

# РАЗРАБОТКА, РЕАЛИЗАЦИЯ И СОПРОВОЖДЕНИЕ ВЕБ-САЙТА НАУЧНОЙ ОРГАНИЗАЦИИ

М. Ю. Колодин

Санкт-Петербургский институт информатики и автоматизации РАН  
199178, Санкт-Петербург, 14-я линия В.О., д.39  
myke@iias.spb.su, <http://myke.da.ru>

---

УДК 681.3

М. Ю. Колодин. Разработка, реализация и сопровождение веб-сайта научной организации // Труды СПИИРАН. Вып. 1, т. 3. — СПб: СПИИРАН, 2003.

**Аннотация.** В работе рассматриваются основные способы организации веб-сервера активной научной организации на примере веб-сервера группы ИГИТО СПИИРАН. Указаны базовые задачи сервера, критерии выбора архитектуры системы и инструментария и описано поэтапное решение поставленной задачи, при этом указаны основные трудности и пути их преодоления, перспективы развития веб-сервера и работы внутри организации, с ним связанной. — Библ. 9 назв.

UDC 681.3

М. Y. Kolodin. Design, implementation and maintenance of scientific organization's web-site // SPIIRAS Proceedings. Issue 1, v. 3. — SPb: SPIIRAS, 2003.

**Abstract.** This paper discusses main ways of organization of a web-server for an active scientific organization, basing on example of web-server of SPIIRAS RGITE group. The principal tasks of the server and criteria for selection of system architecture and tools are stated, and step-by-step solution of the given task is described, with special attention to the main difficulties and ways of their passing and perspectives of further development of web-server and work within organization, associated with it. — Bibl. 9 items.

---

## 1. Общие положения

Наличие должного интернет-представительства — практическая необходимость для любой современной организации, тем более — для научной. Основное назначение веб-сайта научного учреждения или его подразделения (далее — организации) — это сетевое, т.е. публичное, представление результатов научной деятельности. Таковое представление должно, как минимум, включать:

- сведения о наименовании и статусе организации,
- общее назначение и направления деятельности организации,
- сведения о составе организации, направлениях деятельности каждого из ее сотрудников,
- постановки задач и результаты работы,
- сведения о публикациях и конференциях.

В расширенном варианте веб-сайт должен служить средством не только пассивной презентации общих данных и результатов работы организации, но и активно содействовать ее работе: структурированию и активизации внутренней жизни и развитию взаимоотношений с внешним миром.

Рассмотрим все указанные и сопутствующие вопросы на примере веб-сайта научно-исследовательской группы информационных технологий в образовании (ИГИТО) Санкт-Петербургского института информатики и автоматизации Российской Академии наук (СПИИРАН).

## 2. Основные требования

Прежде всего, определим минимальные требования к веб-сайту и его организации. Веб-сайт есть совокупность стандартным образом оформленной и структурированной информации, зарегистрированная и доступная в сети. Физически это набор веб-страниц, стилей, исходной информации, на основе которой формируется динамическое содержание веб-сайта, и программ для преобразования и представления такой информации. В простейшем случае веб-сайт есть совокупность веб-страниц. Для формирования содержания веб-сайта существуют специальные средства, которые будут рассмотрены далее. Все начинается с принятия решения о создании веб-сайта, т. е. организационно нужно решить, что именно будет выполняться, с какими целями, кто и как этим будет заниматься, в какие сроки и с какими затратами ресурсов, материальных и человеческих, и каковы сроки, и кто является ответственным за выполнение работ. Прежде всего, нужно определить хостинг, т. е. размещение веб-сайта в сети, и его регистрацию в соответствующих сетевых (или государственных) органах. Хостинг может быть внешним (платным либо бесплатным, например, в обмен на размещение рекламы) либо внутренним (т. е. на внутреннем компьютере организации, имеющем постоянный IP-адрес и круглосуточный доступ в интернет, желательно, с достаточно широким каналом). В нашем случае мы разместили веб-сайт группы, так же как и веб-сайт института, на внутреннем компьютере в лаборатории, обеспечив достаточный канал для того, чтобы потенциальные пользователи как внутри, так и вне лаборатории могли беспрепятственно и круглосуточно получать и передавать нужную им информацию. Регистрация также может быть платной либо бесплатной, для последней, как правило, есть особые домены, и требуется право на включение в них. В нашем случае, так как институт имеет лицензию на образовательную деятельность, мы зарегистрировали имя для нашего веб-сайта как [www.computer.edu.ru](http://www.computer.edu.ru) и определили для него фиксированный IP-адрес.

Для работы над веб-сайтом требуется участие нескольких ответственных, обязанности которых в простейшем случае может совмещать один человек. Это веб-мастер, дизайнер, художник, контент-автор, администратор сервера, администратор базы (баз) данных, связанных с сервером, руководитель организации. В нашем случае первоначально все функции, кроме последней, выполнял один человек, в дальнейшем, по мере усложнения компьютерной системы, эти функции разделились.

Очень важны вопросы безопасности веб-сайта. Опасность внешнего вторжения или внутреннего повреждения веб-сайта и связанной с ним вычислительной системы, умышленных или по неосторожности, очень высока. До сих пор не существует универсальных средств защиты, и необходимо применять аппаратные, программные и организационные средства защиты, включающие паролирование, программную проверку запросов, использование защитных программных комплексов, и многое другое.

Разработка, формирование и поддержка содержимого веб-сайта подробно и поэтапно описаны ниже, включая вопросы восстановления при сбоях. Пока же отметим, что для кодирования использовались всевозможные средства от текстовых редакторов до программного генерирования текста страниц и отдельных программ.

Реклама — не столь незначительный вопрос, даже для научных организаций, как это могло бы показаться. Интернет содержит миллионы веб-сайтов, и найти среди них нужный порой оказывается чрезвычайно непросто. Нужно

обеспечить возможность для пользователей найти наш веб-сайт. Для этого можно использовать несколько механизмов. Прежде всего, это использование мета-полей “description” и “keywords” в заголовке главной страницы веб-сайта, что дает возможность поисковым серверам зарегистрировать данный веб-сайт в своих базах. Далее, можно самостоятельно зарегистрироваться на таких серверах. И очень важно навести перекрестные ссылки с дружественными или административно вышестоящими либо подчиненными или тематически близкими веб-сайтами, связаться с их веб-мастерами для периодического контакта.

При размещении информации на веб-сайте необходимо учитывать субординацию, особенно для государственных (прежде всего — военных и академических) организаций, а также вопросы авторских прав на всю размещаемую информацию, включая тексты, изображения, используемые для формирования выводимой информации программы, форматы данных и дизайн.

Последнее, что надо учесть, это вопросы закрытия веб-сайта: кто может выполнить такую операцию, каковы последствия, куда денется накопленная там информация. В нашем случае все подобные вопросы тривиальны, но в общем случае они могут возникнуть, и их стоит продумать с самого начала.

По организации данных и способу формирования показываемых веб-страниц сайты можно разделить на несколько категорий: пассивные, активные, интранет, распределенные веб-службы; поскольку с этими категориями непосредственно связаны и прямо им пропорциональны и используемые программные средства и организационные процедуры, эти категории можно трактовать и как уровень сложности веб-сайта.

### 3. Пассивные веб-сайты

Пассивные веб-сайты — простейшие для разработки, реализации и сопровождения. Их основная и практически единственная задача — отображение указанной выше информации об организации. Основным формат таких страниц — HTML, возможно, с дополнительными стилевыми таблицами в формате CSS; некоторое оживление достигается с помощью включений на Javascript. Главные рекомендации здесь: не перегружать первую страницу (как, впрочем, и дальнейшие) излишней графической информацией, дать возможность посетителю с первого взгляда определить назначение веб-сайта, характер и основные составляющие деятельности организации; для академической организации необходимы строгость дизайна, отсутствие вычурных особенностей, единый стиль всех страниц. Очень важно, чтобы был указан адрес веб-мастера, к которому можно обратиться, если возникнут вопросы. Наша первая страница была именно такой, с рисунками-ссылками на основные деятельности и подразделения группы. Обновление информации было строго ручным, по мере необходимости (появление новых сотрудников или групп, формирование годовых отчетов или появление публикаций, пр.). Для возможного восстановления информации использовалось регулярное резервное копирование на другой компьютер.

Уже на этом этапе были сформулированы основные требования к веб-сайту:

- веб-страница ИГИТО (далее — страница) должна содержать, как минимум, всю имеющуюся в группе официальную информацию,
- страница должна иметь современный вид,
- страница должна иметь удобную навигацию,

- страница должна правильно и, по возможности, одинаково показываться основными современными браузерами (MSIE, NN, Opera, Konqueror),
- страница должна быть легко масштабируемой,
- поддержка страницы должна выполняться администратором, возможно, не автором страницы,
- страница должна иметь стиль, близкий к общему стилю веб-сайта СПИИРАН.

Этот этап был успешно выполнен на ранних стадиях формирования и развития группы ИГИТО и компьютерного научно-образовательного центра. Основные использованные средства — HTML-страницы, CSS-стили.

#### 4. Активные веб-сайты

Активные веб-сайты — следующий шаг в развитии. Дело в том, что даже в небольшой организации при активизации ее деятельности информация появляется и обновляется очень часто, практически постоянно, и у веб-мастера нет возможности достаточно оперативно отслеживать изменения и вносить модификации в HTML-код. Вместо этого требуется автоматическое формирование содержимого веб-сайта на основе имеющейся в каждый момент информации. Кроме того, на этом этапе мы столкнулись и с иными требованиями. Прежде всего, нам нужно было согласовать формат и содержимое нашего веб-сайта с веб-сайтом объемлющей организации — института. Кроме того, каждый сотрудник смог (и фактически был обязан) иметь собственную веб-страницу. Далее, множество информации появлялось не в формате HTML, а в простом текстовом формате, и ручной перевод ее в HTML занял бы слишком много времени, которое тратить было тем более жаль, что информация быстро устаревала (оперативная информация о семинарах, встречах, различного вида новости, прочее). Таким образом, естественно появилась задача автоматизации процесса. Исходя из перечисленных требований и с учетом традиций веб-проектирования, был выбран механизм формирования активных страниц с использованием аппарата CGI и языка Perl; такой выбор диктовался и необходимостью переносимости веб-сайта между платформами MS Windows и UNIX (Linux).

Важно, что написанная для выполнения этой задачи программа была независимой от содержимого веб-сайта и могла применяться для динамического формирования разнообразных веб-сайтов. Указания на конкретное содержимое веб-страниц передавалось программе в виде параметров, а при последующих вызовах программа самостоятельно формировала запросы и выводимые страницы на основе текущего состояния информации веб-сайта, организованной в виде дерева каталогов. Кроме того, был реализован специальный упрощенный язык (WEP, “web engine in perl”), на котором фактически и были записаны все страницы веб-сайта; при работе программа разбирала текст на этом языке и формировала соответствующие страницы для вывода. Кроме основной программы, использовались запускаемые по расписанию программы формирования оперативной информации на основе внешних источников, например, получение новостей с различных веб-сайтов (по согласованию с оными) и построение из этой информации веб-страниц, согласованных с основным содержимым нашего веб-сайта. По аналогичным расписаниям выполняется и резервное копирование информации, причем в используемом сжатом формате (без явного HTML-форматирования) эта информация занимает намного меньше места, чем

ранее. На этом этапе ведение веб-сайта института также перешло в наше подразделение, и согласовывать работу стало удобнее и эффективнее.

Мы уточнили требования к странице: главная страница содержит название, фотографию, разделы: институт, группа, новости, подгруппы, учебные планы, архивы, ссылки на ресурсы, выходные данные, новости веб-сайта, группы, института, по специальности, в городе и в мире; учебные планы: базовые, специальные, дополнительные, общая информация по институту, ссылки на официальные страницы, о создании группы, о ее текущей деятельности, результаты работы, подгруппы, семинары и конференции, состав группы, подстраницы сотрудников, объявления, сборник "Труды СПИИРАН" (и аналогичные, т.е. «Издательская деятельность»): положение (приказ, требования), сроки, способы подачи статей, информация о прохождении статей через редакцию и типографию, ссылки на полезные ресурсы, полный комплект лицензионного (бесплатного) программного обеспечения для подготовки и печати статей, инструкции по его установке и использованию, личные страницы сотрудников, выполненные в едином стиле, автоматически, по имеющимся материалам, в частности, фамилия и краткая биография, фото, биография, интересы (научные и прочие), публикации, прочее, по желанию сотрудника; сформирована также богатая библиотека материалов (архив), содержащая некоммерческое программное обеспечение и литературу. Есть возможность работать со страницей на русском либо на английском языке. Есть счетчики посещений веб-сайта и отдельных веб-страниц.

Такой подход имеет свои достоинства и недостатки. Из первых отметим оперативность получения информации, отсутствие необходимости специального программирования каждой веб-страницы отдельно, даже отсутствие необходимости поголовного изучения языка HTML, гарантированное единство стиля всех формируемых страниц, и т.п. Заметим, что уже здесь появляются первые признаки мета-подхода к ведению веб-сайта: единая база, стиль и процедуры формирования выходной информации; однако применение мета-свойств здесь минимальное. Из минусов важнейшими являются ограниченность быстродействия механизма (впрочем, для выполнения запросов с нынешней и предполагаемой в ближайшем будущем интенсивностью вполне достаточного), а также уязвимость механизма для злонамеренного вторжения. Решение последней проблемы стало особенно важным на следующем этапе разработки.

Важным, хотя и незаметным для пользователя, шагом уже на этом этапе стало частичное разделение содержимого сайта, преобразований и представлений содержимого на основе языков XML и XSL. В дальнейшем предполагается значительное увеличение обрабатываемых таким образом данных. Успешно решен вопрос о построении веб-сайта полностью на XML и XSL, причем обработка информации происходит на стороне сервера, в результате чего значительно уменьшается трафик, и вывод не зависит от используемого клиентом браузера.

## 5. Веб-сайты и интранет

После получения простого информационного веб-сайта мы переходим сейчас к разработке и реализации более сложной системы. Нам нужен активный веб-сайт для поддержки и управления деятельностью организации. Таким образом, должна быть сформирована несложная развиваемая интранет-система. Базовым инструментарием по-прежнему остается сервер IIS5 на компьютере с ОС MS Windows 2000 Server и аппарат CGI/Perl (хотя мог бы с не

меньшим успехом использоваться и веб-сервер Apache). Значительное внимание уделяется вопросам безопасности, фактически, каждое обращение к программе содержит имя и зашифрованный пароль пользователя, определяющие его права по получению и модификации информации. Основное отличие от предыдущего этапа — возможность изменения пользователями (посетителями веб-сайта) хранимой на сервере информации; прежде всего, это могут и должны делать сами сотрудники группы. Очень важно, что веб-сервер приобретает все более активный характер, становится все более необходимым в повседневной работе сотрудников, и все больше информации вовлекается в цикл обработки программ веб-сайта. Здесь задействованы базы данных сотрудников, их выполненных, текущих и перспективных работ, публикаций, семинаров, и другие.

Использованный аппарат представляется вполне достаточным, однако требования к веб-сайту возрастают, возрастает и нагрузка на компьютер. Частичным решением может оказаться использование программ типа PerlEx (однако это коммерческая разработка) либо переход на веб-сервер Apache с модулем `mod_perl` (при таком подходе не будет происходить многократная перекомпиляция программы, однократно задействуется память для программы и общих для всех запусков данных).

Кроме того, при предполагаемом дальнейшем расширении возможностей веб-сайта аппарат CGI (или ISAPI) оказывается недостаточным и ограниченным: сильно растут объем и сложность кода, по-прежнему есть ограничения на число обращений к серверу, но, что еще важнее, нет достаточной гарантии безопасности).

## **6. Службы на основе веб-сайтов**

Достаточно сложные распределенные многофункциональные службы на основе веб-сайтов можно получить только на основе мощных современных технологий. Одной из них, вполне соответствующей нашим запросам, является семейство технологий на основе языка Java. При этом используется единый и переносимый веб-интерфейс, а также стандартизованные протоколы для обмена информацией и программами. В частности, так можно обеспечить модульность разработки и веб-сайта, управляемость, масштабируемость, достаточную для наших целей защищенность всей системы и ее отдельных частей. Кроме того, можно реализовать распределенную многозвенную многокомпонентную систему услуг (сервисов), значительно более устойчивую и многофункциональную, чем описанные предыдущие. Однако нужно учесть и объемность (прежде всего, сложность и ресурсоемкость) получаемой системы; что, впрочем, вполне соответствует ее новым возможностям. Такие системы являются полностью динамическими.

На этом и предыдущем этапе очень важна организационная составляющая. Веб-сайт и связанные с ним структуры будут наполняться и развиваться только в том случае, если докажут свою непосредственную ежедневную полезность для большинства сотрудников организации. Сам веб-сайт должен быть максимально независим от того, кто именно изначально или в данный момент времени занимается его разработкой и наполнением: должна быть обеспечена полная преемственность разработки.

Разумеется, указанный этап — не последний. В описываемом представлении недостает гибкости и универсальности представления данных и управления разработкой и сопровождением веб-системы. Представляется, что требу-

ется стандартное средство описания и управления системой; и таким средством являются мета-системы на основе языка (формата) расширяемой гипертекстовой разметки XML и родственные ему. Успех применения XML-технологий на более ранних этапах работы дает уверенность в правильности выбранного подхода. Соответствующие средства поддержки имеются уже в целом ряде языков, прежде всего, в Java, Perl, Python, C++; однако нужны значительные усилия для реализации практически полезных гибких и расширяемых конкретных систем среднего и большого размера. Их разработкой и занимается автор в настоящее время.

## 7. Выводы

Практически всем научным организациям необходим собственный веб-сайт.

Для каждого типа организации нужно выбрать соответствующий ей уровень сложности веб-сайта, в соответствии с объемом, типами (текст, графика, выборки из баз данных) и частотой обновления материалов, а также с используемым для формирования и представления информации программным обеспечением.

При разработке веб-сайта нужно учесть, что организация и ее веб-сайт будут развиваться, и планировать изменения при развитии изначально.

В разработке и сопровождении веб-сайта нет незначительных элементов, и каждый элемент — это отдельная деятельность, нуждающаяся в планировании и четком выполнении.

По мере развития и усложнения реализации большинства веб-сайтов будут неминуемо переходить от домашних внутренних приемов к стандартным промышленным технологиям.

## Литература

- [1] Колодин М. Ю. Перспективы применения и развития мета-систем // Тезисы VI Санкт-Петербургской Международной конференции «Региональная информатика-98». — СПб, 1998. — С. 30-31.
- [2] Колодин М. Ю. Мета-преобразования структурированной информации // Труды V Международной электронной научной конференции «Современные проблемы информатизации в технике и технологиях». — Воронеж: ЦЧКИ, 2000. — С. 87.
- [3] Колодин М. Ю. Представления и преобразования гипертекстовой информации в специализированных расширяемых форматах // Труды VII международной конференции «Региональная информатика-2000». — СПб, 2000. — С. 12.
- [4] Колодин М. Ю. Персональные технологии в научных исследованиях // Тезисы международной конференции «Эволюция инфосферы». — Москва, 1995.
- [5] Колодин М. Ю. Мета-системы в научных исследованиях // Тезисы международной конференции «Региональная информатика-96». — Санкт-Петербург, 1996.
- [6] Кристиансен Т., Торкингтон Н. Perl: библиотека программиста. — СПб: Питер, 2001. — 736 с.
- [7] Валиков А. Н. Технология XSLT. — СПб: БХВ-Петербург, 2002. — 544 с.
- [8] Фролов А. В., Фролов Г. В. Сервер Web своими руками. Язык HTML, приложения CGI и ISAPI, установка серверов Web для Windows. — М.: Диалог-МИФИ, 1997. — 288 с.
- [9] Дунаев С. Intranet-технологии: WebDBC, CGI, CORBA 2.0, Netscape, Suite, Borland, IntraBuilder, Java, Javascript LiveWire. — М.: Диалог-МИФИ, 1997. — 288 с.