

СОЗДАНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ УПРАВЛЕНИЯ ИНФОРМАЦИОННЫМИ РЕСУРСАМИ WEB ПОРТАЛА ВИРТУАЛЬНОЙ БИБЛИОТЕКИ

Информационные ресурсы виртуальных библиотек находятся в постоянном развитии. Проблема осложняется тем, что постоянно растет количество людей, в прямые служебные обязанности которых входит обслуживание процессов публикации информации, растет объем ресурсов, увеличивается номенклатура ресурсов, представленных в виртуальных библиотеках.

Вначале это было под силу отдельным энтузиастам, в основном программистам, которым не составляло труда использовать язык разметки (HTML) для форматирования документов и файловую структуру – для их размещения. Затем появились системы для визуального редактирования документов и поддержки структуры статических (представляющих собой набор HTML-страниц) Web-сайтов: FrontPage, DreamViewer, HomeSite и др. С помощью этих систем можно легко создавать и модифицировать такие сайты, не обладая специальной квалификацией и не вдаваясь в тонкости HTML. Но сейчас когда Internet/Intranet/Extranet технологии являются основой корпоративных и межкорпоративных информационных систем подобный подход по целому ряду причин перестает удовлетворять растущие потребности и актуальной становится разработка так называемых систем Web-публишинга, или систем управления информационными ресурсами порталов.

Цель создания портала – обеспечение информацией и возможность активной коммуникации между библиотекой, партнерами, которые будут участвовать в создании портала, и широким кругом потребителей информации. Портал – это набор программных средств и информационных ресурсов, который обеспечивает персонифицированный и настраиваемый интерфейс, дающий возможность людям находить, и взаимодействовать с другими людьми, а также находить и использовать приложения и информацию в соответствии со своими интересами.

Наиболее критичной характеристикой портала является персонификация для конкретного пользователя. Независимо от того, предназначено клиентское место портала для индивидуального пользователя или для сообщества, портал должен настраивать свой внешний вид, содержание и интерфейс приложений для каждого из них индивидуально. Эта настройка может быть основана на отслеживании индивидуальных привычек пользователя, области его интереса или просто на той роли, которую пользователь выполняет в сообществе.

Вторая по значимости характеристика портала – это способность организации пользовательской рабочей среды, с целью устранения

информационной перегрузки. Пользователи могут иметь доступ к информации в консолидированном виде, а организации хотят могут обеспечить простые средства проектирования дизайна корпоративного рабочего клиентского места в виде, который наилучшим образом организует информацию для сотрудников. Распределение ресурсов обеспечивает разделение некоторых возможностей портала на уровни, доступные разным категориям пользователей. Например, к некоторым уровням могут иметь доступ только администраторы. Пользователи имеют высокий уровень безопасности с минимальными накладными расходами и неудобствами, и распределение ресурсов обеспечивает то, что служебные механизмы находятся вне поля зрения конечного пользователя.

Способность портала обеспечить доступ и отображение информации из множественных гетерогенных источников данных также является критичным фактором.

Портал, помимо всего прочего, должен предоставлять конкретному пользователю специфическую и нужную ему информацию. В результате, пользователи будут тратить меньше времени на просмотр несущественной информации или на переключение между разными интерфейсами для выполнения различных задач. Это должно обеспечивать принцип, в соответствии с которым пользователи обнаруживают информацию, а не находятся в постоянных поисках и режиме просмотра.

Программно-технологическая платформа, выбранная для построения и поддержки портала должна включать и обеспечивать базовый набор служб-компонент. При выполнении этого требования построение портала с любой выбранной архитектурой будет происходить с минимальными затратами сил и времени, за счет использования готового набора служб-компонент.

Базовый набор служб-компонент программно-технологической платформы включает пять основных компонентов:

- Компонент служб представления
- Компонент пользовательских служб
- Компонент управления информацией
- Компонент адаптеров портала
- Компонент Web-инфраструктуры

В качестве основных отличительных технологических принципов предлагаемых решений можно выделить следующие:

- распределенное объектное хранилище данных с динамическим поиском по опубликованным материалам с учетом метаданных,
- гарантия целостности документов,
- открытое авторство,
- возможность семантического связывания документов,
- наличие виртуальной редакции,
- поддержка процесса разработки и отладки компонент системы без ее остановки,
- высокая степень масштабируемости и качественной расширяемости,

- надежность и устойчивость при больших нагрузках при минимальных требованиях к оборудованию и программному обеспечению.

Важным принципиальным моментом при создании системы является ориентация на максимальное использование стандартов и промышленного программного обеспечения. Наиболее перспективным средством в создании при построении такого класса систем прямо считать использование XML.

Для обеспечения распределенности, производительности, расширяемости и масштабируемости в основу архитектуры положена была разработана многоуровневая и многокомпонентная архитектура.

Два базовых понятия – распределенного информационного объекта и трехуровневой архитектуры с общей шиной определяют принципы построения портала.

Под распределенным объектом мы понимаем целостную информационную единицу, однозначно идентифицируемую в системе, и имеющую вполне определенную именно для данного экземпляра объекта внутреннюю структуру. Распределенность означает возможность раздельного хранения частей одного объекта на разных серверах и возможность построения одновременного запроса к разным объектам: находящимся на разных серверах.

Трехуровневая архитектура с общей шиной означает наличие промежуточного слоя (middleware) между интерфейсами ввода-вывода системы и уровнем хранилищ информации.

Такая архитектура построения информационной системы обладает многочисленными преимуществами перед обычным подходом клиент-сервер в тех случаях, когда требуется масштабируемость на (а) большие объемы разнородной информации, (б) высокие нагрузки, а также (в) обеспечение распределенного доступа и (г) высокая гибкость в модификации и развитии системы. Такая схема поддерживается консорциумом OMG (www.omg.org), в который входят практически все ведущие фирмы-производители программного обеспечения и около 30 крупнейших университетов. По технологиям OMG построены такие крупнейшие информационные системы как CNN (www.cnn.com), DPA.

Центральной частью системы является информационное хранилище, в котором находятся все информационные ресурсы, включая как сами документы, так и шаблоны страниц, а также административную информацию.

Рассмотрим жизненный цикл информационных ресурсов системы. Он состоит из трех основных этапов: создание, управление и распространение.

Этап первый – создание ресурса

Данный этап ложится на плечи авторов, которые занимаются разработкой графической, текстовой информации, шаблонов HTML, кода back-end. Для этого они могут применять такие инструменты, как редакторы HTML, текстовые процессоры, визуальные редакторы и другие средства создания объектов.

Предполагается, что авторы могут использовать для этих целей

инструментальные средства, которые они уже имеют и привыкли работать с ними.

Система должна только обеспечить загрузку созданного информационного ресурса в хранилище системы.

При этом автору должна быть предоставлена возможность задать список мета атрибутов для определенного информационного ресурса.

Набор мета атрибутов базируется на стандарте, предложенном Dublin Core Metadata Group. Для определенных групп пользователей (тех, кто имеет достаточный опыт работы с компьютером) в системе может быть использован встроенный редактор HTML страниц.

Этап второй – управление информационными ресурсами

Управление информационными ресурсами – важнейший этап жизненного цикла ресурсов. Сегодня на первый план выдвигаются задачи поддержки актуальности и точности ресурса (необходимость самой свежей информации), поэтому требуется эффективное управление ресурсами.

На данном этапе пользователи в соответствии с назначенными им правами доступа к ресурсам участвуют в процессах создания и публикации нового информационного ресурса.

Управление процессам начинается с момента, когда автор ресурса, выполнив назначенную ему работу, направляет его в хранилище (репозиторий). Передача содержимого протоколируется и следующий вовлеченный в деловой процесс исполнитель автоматически извещается. При этом ведется аудит событий.

Таким образом, пользователи не будучи техническими специалистами, самостоятельно могут участвовать в процессах обновления (актуализации) ресурсов.

Этап третий – распространение ресурсов

Для динамического распространения информации система управления контентом должна обеспечивать компоновку ее элементов на базе прямой и косвенной информации о потребителях информации.

Информационное обеспечение организовано с учетом:

- расширения сфер применения информационных ресурсов;
- развития информационных ресурсов на основе системы классификаторов;
- создания единого информационного пространства сферы образования;

Проектируемая система должна обеспечить единообразие создаваемых информационных ресурсов. Каждый ресурс в системе должен однозначно идентифицироваться и иметь определенную структуру, связанную с его типом.

В системе должна существовать библиотека шаблонов для всех типов ресурсов. Модель шаблона строится на основе стандарта XML.

Каждому типу ресурсов в системе должен соответствовать определенный набор процедур, которые регламентирует базовые действия над данным ресурсом (например, регистрация ресурса в репозитории, дублирование ресурса, экспорт ресурса, и т.д.)

Для однозначного описания данных, сокращения объемов при их хранении, передаче, обеспечения эффективного поиска и идентификации данных, удобства первоначального ввода информации используются классификация и кодирование информации.

Классификация и кодирование обеспечивают:

единообразие в обозначении признаков и характеристик информационных объектов;

упорядочение и группировку информации по определенным признакам;

приведение информации к форме, удобной для обработки с помощью вычислительных средств.

Информационная совместимость базируется на использовании технологии XML.

Основные информационные ресурсы портала должны храниться в реляционных БД данные в которых объединены единым смысловым содержанием.

При выборе базовых программных средств и технологий разработки в приоритет отдается свободно распространяемым программам в соответствии с общепринятыми в научном мире стандартами, что обеспечивает расширяемость при низкой стоимости и высокой надежности, и позволяет активно участвовать в разработке этих продуктов при необходимости.

К настоящему времени разработано программное обеспечение серверного ядра, обеспечивающего основные функции, такие как работа с базой данных, генерация сайта в соответствии со структурой логического дерева, создание такого дерева, поиск и администрирование системы. Для расширения возможностей ядра системы существует набор модулей, позволяющих реализовать дополнительные функции по управлению контентом портала.

Основные возможности ядра системы :

- Поддержка нескольких независимых web-сайтов в рамках одной корпоративной системы; Менеджер навигационного меню – позволяет создавать меню любой сложности
- Поддержка различных типов страниц:
 - Динамические страницы – создаются с помощью встроенного html-редактора
 - Статические страницы – поддерживается использование страниц, написанных во внешнем html-редакторе
 - Страницы-шаблоны – предназначены для упрощения процедуры заполнения данных определенного формата, например, описания товаров и т.п. (с поддержкой XML)
 - "Последовательности страниц" – набор страниц сменяющих друг друга по "цепочке" в определённые моменты времени.
- Встроенный html-редактор. Позволяет создавать и редактировать страницы, используя ограниченный набор элементов – текст, картинка, ссылка. Редактор шаблонов отображения сайта – позволяет выбирать

один из нескольких predetermined шаблонов, а также модифицировать их – выбирать тип шрифта, цвета, кнопки и т.п.

- Возможность указания времени публикации различных элементов системы (страницы, пункты навигационного меню, баннеры и т.д.)
- Системы "быстрых" ссылок внутри сайта.
- Гибкая система назначения прав доступа;
- Полная поддержка нескольких языков как в системе администрирования, так и в режиме on-line.
- Полнотекстовый поиск. В системе предусмотрена возможность поиска по любым типам страниц, имеющихся в системе.
- Stop-word-filter. Позволяет в автоматическом и ручном режиме запускать поиск "запрещенных" слов и выражений, результат поиска формируется в виде списка найденных страниц.