

# Создание структуры документа (X)HTML

*В этой главе...*

- Создание базовой структуры документа (X)HTML
- Определение заголовка документа (X)HTML
- Создание полноценного документа (X)HTML

**О**бщая структура документа (X)HTML включает заголовок и тело документа. *Заголовок* передает информацию браузеру об общих характеристиках документа, а *тело* содержит информационное наполнение документа, отображаемое в окне браузера. Для создания документа (X)HTML прежде всего нужно очертить его структуру.

В этой главе мы рассмотрим наиболее важные элементы, используемые для определения структуры документа (X)HTML, в том числе его тело и заголовок. Будет также показано, как сообщить браузеру, какая версия языка HTML (или XHTML) используется. Несмотря на то что эта информация не видима конечному пользователю, браузеры используют ее для корректного отображения содержимого документа.

## Структура документа

Невозможно найти две одинаковые страницы (X)HTML — все они включают уникальные комбинации тегов и содержимого. В то же время любая правильно спроектированная страница (X)HTML требует наличия одной и той же базовой структуры, включающей в себя:

- ✓ предложение, идентифицирующее документ как документ (X)HTML;
- ✓ заголовок документа;
- ✓ тело документа.

При создании любого документа (X)HTML нужно начинать именно с этих трех конструкций — наполнить их содержимым и разметкой можно в любой момент.



Так как базовая структура обязательна для любого документа (X)HTML, ее постоянное создание с нуля отнимает время, загружая вас однообразной работой. Большинство специализированных редакторов HTML автоматически формирует базовую структуру при создании нового документа.

# Маркировка документа (X)HTML

В начале любого документа HTML содержится *объявление типа документа* (DTD — Document Type Declaration) — DOCTYPE. В этом объявлении описывается, какая версия языка HTML или XHTML использована, что в свою очередь помогает браузеру правильно интерпретировать весь документ. В этой главе мы будем использовать спецификацию XHTML 1.0, поскольку это самая последняя из спецификаций, которую поддерживают *все* браузеры и утилиты редактирования.

## Создание объявления типа документа HTML

Если вы собрались использовать в документе спецификацию HTML 4.01 (а не XHTML), у вас есть выбор из трех возможных объявлений типа документа.

- ✓ **HTML 4.01 Transitional.** Это наиболее полная версия HTML, которая включает в себя как все структурные элементы HTML, так и все теги представления.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
```

- ✓ **HTML 4.01 Strict.** Наиболее строгая версия HTML, исключая все теги представления и использующая для управления отображением таблицы стилей CSS.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 //EN"
"http://www.w3.org/TR/html4/strict.dtd">
```

- ✓ **HTML 4.01 Frameset.** Эта версия создана на основе HTML 4.01 Transitional с добавлением тегов поддержки фреймов.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"
"http://www.w3.org/TR/html4/frameset.dtd">
```



Применение версии HTML 4.01 Frameset в настоящее время допускается, хотя и не рекомендуется по причине практически полного отказа от фреймов при создании сайтов.

## Создание объявления типа документа XHTML

Для создания документа XHTML можно воспользоваться одним из следующих объявлений типа документа.

- ✓ **XHTML 1.0 Transitional.**

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">
```

- ✓ **XHTML 1.0 Strict.**

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/xhtml1-strict.dtd">
```

### ✓ XHTML 1.0 Frameset.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"  
"http://www.w3.org/TR/xhtml1/xhtml1-frameset.dtd">
```



Подобно фреймам HTML, использование фреймов XHTML не рекомендуется.



Описания типов документов спецификации XHTML 1.0 аналогичны типам HTML 4.01 (см. главу 1). Определения документов HTML DTD подробно описаны по адресу [www.w3.org/TR/html401/sgml/dtd.html](http://www.w3.org/TR/html401/sgml/dtd.html); определения XHTML DTD — по адресу [www.w3.org/TR/xhtml1/dtds.html](http://www.w3.org/TR/xhtml1/dtds.html).

## Тег <html>

После определения версии языка, которой следует документ, вставьте в текст (X)HTML тег <html>. Он должен охватывать все остальные теги документа (X)HTML.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
"http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">  
  
<html>  
  
</html>
```

## Создание пространства имен XHTML

*Пространством имен* называют набор имен, используемых тегами и атрибутами документа XML. Язык XHTML использует набор имен XHTML. Таким образом, объявление его пространства имен выглядит следующим образом.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
"http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">  
  
<html xmlns="http://www.w3.org/1999/xhtml">  
  
</html>
```



Не нагружайте себя осмыслением значения пространства имен. Если бы вы работали с другими словарями XML, это имело бы смысл. В случае же работы с XHTML главное знать, *как* включать его пространство имен, что и было продемонстрировано в предыдущем фрагменте.

## Создание заголовка документа

*Заголовок* документа (X)HTML — первый из двух его основных компонентов (вторым является *тело*). В нем содержится общая информация о документе, в том числе его заглавие и *метаданные*: ключевые слова, информация об авторе и описание. Если в документе используются таблицы стилей, информация о них также помещается в заголовок документа.



В главе 9 подробно описан процесс создания таблиц стилей CSS и их включения в документы (X)HTML.

Тег `<head>`, определяющий заголовок документа, следует непосредственно за открывающим тегом `<html>`.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <meta http-equiv="Content-Type" content="text/html;
      charset=UTF-8" />
  </head>
</html>
```

Обратите внимание на то, что тег метаданных, определяющий символьную кодировку, следует обязательно включать в код веб-страницы, чтобы обеспечить положительные результаты верификации на сайте [validator.w3.org](http://validator.w3.org).

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
```

Не забывайте об этом теге!

## Присвоение заголовка странице

Любая страница (X)HTML должна содержать описательный заголовок, который сообщает тематику данной страницы. Этот текст отображается в заголовке окна браузера (рис. 4.1). Заголовок страницы должен быть кратким и в то же время информативным. (К примеру, заголовок *Консультационный центр Шурика* более информативен, чем *Моя домашняя страница*.)

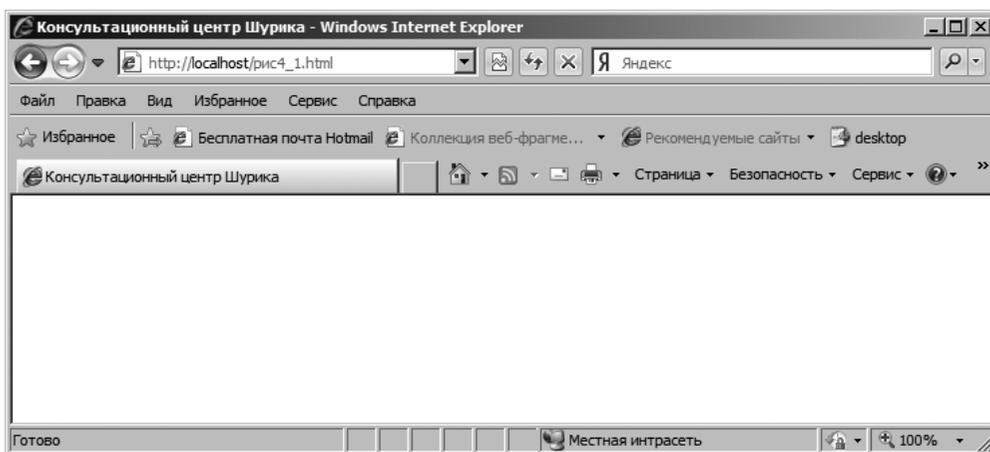


Рис. 4.1. Заголовок страницы (X)HTML отображается в заголовке окна браузера

Определение заголовка создается с помощью тега `<title>`, вложенного в тег `<head>`.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Консультационный центр Шурика</title>
  </head>
</html>
```



Заголовки страниц используются поисковыми службами для индексации веб-страниц, что позволяет включить страницу в область поиска глобальных поисковых порталов. Заголовок может стать тем первым текстом, который потенциальный посетитель прочитает о странице на поисковом портале. Информация о вашей странице в результатах запроса на поисковом портале будет отображаться наряду с сотнями других страниц, а это значит, что, создавая броский заголовок, вы получаете реальный шанс привлечь внимание потенциальных посетителей страницы.



Заголовок страницы также используется при создании закладок и журнала в браузере. Сделайте его кратким и ярким.

## Определение метаданных

*Метаданными* называют сведения о данных. В контексте Интернета это означает объяснение данных, содержащихся на веб-странице. Метаданные веб-страницы могут включать в себя:

- ✓ ключевые слова;
- ✓ описание страницы;
- ✓ информацию об авторе страницы;
- ✓ название программы, использованной для создания страницы.

## Теги и атрибуты

Каждый элемент метаданных определяется с помощью:

- ✓ тега `<meta />`;
- ✓ атрибутов `name` (имя) и `content` (содержимое).

Например, следующие элементы создают список ключевых слов и описание страницы консультационной службы.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">
```

```

<head>
<title>Консультационная служба Шурика</title>
  <meta http-equiv="Content-Type" content="text/html;
    charset=UTF-8" />
  <meta name="keywords"
    content="IT консультация, эксперт по сетям, MCSE" />
  <meta name="description"
    content="Обзор информационных сервисов,
    предоставляемых Шуриком" />
</head>
</html>

```

## Применение пользовательских имен

При изучении спецификации (X)HTML можно прийти к выводу о том, что:

- ✓ не предопределяются типы данных, которые можно включать в веб-страницу;
- ✓ жестко не закрепляются имена разных элементов метаданных, таких как ключевые слова и описания.

Исходя из этого, в приведенном выше примере можно было вместо имен элементов ключевых слов и описаний, `keywords` и `description`, использовать любые другие, например `kwrd` и `desc`.

```

<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Консультационная служба Шурика</title>
    <meta http-equiv="Content-Type" content="text/html;
      charset=UTF-8" />
    <meta name="kwrd"
      content="IT консультация, эксперт по сетям, MCSE" />
    <meta name="desc"
      content="Обзор сервисов для специалистов IT,
      предоставляемых Шуриком" />
  </head>
</html>

```



Возникает вопрос: если вы вольны в определении как имен (атрибут `name`), так и содержимого (атрибут `content`) метаданных, как система узнает, каким образом ей их обрабатывать? Ответ прост: система не обрабатывает неизвестные ей метаданные. Любой поисковый механизм работает по-своему, однако практически все они обращаются к метаданным, именами которых являются `keywords` и `description`. В то же время многие поисковые механизмы могут не обращать внимания на другие элементы метаданных, которые включены в страницу.



Многие разработчики веб-страниц включают метаданные в следующих целях:

- ✓ таким образом они оставляют послания людям, которые захотят посмотреть исходный код страницы;
- ✓ для подготовки страницы к будущим поколениям браузеров и поисковых механизмов.



Несмотря на то что ключевые слова и описания страниц не являются обязательными элементами, именно они обычно используются поисковыми системами для сбора информации о веб-страницах. Постарайтесь включить в метаданные как можно более точную и подробную информацию, если хотите, чтобы вашу страницу в процессе поиска не обходили стороной разные поисковые системы.

## Автоматическая переадресация на другую страницу

С помощью метаданных на страницу можно вставлять инструкции, указывающие браузерам, каким образом следует обрабатывать данную веб-страницу. В частности, принято использовать теги `<meta />` для автоматической переадресации посетителей с одной веб-страницы на другую. С такой задачей вы можете столкнуться, когда будете перемещать свой сайт с одного веб-сервера на другой. В этом случае на старом веб-сервере вы помещаете страницу, в которой декларируете сам факт перемещения, и вставляете тег `<meta />`, автоматически открывающий в окне браузера страницу в новом месте.

Для того чтобы использовать тег `<meta />` с целью отправки сообщения, выполните два действия.

1. **Вместо атрибута `name` используйте атрибут `http-equiv`.**
2. **Из списка predefined значений этого атрибута, представляющих инструкции браузеру, выберите нужное.**

Эти значения основаны на инструкциях, которые используются в заголовке HTTP, однако изменение заголовка HTTP в документе значительно сложнее, чем внедрение инструкций в саму веб-страницу.

Для отправки инструкций браузеру, переадресующих пользователя с одной страницы на другую, выполните описанные ниже действия.

1. **Используйте тег `<meta />` с атрибутом `http-equiv="refresh"`.**
2. **В атрибуте `content` укажите количество секунд, по истечении которых отображаемая страница будет обновлена новой, содержащейся по указанному в атрибуте `url` адресу URL.**

В приведенном ниже примере выделенная полужирным шрифтом строка определяет обновление страницы через 15 секунд с переходом к странице `www.w3.org`.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">

<bhtml xmlns="http://www.w3.org/1999/xhtml">
```

```

<head>
  <title>Все о разметке</title>
  <meta http-equiv="Content-Type" content="text/html;
    charset=UTF-8" />
  <meta http-equiv="refresh" content="15;
    url=http://www.w3.org" />
</head>

<body>
  <p>Эта страница находится в стадии разработки.
  Пока вы можете посетить <a href="http://www.w3.org">сайт
  W3C</a>, где найдете ссылки на ресурсы, посвященные
  разметке.
  </p>

  <p>Через 15 секунд вы будете автоматически
  переадресованы на эту страницу </p>
</body>
</html>

```



Осторожно используйте метаданные при перенаправлении на другую веб-страницу. Некоторые поисковые системы в случае обнаружения переадресации с помощью метаданных приходят к выводу о том, что ваш сайт рассылает спам. В результате ваш веб-сайт либо веб-страница *исключается* из списка поисковой системы. Если вы хотите стать профессионалом в области использования метаданных для переадресации, поднимитесь на следующий уровень и попытайтесь выполнить переадресацию с помощью HTTP-кода с номером 301. При этом сервер осуществляет переадресацию на основе файла \*.htaccess, находящегося в корневом каталоге вашего веб-сервера. Поскольку описание серверных перенаправлений выходит за рамки этой книги, воспользуйтесь поиском в Google для получения дополнительной информации.



Не все старые браузеры умеют обращаться с тегом `<meta />`, содержащим атрибут `http-equiv`, и могут не выполнять автоматическую переадресацию. По этой причине обязательно вставляйте на страницу переадресации текст, описывающий проблему, и явную ссылку на новую страницу, чтобы пользователь смог выполнить данную операцию вручную. Создание ссылок с помощью тегов `<a>` описывается в главе 6.

Когда откроется страница, содержащая описанный выше код HTML, пользователь увидит то, что показано на рис. 4.2.



Атрибут `http-equiv` может использоваться в теге `<meta />` и для множества других целей, в частности для установки времени жизни страницы и определения использованного на ней набора символов (т.е. языка). Полный список операций, которые можно выполнить с помощью атрибута `http-equiv`, приведен на странице словаря тегов языка HTML:

<http://vancouver-webpages.com/META/metatags.detail.html>

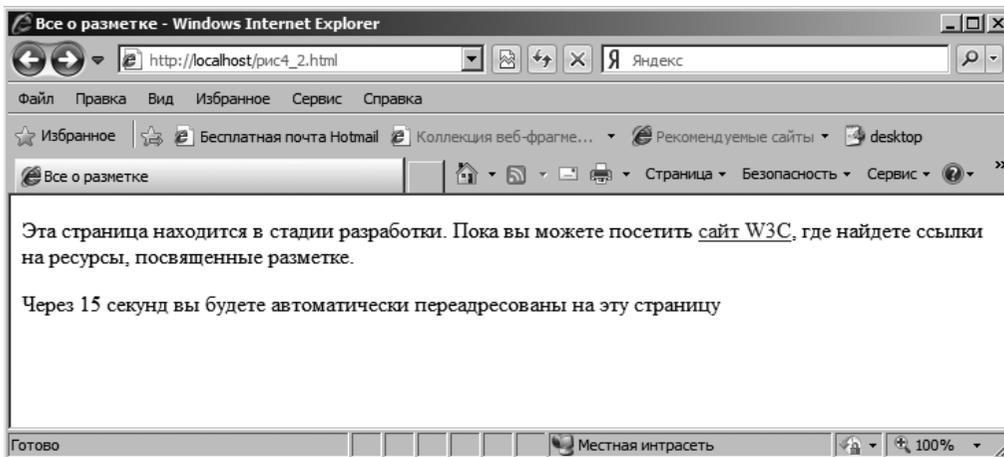


Рис. 4.2. При использовании тега `<meta />` для автоматической переадресации вставляйте на страницу ссылку, позволяющую пользователю выполнить переход вручную

## Создание тела документа (X)HTML

После того как будет создан заголовок страницы, определены ее заглавие и метаданные, можно приступать к наполнению страницы содержимым: разметкой и текстом, отображаемым в окне браузера. Все содержимое страницы хранится в теге тела документа `<body>`.



Если хотите, чтобы в окне браузера что-то отображалось, поместите это содержимое в область действия тега `<body>`, как показано ниже.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
"http://www.w3.org/TR/xhtml1/xhtml1-transitional.dtd">

<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <meta http-equiv="Content-Type" content="text/html;
      charset=UTF-8" />
    <title>Консультационная служба Шурика</title>
    <meta name="kwrд"
      content="IT консультация, эксперт по сетям, MCSE" />
    <meta name="desc"
      content="Обзор сервисов для специалистов IT,
      предоставляемых Шуриком" />
  </head>

  <body style="color: white;
    background-color: teal;
    font-size: 1.2;
    font-family: sans-serif">
    <h1>Консультационный центр Шурика</h1>
```

```
<p>Шурик имеет двадцатилетний опыт проведения консультаций по самым различным вопросам информационных технологий. В нашем центре вы можете получить квалифицированную помощь по самым различным вопросам, от создания и конфигурирования сети до написания технической документации. С помощью нашего консультационного центра вы сможете создать действительно инновационную информационную инфраструктуру.</p>
```

```
<p>Дополнительную информацию вы можете получить, обратившись к Шурику по адресу <a href="mailto:shurik@conquestmedia.com">Шурик</a>.</p>
```

```
</body>  
</html>
```

На рис. 4.3 показано, как браузер отображает эту страницу, в частности:

- ✓ содержимое тега `<title>` отображается в заголовке окна;
- ✓ тег `<meta />` вообще не оказывает никакого влияния на содержимое страницы;
- ✓ в окне браузера отображается только текст, содержащийся в тегах заглавия (`<h1>`) и абзаца (`<p>`) внутри тега `<body>`.

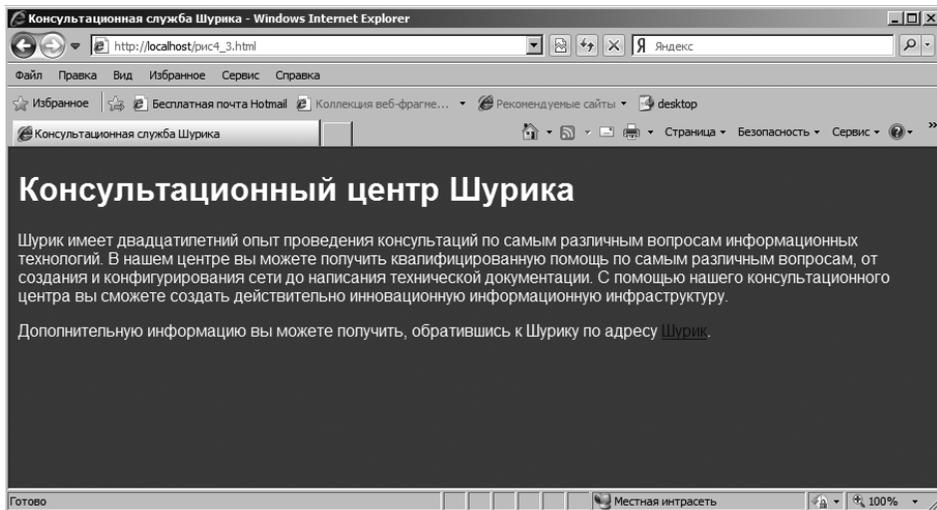


Рис. 4.3. В окне браузера отображается содержимое тега `<body>`