

КОНЦЕПЦИЯ СОЗДАНИЯ ПОРТАЛЬНОЙ СИСТЕМЫ В СЕТИ ИНТЕРНЕТ ДЛЯ ИНФОРМАЦИОННОГО ОБСЛУЖИВАНИЯ АПК

Г.В. Попков

В статье рассматриваются программные средства для создания многопрофильных порталных систем для их использования в сельской местности. Предложенные технологии позволят наиболее эффективно работать с массивами информации, расположенными на портале различным категориям интернет пользователей.

Ключевые слова: информационные технологии, интернет, СУБД, php, drupal.org, портал.

Федеральная целевая программа «Социальное развитие села до 2012 г.» указывает, что для достижения максимальной эффективности продовольственной программы и повышения общего уровня жизни на селе необходимо пересмотреть отношение к сельскому хозяйству в стране в кратчайшие сроки. Для этого предусмотрен комплекс мер поддержки отечественного сельскохозяйственного производителя. В частности, обозначена потребность разработки информационно-телекоммуникационной системы, предназначенной для оперативного представления баз знаний по всем отраслям сельского хозяйства. В числе заинтересованных пользователей – производители и участники рынка сельскохозяйственной продукции, ученые и специалисты сельскохозяйственного профиля, фермеры, широкие слои населения, нуждающиеся в получении оперативной и актуальной информации, пригодной для использования в практической деятельности.

Подобная система должна базироваться на использовании современных сетевых и информационных технологий, прежде всего Web-технологий, и представлять собой единое Интернет-пространство знаний, необходимых для производства сельскохозяйственной продукции и включающих как информацию сельскохозяйственной направленности, так и сведения из смежных отраслей, например, о средствах производства, рынках сбыта, трудовых ресурсах, транспортной системе, перерабатывающих мощностях, обучающих центрах, социальных и культурных организациях.

Известно, что количество информации, производимой человечеством, удваивается каждые три года. За год вырабатывается свыше пяти экзбайтов (миллионов терабайт)

информации. При этом развитие информационных технологий привело к тому, что в развитых странах большая часть информации (более 90 %) хранится в электронном виде. На долю фото, видео, кино, бумажных и прочих носителей приходится менее 7 % информации [1].

На данном этапе количество зарегистрированных доменных имен превысило 2,5 млн., количество Веб-сайтов – 230 млн. При этом даже самые крупные информационно-поисковые системы, такие как Google, в процессе индексации охватывают не более 30-50 % доступных ресурсов. Доступные ресурсы, в свою очередь, – это лишь видимая часть огромного массива информации. Объем ресурсов, не поддающихся классическим системам поиска по индексации, очень велик. Это, прежде всего, информация в различных базах данных, файлах специфических форматов, нераспознаваемых обычно информационно-поисковыми системами. Таким образом, традиционные средства поиска информационных ресурсов не справляются с задачей нахождения большей части информации.

Для поиска информации в Интернете, чаще всего представленной в формате HTML, используются специальные средства – сетевые информационно-поисковые системы. Оказалось, что СУБД, широко применяющиеся для прикладных задач, непригодны для поиска информации в Интернете. Не оправдали надежд в данной сфере и системы искусственного интеллекта. В настоящее время, как правило, эффективно работают системы, опирающиеся на использование частотно-лингвистических и эвристических методов.

До появления информационно-поисковых систем средствами навигации в Интернете были каталоги, организованные по образцу и подобию библиотечных, не требующих от

пользователей особых навыков. По найденной в каталоге гиперссылке пользователь попадал непосредственно на страницу с требуемой информацией.

Эволюция поиска информации в Интернете привела к появлению понятия портала как интегрированного web-сайта, организованного в виде системного многоуровневого объединения различных информационных ресурсов и сервисов, ориентированных на определенную целевую аудиторию, и играющего роль навигационной системы.

В настоящее время РАСХН приступила к разработке собственного портала. Эта работа станет первым этапом построения ИТС «Электронное сельское хозяйство России», которая будет иметь трехуровневую структуру для различных групп пользователей: институтов, отделений и Президиума РАСХН. Цель разработки – создание эффективной системы доведения результатов научных исследований до потребителей посредством Web-представительства РАСХН и его организаций.

Идея создания многофункциональных порталов в сети Интернет возникла не сегодня. Компании-разработчики заинтересованы в предоставлении удобного доступа к различным приложениям, источникам информации, базам данных, представляющим интерес для широкого круга пользователей сети Интернет. Однако еще пять-шесть лет назад не существовало технологической возможности решения задач такого класса. Классический компьютерный интерфейс не мог полностью и гибко интегрировать и персонализировать информацию пользователей Интернет-порталов.

Для обеспечения доступа к разнообразной информации требуется гибкая технологическая модель. Этой моделью стали новые веб-технологии. Портальные продукты появились на рынке с середины 1990-х гг. Они предназначались для постройки систем на базе корпоративных интрасетей и довольно быстро были приняты пользователями. Возможно, решающую роль в этом процессе сыграла тогдашняя мода на все, что связано с Интернетом.

Мода на Интернет до некоторой степени исчерпала себя в том виде, в каком пользователи привыкли его видеть, однако в последнее время рынок порталов переживает настоящий бум. В последние годы в составе порталов появились весьма продвинутые функции управления разнообразной инфор-

мацией, такие как управление контентом; документооборот и управление знаниями; социальные сети. Кроме того, развитие в последние годы языка XML, ставшего языком интеграции, позволило упростить и ускорить процесс создания порталов, а также реализовать обмен информацией в унифицированном виде между различными программными платформами.

Информационные порталы должны играть роль центрального шлюза для доступа к огромному объему информации, сосредоточенной на одном ресурсе в сети Интернет. Неоспоримым преимуществом порталного сервера является центральная точка контроля безопасности и удобный доступ к сервисам, а также привычный и удобный дизайн.

Для поддержки этих процессов информационный портал включает множество компонентов и различных категорий ПО: архивное ПО и приложения для управления документами; ПО для координирования совместной работы; поисковые системы; каталожные услуги и средства контроля за авторизацией и доступом; средства интеграции приложений, процессов и данных. Информационная система включает функции, позволяющие создавать и управлять совокупностями данных, как-то: получение информации из разнообразных источников, таких как поисковые Web-роботы; индексирование материалов; организация материалов в структуру, соответствующую потребностям и структуре портала; поддержка версий и архивов; предоставление постоянного доступа к информации с помощью продвинутых, простых в использовании поисковых и просмотрных инструментов. Специальные программные продукты, включенные в приложения, позволяют поддерживать инфраструктуру управления знаниями, строить и создавать коллекции структурированных и неструктурированных данных, обеспечивать доступ к контенту, в частности, с помощью поиска и анализа данных.

По мере развития технологий, лежащих в основе порталов, появляются новые методы использования порталов большим количеством пользователей. Возможности порталов обычно развиваются по стандартной модели: сначала они предоставляют доступ к неинтегрированным данным и контенту, затем начинают объединять людей и различную информацию, потом унифицируют доступ к структурированным и неструктурированным данным и, наконец, интегрируют при-

КОНЦЕПЦИЯ СОЗДАНИЯ ПОРТАЛЬНОЙ СИСТЕМЫ В СЕТИ ИНТЕРНЕТ ДЛЯ ИНФОРМАЦИОННОГО ОБСЛУЖИВАНИЯ АПК

ложения для поддержки бизнес-процессов, информационной и развлекательной части портала.

Значительный и устойчивый рост рынка порталов вызвал цепную реакцию. Наиболее дальновидные производители ПО почувствовали рост покупательского спроса и не замедлили наполнить рынок новыми продуктами.

Особенностью этой бизнес-сферы является наличие множества компаний, пришедших из других областей. Первыми на этом рынке появились компании, специализирующиеся на управлении документами и поставке хранилищ данных и СУБД, владеющие технологиями обработки, хранения и извлечения структурированных данных. Позднее пришли компании, работавшие в области интернет-коммерции, сильными сторонами которых являются дизайн и функции поддержки транзакций, а также навыки работы с HTML, PHP, XML и взаимодействием сетей. Еще позднее присоединились компании, специализирующиеся на полнотекстовом поиске и создании структурированных архивов полнотекстовых данных, в которых можно проводить продвинутый текстовый поиск и осуществлять просеивание. Проявили интерес к информационным порталам и провайдеры контента в Интернете, такие как Yahoo и AltaVista Business Services. Их конкурентное преимущество – умение организовывать поиск и категоризацию данных из интернета. Последними на этот рынок вышли крупные поставщики программного обеспечения, такие как IBM, SAP, BEA, Sun, Oracle и Microsoft. Данные компании предлагают свои решения CMS-систем (Content Management Systems) для создания многофункциональных web-порталов.

На данном этапе развития IT-рынка огромную конкуренцию мощным вендорам составляет один противник, в борьбе с которым едины все поставщики portalного ПО, – собственные разработки. Это серьезный противник. Опрос 49 крупнейших корпораций, проведенный компанией Forrester в 2009 г., выявил, что около половины заказчиков намерены приобрести порталы, оставшаяся часть предпочитает создавать их самостоятельно.

Достаточно остро стоит проблема безопасности portalных систем, в частности, по-прежнему не решены вопросы аутентификации пользователей при едином входе в систему (SSO, single sign-on). Кроме того, стан-

дарты безопасности для порталов, такие как Security Assertion Markup Language (SAML) на базе XML, продолжают изменяться. В настоящее время активно внедряется стандарт SAML, разработанный технической группой Oasis. Он должен упростить процесс обмена информацией об идентификации и управление доступом между поставщиками SSO-систем и порталом.

Одна из основных технологических проблем заключается в том, что для создания и поддержания портала, связанного с большим объемом функционала, требуются большие ресурсы, способные поддерживать процессы извлечения данных. Основные методы решения задачи обеспечения доступа к информации для пользователей портала разрабатываются специалистами в области безопасности и связаны с распространением портлетов. Портлеты, или portalные апплеты, – это своеобразные интерфейсы для подключения корпоративных приложений к portalу, элементы программ, исполняющие роль плагина для portalной структуры.

Для решения этой проблемы разрабатываются стандарты Java Portlet API и Web Services for Remote Portals (WSRP). Java Portlet API обеспечивает взаимодействие между портлетами и порталами. Все портлеты, написанные в этом стандарте, будут отделены от окружающей portalной структуры и смогут работать со всеми поддерживающими его portalными серверами, аналогично работе сервлетов на различных серверах приложений. Как и при работе с прочими Java-апплетами, ограничения в большей степени связаны с внутренними установками, чем с возможностями продукта, и позволяют сделать на portalе все необходимое.

Еще одна глобальная проблема связана с обеспечением доступа в сеть Интернет в сельских районах, где зачастую слабо развита инфраструктура телефонной связи, а в некоторых селах до сих пор есть только почтовое отделение и телеграф. Говорить о выделенных Интернет-линиях в данном случае не приходится. В настоящее время 30 % населенных пунктов в сельской местности имеют только групповые пункты доступа к телефонной связи. Как правило, это отделения почты России либо здания администраций районов. В этой ситуации проблема доступа к ресурсам портала в большей степени зависит от организации абонентского доступа для населения сельских районов.

В подобных случаях наиболее доступным видом доступа оказывается коммутируемый Dial-Up доступ через телефонную сеть села, деревни, сельского района. Следовательно, разработчики портала должны предусмотреть применение технологий, максимально быстро работающих по медленным телефонным каналам ТЧ ТФОП России.

За последние годы сформировались негласные требования относительно разделов, присутствие которых обязательно на многофункциональном портале. Несоблюдение этих требований может существенно снизить посещаемость сайта и даже негативно повлиять на репутацию ресурса, несмотря на то что эти требования можно считать, скорее, правилами хорошего тона, нежели стандартами.

К указанным разделам обычно относят: новости; карту сайта или инструмент поиска внутри сайта; средства обратной связи с посетителями (как минимум, адрес электронной почты или форму для комментария непосредственно на сайте); средства навигации, позволяющие посетителю понять, в каком разделе сайта он находится (заголовки типа «Новости > Архив новостей > 2005 > Январь > 25 января»); координаты компании (адрес, телефон, схема проезда); правила и гарантии использования личной информации посетителей, заполняющих различные анкеты и опросные листы (Privacy policy); для сайтов, принимающих платежи по кредитным картам, — сведения о безопасности данной операции. Реже встречаются инструкции и фрагменты HTML-кода для упрощения размещения ссылок на собственный сайт для владельцев других сайтов. В последнее время раздел новостей все чаще содержит ссылку на новости в формате RSS, реже — на версию для WAP, КПК. За последние годы посетители сайтов привыкли к общепринятому расположению элементов интерфейса (меню слева, ссылки справа, еще одно меню сверху), не говоря уже о том, что любой сайт должен корректно отображаться в различных видах и версиях браузеров. Как правило, посетители придают гораздо большее значение наличию на сайте общепринятых элементов в общепринятых местах, нежели дизайну и художественному оформлению. Таким образом, значение веб-дизайна постепенно отходит на второй план.

Не секрет, что одна из главных тенденций в области ИТ-угроз последнего десятилетия — это увеличение числа вредоносных про-

грамм, использующих Интернет в качестве канала распространения. По данным отчета Symantec об интернет-угрозах "Internet security threat report", в последнее время наблюдается рост числа коммерческих вредоносных программ и технологий, разрабатываемых для обогащения и преследующих такие цели как онлайн-мошенничество (фишинг, фарминг) или кража конфиденциальной информации.

Указанные проблемы приводят к более серьезным, нежели десять лет назад, требованиям к защищенности web-приложений и технологий их создания, а также к средствам защиты данных в готовом ПО, используемом при создании сложных web-решений.

На примере разработки проекта справочно-информационной системы можно продемонстрировать пути решения нескольких проблем. Во-первых, это задача выбора оптимальной CMS, СУБД, операционной системы, серверной платформы. Данный проект реализован на CMS open-source www.drupal.org. В качестве языка программирования использована технология PHP (Personal Home Pages), которая использует CGI-приложения, интерпретирующие внедренный в HTML-страницу код на скриптовом языке, а также AJAX-технология (Asynchronous JavaScript and XML) создания web-приложений. Использование AJAX-технологии позволяет не производить полную перезагрузку web-страницы в ответ на каждый клик пользователя, а догружать с web-сервера только нужные пользователю данные. AJAX-технология реализуется с помощью генерации на сервере DHTML и соответствующего кода на скриптовых языках для выполнения в клиентском приложении с помощью большинства доступных серверных технологий (PHP, ASP .NET 2.0, Ruby on Rails и др.). В качестве СУБД используется MySQL. Для организации платежей было введено понятие "персонального виртуального офиса" и персонального кошелька с использованием протокола XML. Таким образом, решена задача сбора платежей пользователей через системы платежных терминалов и мультикасс (cash-in).

В большинстве случаев средства управления информационным наполнением web-сайтов и web-порталов позволяют осуществлять централизованное управление хранением данных (чаще всего это фрагменты документов); отделять содержание от представления (от дизайна сайта); автоматизировать

КОНЦЕПЦИЯ СОЗДАНИЯ ПОРТАЛЬНОЙ СИСТЕМЫ В СЕТИ ИНТЕРНЕТ ДЛЯ ИНФОРМАЦИОННОГО ОБСЛУЖИВАНИЯ АПК

управление жизненным циклом данных; применять перечисленные средства пользователями для решения различных задач.

В состав CMS-средств обычно входят средства управления документами, такие как блокировка при изъятии для редактирования; контроль версий; защищенный доступ; организация библиотек и каталогизации; средства редактирования фрагментов документов и превращения документов на бумажных носителях в электронной формат; управление данными и архивами; возможность коллективной работы над документами; организация документооборота.

Средства управления информационным наполнением, инструменты создания порталов, как правило, позволяют предоставлять персонализированный доступ к тому или иному виду информации, организовывать эффективную работу с данными, определять период их актуальности, осуществлять индексацию и категоризацию.

Спектр технологий создания веб-решений достаточно широк, от написания собственных приложений, основанных на применении скриптовых языков и CGI-кода, до внедрения готовых средств создания порталов и управления информационным наполнением. Последний способ создания корпоративных сайтов, возможно, станет доминирующим в ближайшее время за счет низкой стоимости внедрения и сопровождения.

Возможные разделы портала:

- новостные ленты,
- торговая площадка,
- каталог продукции,
- каталог компаний,
- справочная информация о районе,
- административная информация,
- он-лайн приемные,

- история края,
- помощник фермера,
- тематические разделы, посвященные сельскохозяйственной тематике,
- развлечения,
- форумы, элементы социальных сетей,
- система профилей пользователей,
- личные странички пользователей.

Создание интернет-ресурсов, ориентированных в первую очередь на жителей сельских районов страны, не только актуальная, но и технически реализуемая задача. Создание подобных ресурсов поможет повысить уровень информационной просвещенности жителей села, автоматизировать управление сельхозпроизводством, развить товарооборот в сельскохозяйственных районах. Важным социальным аспектом обсуждаемого процесса станет повышение уровня образования молодежи, участвующей в наполнении контента подобных ресурсов. Работа над интернет-ресурсами для работников сельского хозяйства будет способствовать выполнению программы "Электронная Россия".

СПИСОК ЛИТЕРАТУРЫ

1. Гольдштейн, Б.С. Сети связи / Б.С. Гольдштейн, Н.А. Соколов, Г.Г. Яновский. – СПб.: БХВ-Петербург, 2010. – 400 с.
2. Ландэ, Д.В. Поиск знаний в INTERNET / Д.В. Ландэ. – М.: Диалектика, 2005. – 272 с.
3. Меденников, В.И. Моделирование оптимальных информационных систем в сельском хозяйстве / В.И. Меденников, С.Г. Сальников // Вестник МГАУ им. В.П. Горячкина. – 2005. – № 3.

Попков Г.В., к.т.н., науч. сотр., Институт вычислительной математики и математической геофизики СО РАН, тел. 8(383) 330-96-43, E-mail: glebpopkov@rambler.ru.