимый размер хранилища ЭБ.

11. Возможность определения множественных коллекций в рамках одной инсталляции системы

Определение перечисленных параметров позволяет получить представление о подсистеме хранения ресурсов в ЭБ, а также сравнивать между собой основные характеристики различных ЭБ.

2. Метаданные

Используемые в информационных системах метаданные описывают различные характеристики обрабатываемых и хранимых в системе информационных ресурсов. Здесь нас будут интересовать только те метаданные, которые описывают собственно информационный ресурс, а не детали его существования в конкретной информационной системе. Эти метаданные должны появляться в информационной системе или ЭБ в момент погружения в нее электронного документа и содержать полное описание этого электронного документа в той степени детализации, которую способна обеспечить эта ЭБ.

По способу создания в ЭБ метаданных, или каталогизации информационных ресурсов, можно характеризовать ЭБ следующими характеристиками

1. Возможность ручного ввода метаданных через административные интерфейсы взаимодействия ЭБ с ее пользователями.
2. Возможность пакетной загрузки метаданных в фиксированных форматах и схемах, представленных в стандартизованном виде.
3. Возможность заимствования метаданных из других информационных систем и ЭБ, в которых обрабатываемый информационный ресурс (электронный документ) был уже представлен.
4. Возможность автоматической каталогизации информационного ресурса в момент его погружения в ЭБ, например, автоматическое извлечение тегов из файлов изображений (TIFF, JPG) или из документов MS Word (DOC, DOCX, RTF).

Совокупность всех используемых для описания информационного ресурса в ЭБ элементов метаданных образует набор метаданных, который может регламентироваться определенной схемой метаданных. При этом схемы метаданных могут быть локальными и глобальными. Глобальные схемы определяются соответствующими международными стандартами или рекомендациями уважаемых организаций. По отношению к схемам метаданных ЭБ характеризуют следующие параметры.

5. Список поддерживаемых схем метаданных для внутренней обработки документов в ЭБ.
6. Список автоматически заполняемых элементов метаданных, например, дата ввода документа в ЭБ.
7. Наличие механизмов автоматического преобразования метаданных между различными схемами.
8. Возможность получения метаданных в различных форматах (ISO-2709, GRS-1, XML и др.).
9. Возможность расширения номенклатуры поддерживаемых схем и форматов метаданных штатными средствами управления ЭБ.
10. Список ограничений на возможности расширения поддерживаемых схем метаданных.
11. Поддержка кодировки UNICODE для метаданных

3. Поиск информации

Параметры, отражающие поисковые возможности ЭБ, определяют одну из важнейших ее характеристик. При этом различают два вида поиска: поиск по метаданным и поиск по электронным документам. Соответствующие характеристики ЭБ приведены ниже

- Поддержка поиска по метаданным.
- Поддержка поиска по электронным документам.

Однако указание этих параметров определяет очень грубую характеристику поисковых возможностей ЭБ. Так поиск по метаданным должен быть уточнен или указанием схем метаданных, или указанием поисковых атрибутов при условии поддержки модели поиска Z39.50 (ISO-23950).

4. Интерфейсы доступа к информации

Доступ к хранимым в ЭБ электронным ресурсам осуществляется пользователем посредством пользовательских интерфейсов. Эти интерфейсы могут обеспечивать локальный и сетевой режим работы с ЭБ. При этом локальные интерфейсы доступа к ресурсам не отрицают возможности сетевого доступа, который в этом случае может быть обеспечен внешними по отношению к ЭБ средствами. Например, пусть при создании ЭБ разработчики снабдили ее графическими интерфейсами, реализованными в виде специализированных локальных приложений. Эти пользовательские приложения обеспечивают полнофункциональный доступ к ЭБ, которая функционирует на том же компьютере, т.е. локально. Однако, даже в этом локальном варианте можно обеспечить сетевой доступ к ЭБ посредством возможностей самой ОС через X-Windows, Remote Desktop, VNC и др.

В последнее время наибольшей популярностью пользуются интерфейсы доступа к ЭБ на основе WWW. При этом в качестве клиентского приложения выступает обычный WEB-браузер. При работе в сетевом режиме «клиент-сервер» такой подход требует или реализации в серверной части ЭБ сервера WWW или, наоборот, реализации ЭБ в виде отдельного приложения самостоятельного сервера WWW. Последний вариант соответствует также выполнению ЭБ в среде так называемого сервера приложений.

При любой реализации интерфейсы доступа к ЭБ должны обеспечивать функциональность, соответствующую возможностям ЭБ.

5. Интерфейсы ввода информации

Ввод информации в ЭБ обеспечивает ее пополнение новыми информационными ресурсами. В разделе должны присутствовать следующие параметры

- Ввод информации обеспечивается собственными интерфейсами ЭБ
- Ввод информации обеспечивается внешними по отношению к ЭБ средствами
- ЭБ содержит интерфейсы для интерактивного ввода данных
- ЭБ поддерживает возможность ввода данных в пакетном режиме
- Наличие возможности загрузки архивированных файлов
- Наличие возможности загрузки файлов из источника по URL

- Список схем метаданных, ввод в которых поддерживается ЭБ
- Возможность изменения списка метаданных при вводе информации
- Возможность изменения интерфейсов ввода информации при расширении списка элементов метаданных
- Наличие автоматически заполняемых полей
- Возможность заполнения полей данными из предлагаемого списка (словаря)
- Возможность установка различных параметров процесса ввода информации для каждой коллекции
- Наличие динамической настройки интерфейсов ввода информации
- Наличие отдельной рабочей области для процесса ввода информации
- Использование ролей для процесса ввода информации
- Возможность настройки ролей в пределах коллекции для процесса ввода информации
- Уведомление пользователей по E-mail о завершении процесса ввода информации
- Уведомление администраторов по E-mail о завершении процесса ввода информации
- Возможность просмотра пользователями списка незавершенных процессов ввода информации
- Возможность просмотра пользователями списка завершенных процессов ввода информации
- Возможность просмотра администраторами списка процессов ввода информации
- Наличие запроса на ввод лицензии
- Хранение лицензии вместе с контентом
- Возможность интерактивного заимствования описаний из других систем
- Возможность редактирования введенных ранее описаний
- Обновление и индексирование введенных данных в реальном времени

6. Архитектура

Архитектура ЭБ явно не определяет ее функциональные характеристики. Однако неявно все характеристики ЭБ определяются именно ее архитектурой. Наиболее адаптируемы к различным условиям функционирования решения, которые основаны на модульной архитектуре. Это означает, что существует некое ядро системы, в котором реализованы лишь основные функции ЭБ. Основная же функциональность при этом обеспечивается подключением дополнительных динамических модулей, взаимодействующих с ядром через соответствующие API в рамках одного процесса, или посредством документированного протокола обмена данными в рамках многозадачной или многоагентной системы.

7. Аппаратная среда исполнения

Любая информационная система функционирует на конкретной аппаратуре. Этот раздел включает описание необходимого аппаратного минимума для функционирования системы (тип процессора, разрядность, минимальная оперативная память и т.д.).

8. Программная среда исполнения

ЭБ как исполняемое приложение или комплекс исполняемых приложений для функционирования требует некой среды. Этой средой могут быть: собственно операционная система, настройка над операционной системой - некоторый framework, обеспечивающий взаимодействие ЭБ с внешними динамическими библиотеками системных и прикладных объектов, например, среда .NET, а также специальные приложения, выполняющие функции интерпретаторов универсальных программных кодов, например Java, или интерпретируемых языков программирования, например, PHP. В роли среды исполнения приложений ЭБ могут выступать так называемые «сервера приложений», которые обеспечивают достаточно универсальную среду для выполнения разнообразных задач. Следует заметить, что первичной программной средой для всех приложений, в том числе и для серверов приложений, является операционная система. Поэтому зависимость от типа операционной системы является, пусть даже иногда и косвенной, характеристикой ЭБ.

9. Способ хранения и обработки данных и метаданных

Собственно данные (электронные документы) в ЭБ могут храниться как в файловой системе в виде отдельных файлов, так и в той или иной СУБД, заполняя соответствующие поля в таблицах реляционных СУБД или соответствующие элементы в СУБД иерархических. Впрочем, хранение документов в иерархической СУБД формально мало чем отличается от хранения в файловой системе, т.к. любая современная файловая система реализуется как иерархическая СУБД. В зависимости способа хранения документов в ЭБ могут быть реализованы различные механизмы доступа к собственно информационным ресурсам в целях управления электронными документами и их индексирования для реализации функций полнотекстового поиска.

10. Ограничение доступа к ресурсам

Любая ЭБ требует управления ограничением доступа к своим информационным ресурсам. При этом могут использоваться различные технологии, механизмы и приемы, но конечная их цель одна – каждый пользователь ЭБ должен получить только тот уровень доступа к информационному ресурсу, на который он имеет право в силу своих персональных, групповых или территориальных полномочий. Поэтому задача управления доступом к ресурсам ЭБ сводится как правило к следующим задачам:

- определение статуса информационного ресурса и правил его использования;
- определение персональных характеристик конкретного пользователя в момент его обращения к сервисам ЭБ (идентификация);
- проверка достоверности идентификации пользователя (аутентификация);
- определение уровня прав пользователя на запрашиваемый информационный ресурс (авторизация);
- принятие решения о доступе к ресурсу.

Каждая из этих задач порождает свой список характеристик, определяющих способы ее решения для конкретной ЭБ. Так при определении статуса и правил использования информационного ресурса можно говорить о нескольких уровнях полномочий для пользователей:

- Создание ресурса, т.е. помещение электронного документа в хранилище
- Удаление ресурса
- Редактирование метаданных
- Просмотр укороченного набора метаданных
- Просмотр полного набора метаданных
- Возможность исполнения пакетных операций импорта и экспорта метаданных
- Извлечение и просмотр первичного ресурса (электронного документа).
- Возможность исполнения пакетных операций импорта и экспорта данных электронных документов.

Эти полномочия могут фиксироваться в специальных элементах метаданных, хранятся и обрабатываются в соответствии с общими для данной ЭБ правилами работы с метаданными.

Идентификация пользователя, как правило, производится методом сверки явно или неявно указанных пользователем идентификационных параметров с эталонными параметрами, к которым возможен доступ из ЭБ. Механизмы идентификации пользователей необходимы для установления личности пользователя и проверки достоверности явно сообщаемых им идентификационных параметров при обращении к тем или иным сервисам ЭБ. Примерами могут служить

- доступ к реестру операционной системы для получения информации о пользователях ЭБ и их аутентификации;
- доступ к специализированной или универсальной СУБД, хранящей информацию о зарегистрированных в системе пользователях и группах пользователей;
- доступ к каталогам LDAP, в которых хранится аналогичная информация.

Заметим, что пользователем ЭБ может выступать или человек, или компьютер. В первом случае в качестве идентификационного параметра выступает имя (синоним имени) или код пользователя, а во втором – сетевой идентификатор компьютера (IP-адрес или DNS имя). Немаловажной характеристикой идентификации пользователей является ее масштабируемость и интероперабельность, т.е. возможность неограниченного расширения списка зарегистрированных пользователей и возможность использовать внешние для данной ЭБ механизмы идентификации, в том числе через среду Интернет. Эти же свойства ЭБ существенны и для связанных с идентификацией механизмов аутентификации.

11. Интероперабельность

Под интероперабельностью ЭБ следует понимать их способность к взаимодействию с другими системами, в том числе с другими электронными библиотеками, информационными и поисковыми системами. При этом можно обсуждать возможность взаимодействия на различных уровнях.

- Уровень данных - импорт и экспорт данных в принятых форматах
- Уровень метаданных – импорт и экспорт метаданных в стандартных внешних форматах и схемах
- Уровень пользовательских и административных интерфейсов - наличие интуитивно понятных пользовательских интерфейсов доступа к информации и возможность использования их для доступа к другим системам

- Уровень сетевых протоколов - реализация доступа к информации по стандартным протоколам (HTTP, FTP, Z39.50)
- Уровень организационных правил и политик

12. Сбор статистики

Любая информационная система должна собирать информацию о текущей активности, загрузке, об использовании информационных ресурсов, об ошибках и внештатных ситуациях. Уровень детализации этой информации для разных систем может быть совершенно различен. Наиболее продвинутые системы имеют встроенные средства для просмотра этой информации и созданию соответствующих отчетов по различным критериям отбора статистической информации. Для некоторых систем можно воспользоваться сторонними разработками для анализа статистики и получения отчетов. В качестве параметров, характеризующих подсистему сбора и обработки статистической информации в ЭБ, можно выбрать следующие.

- Присутствует встроенная подсистема сбора информации
- Используются системные механизмы сбора информации
- Используются внешние подсистемы сбора информации
- Имеются встроенные средства обработки статистической информации
- Имеются средства событийного управления

13. Реализация, лицензирование и поддержка

В этом разделе группируются параметры, описывающие интеллектуальную ответственность, выходные данные, тип лицензирования, процедуру инсталляции и системную поддержку. Список параметров приведен в Приложении.

14. Стандарты

В этом разделе перечисляются поддерживаемые стандарты и рекомендации с указанием номеров версий.

Характеристики локальной инсталляции

Как любой сетевой ресурс каждая локальная инсталляция программного комплекса, на базе которого создается электронная библиотека, имеет свои отличительные характеристики

- Название ЭБ
- Указание ответственного за инсталляцию ЭБ
- Указание ответственного за содержание ЭБ
- Указание способов доступа к информации ЭБ
- Указание ограничений на доступ к информации ЭБ
- Указание географического места расположения ЭБ
- Технические характеристики инсталляции

Характеристики информационного контента

Фактическое информационное наполнение ЭБ характеризуется параметрами, которые не относятся к ее техническим характеристикам. Ниже перечислены группы наиболее значимых параметров ЭБ, характеризующих ее информационное наполнение.

- Количество различных коллекций и баз данных
- Количество элементов или записей в каждой коллекции и базе данных
- Характеристики, в том числе и тематическая направленность каждой коллекции и базы данных, в том числе описание ограничений на доступ.
- Описание связей между коллекциями, базами данных и записями
- Описание словарей, справочников, рубрикаторов и тезаурусов, используемых при формировании информационного контента ЭБ
- Описание возможностей поиска информации
- Описание возможностей навигации
- Описание возможностей извлечения и просмотра информации

О специальных сервисах

Возникает вопрос, зачем нужно обсуждаемое выше детализированное описание информационной системы или электронной библиотеки и как это описание может быть использовано при реальной работе.

Описание `<softwarInfo>` как паспорт пакета программного обеспечения было бы очень полезно администраторам информационных систем при планировании развертывания той или иной системы. Сравнивая между собой эти формальные описания можно было бы получать объективную сравнительную оценку соответствующих систем и, исходя из этого, принимать решение о развертывании. Отсутствие формализованных единообразных описаний существенно затрудняет этот непростой процесс и заставляет администраторов самостоятельно строить соответствующие таблицы сравнения на основе программной документации.

Описания `<installInfo>` и `<contentInfo>` были бы полезны при сравнении инсталлированных и развернутых информационных систем.

Для последних категорий можно отметить еще один момент. Эволюция мировой информационной инфраструктуры имеет тенденцию к интеграции разрозненных информационных систем в нечто единое, но распределенное (DATA-GRID - ?). Однако, заставить совокупность отдельных информационных систем функционировать как нечто связанное можно лишь основе их полной интероперабельности. Эта интероперабельность должна кроме всего прочего (стандарты, протоколы, запросы, схемы, форматы и т.п.) включать возможность взаимного информирования систем о своих функциональных возможностях и о своем информационном наполнении. Без этого информирования невозможно обеспечить свойство адаптивности информационной системы при интеграции ее в какой-либо «DATA-GRID». Информация, представленная в структурах `<installInfo>` необходима для обеспечения функциональной адаптивности, а в структурах `<contentInfo>` - адаптивности информационной.

Справедливости ради следует заметить, что на обеспечение функциональной адаптивности информационных систем в части WEB-сервисов направлена технология на основе WSDL [9], в части описания программных интерфейсов – IDL. На обеспечение информационной адаптивности информационных систем ориентированы технологии на основе XML, RDF, OWL. Каждая из упомянутых технологий решает частную задачу обеспечения адаптивности для специальных (на основе XML) систем. Однако задача обеспечения адаптивности информационных систем, несомненно, намного шире.

Что касается специальных информационных систем, то существуют упоминавшиеся выше решения для систем на основе Z39.50 с использованием специальных стандартизованных системных сервисов (Explain). Через сервис Explain системы на основе Z39.50 могут обмениваться информацией, составляющей в нашем случае содержание структур <installInfo> и <contentInfo>. Более того, минимальная информация о функциональных возможностях любого сервера Z39.50 присутствует в APDU в виде битовых строк при инициализации любого сеанса Z39.50, что обеспечивает минимальный уровень адаптивности для субъектов сетевого взаимодействия в Z39.50. Идеология систем на основе Z39.50 была частично сохранена при попытке реализовать функциональность Z39.50 в WEB-системах в виде технологий XML/SOAP/SRW – ZeeRex (см. <http://explain.z3950.org>).

В заключение следует заметить, что сегодня не существует готовых технологических решений для обеспечения адаптивности разнородных информационных систем. Не исключено, что первым шагом для создания таких решений может быть создание детализированного описания произвольной информационной системы и предоставление этой информации или в виде фиксированного файла для распространяемых пакетов программного обеспечения, или в виде ответа на специальный запрос к соответствующей системе для инсталлированных систем.

Литература

1. Жижимов О.Л., Мазов Н.А., Федотов А.М. О некоторых отличиях электронных библиотек от хранилищ цифрового контента традиционных библиотек [Электронный ресурс] / О.Л.Жижимов // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: материалы конф. – Электрон. дан. – М.: ГПНТБ России, 2010. – 1 электрон. опт. диск (CD-ROM). – Систем. требования: IBM PC, Windows 2000 или выше. – Загл. с этикетки диска. – ISBN 978-5-85638-139-8. – № гос. регистрации 0321000673. - <http://www.gpntb.ru/win/inter-events/crimea2010/disk/131.pdf>
2. Резниченко В.А., Проскудина Г.Ю., Овдий О.М. Создание цифровой библиотеки коллекций периодических изданий на основе Greenstone. // Электронные библиотеки. 2005. – 8. Вып. 6. <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2005/part6>.
3. Новицкий А.В., Резниченко В.А., Проскудина Г.Ю. Создание научных архивов с помощью системы EPrints.// Электронные библиотеки. 2006. –Том 9. Вып. 4. <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2006/part4/Novitski>
4. Кудим К.А., Проскудина Г.Ю., Резниченко В.А. Сравнение систем электронных библиотек EPrints 3.0 и DSpace 1.4.1 // Электронные библиотеки: перспективные методы и технологии, электронные коллекции. Труды Девятой Всероссийской научной конференции RCDL'2007 (Переславль-Залесский, Россия, 15-18 октября 2007 г.).

Переславль-Залесский: изд.-во «Университет города Переславля», 2007. - ISBN 978-5-901795-09-5. – с.241-252.

5. A Guide to Institutional Repository Software. 3rd Edition // Open Society Institute. – August 2004. – p.28.
6. Dion Hoe-Lian Goh, Alton Chua, Davina Anqi Khoo, Emily Boon-Hui Khoo, Eric Bok-Tong Mak and Maple Wen-Min Ng. A checklist for evaluating open source digital library software. // Online Information Review, Vol. 30 No. 4, 2006. - pp. 360-379.
7. Cyzyk Mark, Choudhury Sayeed. A Survey and Evaluation of Open-Source Electronic Publishing Systems // The Open Society Institute (OSI). - Technical Report. - 28 Apr 2008. - p.70 - [<http://jhir.library.jhu.edu/handle/1774.2/32737>].
8. ANSI/NISO Z39.50-1995. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification / Z39.50 Maintenance Agency Official Text for Z39.50-1995. - July 1995.
9. Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language // W3C Recommendation 26 June 2007. - <http://www.w3.org/TR/wsdl20>.

Приложение 1. Таблица параметров, включаемых в структуру <softwarInfo>.

В таблице дополнительно указаны ссылки на другие рубрикаторы (см. список литературы)

Название параметра	[5]	[6]
1. Хранимые ресурсы		
1.1 Смысловой тип хранимых документов – монография, статья, презентация, изображение, видеозапись и пр.		
1.2 Поддерживаемые физические форматы представления документов – документы PDF, MS Word, MS PowerPoint, MS Excel, TeX, текстовые документы, файлы в графических растровых или векторных форматах, файлы в ауди- и видеоформатах, и т.п.	13.2	2.2.
1.3 Поддерживаемые типы текстового кодирования документов (языки и кодовые таблицы)		
1.4 Поддерживаемые типы сжатия документов, работа с архивами (zip, gz, rar и пр.) «на лету».		
1.5 Максимальный размер обрабатываемых документов с привязкой к формату представления.		
1.6 Возможность распознавания внутренней структуры документов.		
1.7 Возможность преобразования документов из одного физического формата в другой (pdf – txt, doc - txt, doc - rtf, jpg – txt, jpg – gif и т.п.) по требованию пользователя.		
1.8 Максимальное количество документов, сопоставляемых с единицей учета и хранения в ЭБ.		
1.9 Максимальное количество документов, обработка которого допустима в ЭБ.		
1.10 Максимальный суммарный размер всех документов, т.е. максимально допустимый размер хранилища ЭБ.		
1.11 Возможность определения множественных коллекций в рамках одной инсталляции системы	10.1	1.1.1
2. Метаданные	14	3
2.1. Возможность ручного ввода метаданных через административные интерфейсы взаимодействия ЭБ с ее пользователями.		
2.2. Возможность пакетной загрузки метаданных в фиксированных форматах и схемах, представленных в стандартизованном виде.		
2.3. Возможность заимствования метаданных из других информационных систем и ЭБ, в которых обрабатываемый информационный ресурс (электронный документ) был уже представлен.		
2.4. Возможность автоматической каталогизации информационного ресурса в момент его погружения в ЭБ, например, автоматическое извлечение тегов из файлов изображений (TIFF, JPG) или из документов MS Word (DOC, DOCX, RTF).		
2.5. Список поддерживаемых схем метаданных для внутренней обработки документов в ЭБ.	14.1	3.2
2.6. Список автоматически заполняемых элементов метаданных, например, дата ввода документа в ЭБ.	14.7	3.4

2.7.	Наличие механизмов автоматического преобразования метаданных между различными схемами.		
2.8.	Возможность получения метаданных в различных форматах (ISO-2709, GRS-1, XML и др.).		
2.9.	Возможность расширения номенклатуры поддерживаемых схем и форматов метаданных штатными средствами управления ЭБ.	14.2	
2.10.	Список ограничений на возможности расширения поддерживаемых схем метаданных.		
2.11.	Поддержка кодировки UNICODE для метаданных	14.8	3.5
2.12.	Обновление метаданных и переиндексирование «на лету»	8	3.1
3.	Поиск информации, навигация, сортировка	17	4
3.1.	Поддержка поиска по метаданным		
3.1.1.	Поиск по всем метаданным	17.2	4.2
3.1.2.	Список элементов метаданных, по которым может осуществляться поиск	17.3	4.3
3.1.3.	Поддержка поисковой модели Z39.50 (ISO-23950)		
3.1.3.1.	Список наборов поисковых атрибутов		
3.1.3.2.	Список поисковых атрибутов в каждом поддерживаемом наборе		
3.1.3.3.	Список комбинаций поисковых атрибутов		
3.1.4.	Поддержка географического поиска		
3.1.4.1.	Поиск по географическим названиям		
3.1.4.1.1.	Использование словаря (тезауруса) географических названий		
3.1.4.1.2.	Использование сервисов геокодирования		
3.1.4.2.	Поиск по географическим координатам		
3.1.4.2.1.	Список допустимых графических примитивов: точка, линия, окружность, четырехугольник, полигон		
3.2.	Поддержка поиска по электронным документам.		
3.2.1.	Поддержка текстового поиска	17.1	4.1
3.2.1.1.	Использование булевой логики	17.1.1	4.1.1
3.2.1.2.	Усечения и текстовые шаблоны	17.1.2	4.1.2
3.2.1.3.	Поиск по фразам		4.1.3

3.2.1.4. Учет близости		4.1.4
3.2.1.5. Поиск по всем словоформам с учетом орфографии	17.1.3	
3.2.2. Поддержка поиска по бинарному шаблону		
3.2.3. Поддержка поиска по графическим шаблонам		
3.2.4. Навигация по найденной информации	17.4	4.4
3.2.5. Сортировка результатов поиска	17.5	4.5
4. Интерфейсы доступа к информации		
4.1. Поддержка локальных пользовательских интерфейсов доступа к информации		
4.2. Поддержка пользовательских интерфейсов доступа к информации на основе WWW		
4.3. Поддержка пользовательских интерфейсов доступа к информации на основе Z39.50		
4.4. Поддержка пользовательских интерфейсов доступа к информации на основе FTP		
4.5. Поддержка пользовательских интерфейсов доступа к информации на основе специальных приложений		
4.6. Поддержка пользовательских настроек интерфейсов		
4.7. Поддержка многоязычных интерфейсов	16.3	9.3
4.8. Поддержка дискуссионного форума	16.5	
4.9. Наличие отдельной домашней страницы для каждой коллекции	10.1.2	1.1.3
4.10. Соответствие пользовательских интерфейсов функциональным возможностям ЭБ		
5. Интерфейсы ввода информации		
5.1. Ввод информации обеспечивается собственными интерфейсами ЭБ		
5.2. Ввод информации обеспечивается внешними по отношению к ЭБ средствами		
5.3. Наличие интерфейсов для интерактивного ввода данных		
5.4. Наличие возможности ввода данных в пакетном режиме		
5.5. Наличие возможности загрузки архивированных файлов	12.1	2.1.1
5.6. Наличие возможности загрузки файлов из источника по URL	12.2	2.1.2
5.7. Список схем метаданных, ввод в которых поддерживается ЭБ	14.1	

5.8.	Возможность изменения списка метаданных при вводе информации	14.6	3.3
5.9.	Возможность изменения интерфейсов ввода информации при изменении списка элементов метаданных		
5.10.	Возможность настройки автоматически заполняемых полей	14.7	3.4
5.11.	Возможность заполнения полей данными из предлагаемого списка (словаря или тезауруса)		
5.11.1.	Работа со статическими списками		
5.11.2.	Работа с динамическими списками из внешних СУБД		
5.12.	Возможность установка различных параметров процесса ввода информации для каждой коллекции	10.1.1	1.1.2
5.13.	Наличие динамической настройки интерфейсов ввода информации		
5.13.1.	По типу вводимого документа		
5.13.2.	По списку коллекций		
5.13.3.	По именованной группе коллекций или разделу		
5.14.	Наличие отдельной рабочей области для процесса ввода информации	10.2.1	1.2.1
5.15.	Использование ролей для процесса ввода информации	10.2.2	1.2.2
5.16.	Возможность настройки ролей в пределах коллекции для процесса ввода информации	10.2.3	1.2.3
5.17.	Уведомление пользователей по E-mail о завершении процесса ввода информации	10.3.1	1.3.1
5.18.	Уведомление администраторов по E-mail о завершении процесса ввода информации	10.3.2	1.3.2
5.19.	Возможность просмотра пользователями списка незавершенных процессов ввода информации	10.3.3.1	1.4.1
5.20.	Возможность просмотра пользователями списка завершенных процессов ввода информации	10.3.3.2	1.4.2
5.21.	Возможность просмотра администраторами списка процессов ввода информации	10.3.3.3	1.4.3
5.22.	Наличие запроса на ввод лицензии	10.3.4.1	
5.23.	Хранение лицензии вместе с контентом	10.3.4.2	
5.24.	Возможность интерактивного заимствования описаний из других систем		
5.25.	Возможность редактирования введенных ранее описаний		
5.26.	Обновление и индексирование введенных данных в реальном времени	15	3.1
6. Архитектура			

6.1.	Реализация ЭБ в виде модульной системы		
6.1.1.	Наличие документированных API межмодульного взаимодействия	8	
6.1.2.	Наличие шаблонов и примеров для разработки дополнительных модулей		
6.2.	Реализация двухуровневой архитектуры «клиент-сервер»		
6.3.	Реализация трехуровневой архитектуры «клиент - сервер приложений - сервер баз данных»		
7.	Аппаратное окружение	2	
7.1.	Минимальные требования к аппаратной платформе	2.1	
7.1.1.	Процессор	-	
7.1.2.	Разрядность процессора	-	
7.1.3.	Минимальная оперативная память	-	
7.1.4.	Минимальный объем дисковых накопителей	-	
7.1.5.	Наличие сетевых интерфейсов	-	
7.2.	Поддержка SAN	2.2	
7.3.	Поддержка NAS	-	
8.	Программное окружение	3	
8.1.	Возможность выполнения под управлением разных операционных систем	3.1	
8.1.1.	Возможность выполнения в среде Windows		
8.1.2.	Возможность выполнения в среде UNIX		
8.2.	Разрядность среды исполнения (16, 32, 64, 128)		
8.3.	Необходимость наличия дополнительного окружения		
8.3.1.	Исполнение в среде .NET, номер версии		
8.3.2.	Исполнение в среде серверов приложений, тип		
8.1.1.	Исполнение в среде WWW-сервера (CGI)		
8.1.2.	Исполнение в среде PHP		
8.1.3.	Исполнение в среде PERL		

8.3.3. Исполнение в среде ASP		
8.3.4. Исполнение в среде JAVA		
8.4. Необходимость наличия дополнительных сервисов		
8.1.4. Серверы баз данных	3.3	
8.4.1. WEB серверы	3.4	
8.4.2. Серверы приложений на основе JAVA	3.5	
8.4.3. Search engine	3.6	
8.1.5. Другие	3.7	
8.5. Язык программирования	3.2	
9. Способ хранения и обработки данных и метаданных		
9.1. Метаданные хранятся в файловой системе		
9.2. Метаданные хранятся в СУБД		
9.2.1. Иерархическая СУБД		
9.2.2. Реляционная СУБД		
9.2.3. Внутренняя специализированная СУБД		
9.3. Документы хранятся в файловой системе		
9.4. Документы хранятся в СУБД		
9.4.1. Иерархическая СУБД		
9.4.2. Реляционная СУБД		
9.4.3. Внутренняя специализированная СУБД		
10. Ограничение доступа к ресурсам		5
10.1. Поддержка пользователей – персоны		
10.1.1. Идентификация персон механизмами ОС		
10.1.2. Идентификация персон в СУБД		
10.1.3. Идентификация персон в каталогах LDAP		

10.2. Поддержка пользователей – компьютеры		
10.2.1. Идентификация механизмами ОС		
10.2.2. Идентификация механизмами СУБД		
10.2.3. Идентификация в каталогах LDAP		
10.3. Поддержка групп пользователей		
10.4. Аутентификация		
10.4.1. Аутентификация механизмами ОС		
10.4.2. Аутентификация механизмами СУБД		
10.4.3. Аутентификация механизмами LDAP		
10.5. Авторизация		
10.5.1. Авторизация механизмами ОС		
10.5.2. Авторизация механизмами СУБД		
10.5.3. Авторизация механизмами LDAP		
10.5.4. Авторизация внутренними механизмами ЭБ		
10.6. Поддержка уровней доступа	9.5	5.3
10.6.1. Создание ресурса, т.е. помещение электронного документа в хранилище		
10.6.1.1. Отдельный электронный документ		
10.6.1.2. Коллекция, группа документов		
10.6.1.3. Группа коллекций		
10.6.2. Удаление ресурса		
10.6.2.1. Отдельный электронный документ		
10.6.2.2. Коллекция, группа документов		
10.6.2.3. Группа коллекций		
10.6.3. Редактирование метаданных		
10.6.3.1. Отдельный электронный документ		

10.6.3.2. Коллекция, группа документов		
10.6.3.3. Группа коллекций		
10.6.4. Просмотр укороченного набора метаданных		
10.6.4.1. Отдельный электронный документ		
10.6.4.2. Коллекция, группа документов		
10.6.4.3. Группа коллекций		
10.6.5. Просмотр полного набора метаданных		
10.6.5.1. Отдельный электронный документ		
10.6.5.2. Коллекция, группа документов		
10.6.5.3. Группа коллекций		
10.6.6. Возможность исполнения пакетных операций импорта и экспорта метаданных		
10.6.7. Извлечение и просмотр первичного ресурса (электронного документа)		
10.6.8. Возможность исполнения пакетных операций импорта и экспорта данных электронных документов		
10.7. Поддержка групповых привилегий		
10.8. Поддержка ролевых групп		5.4
10.9. Поддержка ACL (Access Control List)		
11. Интероперабельность		
11.1. Поддерживается импорт данных	12.3	2.1.3
11.2. Поддерживается экспорт данных	12.5	2.1.5
11.3. Поддерживается обмен данными с другими системами		
11.4. Поддерживается импорт метаданных	12.4	2.1.4
11.5. Поддерживается экспорт метаданных	14.4	
11.6. Поддерживается обмен метаданными с другими системами		
11.7. Пользовательские интерфейсы допускают возможность работы с другими ЭБ и информационными системами		
11.7.1. Доступ по фиксированной URL (гиперссылка)	12.2	

11.7.2. Доступ к внешним справочникам при вводе информации		
11.7.3. Доступ к каталогам других ЭБ		
11.7.4. Прозрачный доступ ресурсам других ЭБ		
11.8. Поддержка сетевых протоколов доступа к ресурсам ЭБ		
11.8.1. HTTP		
11.8.2. Z39.50		
11.8.3. FTP		
11.8.4. LDAP		
11.8.5. Другие		
12. Сбор статистики	11	
12.1. Присутствует встроенная подсистема сбора информации	11.1	6.1
12.2. ЭБ использует системные механизмы сбора информации		
12.3. ЭБ использует внешние подсистемы сбора информации		
12.4. ЭБ имеет встроенные средства обработки статистической информации	11.1	6.1
12.4.1. Поддерживаются интерфейсы для просмотра информации		
12.4.1.1. Консольные текстовые		
12.4.1.2. Специальные графические		
12.4.1.3. Просмотр через WEB		
12.4.2. Поддерживается подсистема для анализа информации и получения отчетов	11.2	6.2
12.5. ЭБ имеет средства событийного управления		
13. Реализация, лицензирование и поддержка		
13.1. Название пакета		
13.2. Авторы пакета		
13.3. Место разработки		
13.4. Тип лицензирования (Open source license)	1.3	

13.5. Номер версии	1.5	
13.6. Дата выхода версии	1.4	
13.7. Инсталляционная база	6	
13.7.1. Количество инсталляций	6.1	
13.7.2. Географический охват инсталляций	6.2	
13.8. Процедура инсталляции	7	
13.8.1. Наличие программы (скрипта) автоматической инсталляции	7.1	
13.8.2. Наличие программы (скрипта) обновления операционной системы	7.2	
13.8.3. Возможность обновления системы без изменения пользовательских настроек	7.3	
13.9. Требования к обслуживающему персоналу	5	
13.9.1. Системный администратор UNIX	5.1	
13.9.2. Системный администратор Windows		
13.9.3. Java программист	5.2	
13.9.4. PERL программист	5.3	
13.9.5. Python программист	5.4	
13.9.6. .NET/ASP программист		
13.9.7. PHP программист		
13.9.8. SQL программист		
13.10. Системная поддержка	22	12
13.10.1. Наличие документации	22.1	12.1
13.10.2. Наличие сервиса обновления версий		
13.10.2.1. Обновления через интернет		
13.10.2.2. Пакетное обновление		
13.10.3. Наличие сервера списков рассылки сообщений (Listserv), наличие форума	22.2	12.2
13.10.4. Поддержка запросов к разработчикам	22.3	12.3

13.10.5. Формальная поддержка (Help desk)	22.4	12.4
14. Стандарты		
14.1. Соответствие OAI-RMN, номер версии	1.1	8.1
14.2. Соответствие Z39.50	1.2	8.2
14.3.		
14.4.		