

## Глава 2. Структура HTML-документа

### 2.1 Создание HTML-документа

Документы в формате HTML являются обычными текстовыми файлами с расширением ".htm" или ".html", содержащими специальные теги (управляющие элементы) разметки. Теги разметки указывают браузеру, как надо вывести страницу. Поэтому HTML-документ можно создавать с помощью как специальных HTML-редакторов так и с помощью любого текстового редактора.

В нашем случае для изучения основ HTML достаточно воспользоваться, например, редактором Notepad (Блокнот), входящий в стандартные средства операционной системы Windows.

### 2.2. Теги

**Тег (tag)** - способ разметки текста.

Один из переводов английского слова tag - это этикетка (бирка). Для чего нужны бирки? Чтобы пометить что-нибудь, то есть, при помощи бирки можно сообщить об обычной вещи дополнительную информацию.

В HTML-документе тег - это метка, которая прикрепляется к участку простого текста и служит для сообщения о нем дополнительной информации, например, что участок текста надо вывести жирным шрифтом.

Когда браузер получает символы простого текста, он их просто отображает на экране, например, если текст содержит символы "Hello world!", то и на экране появляются именно они. Совершенно другие события должны происходить, когда браузер встречает тег. Символы, составляющие тег, браузер не должен выводить на экран, а должен анализировать. Потом результаты этого анализа он должен использовать для другой информации, к которой прикреплен тег.

Чтобы это произошло, некоторым символам должна быть дана роль специальных, управляющих символов. Главными из таких символов являются открывающая и закрывающая угловые скобки:

"<" и ">".

Синтаксис
<b>&lt;тег&gt;</b> элемент страницы

Создадим простейший пример.

Откройте текстовый редактор и напишите текст:

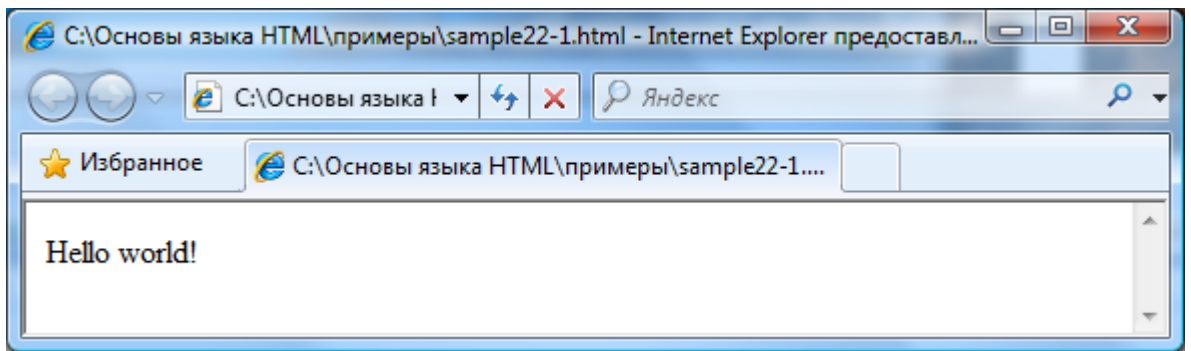
Hello world!

Сохраните его под именем:

sample22-1.html

Содержимое файла sample22-1.html:  
Hello world!

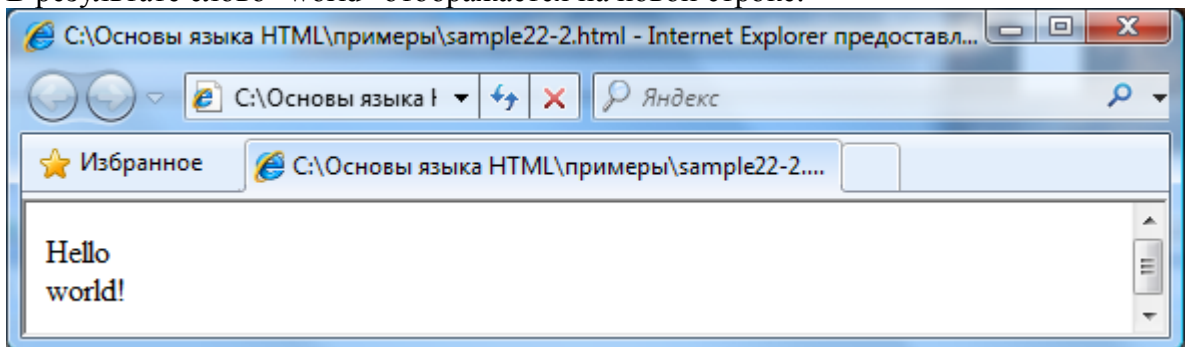
В результате получаем:



Создаем ещё один файл.  
Введём управляющий тег `<br>`.

Содержимое файла sample22-2.html:  
Hello`<br>`world!

В результате слово "world" отображается на новой строке:



Имена тегов можно записывать как маленькими, так и большими буквами. Так, чтобы достичь того же результата, что и выше, можно было использовать большие буквы.

### 2.3. Теги-контейнеры

Некоторые теги состоят из пары элементов - открывающего тега и закрывающего тега. Закрывающий тег дополнительно содержит после открывающей угловой скобки символ `/`. Например:

`<b>элемент страницы</b>`

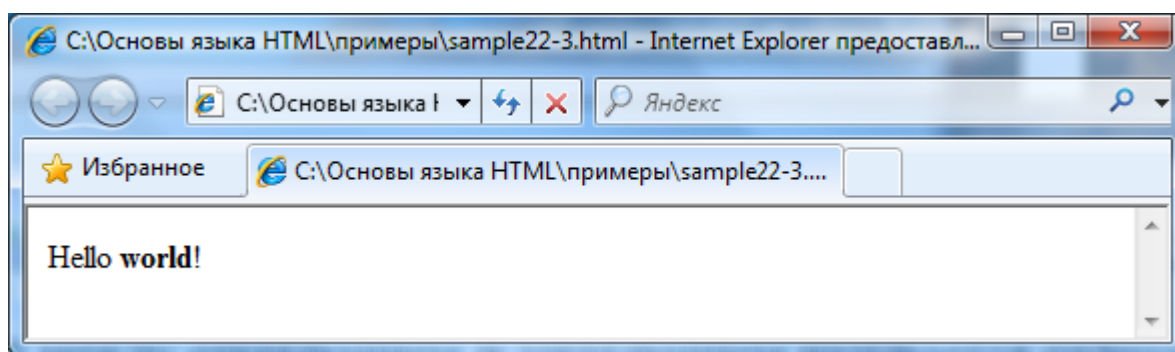
Текст, расположенный между открыванием и закрыванием называется содержимым тега. Теги, у которых есть содержимое, называются **тегами-контейнерами**.

Синтаксис
<code>&lt;тег&gt;элемент страницы&lt;/тег&gt;</code>
Общий случай:
<code>&lt;тег параметр1="значение1" параметр2="значение2" ... &gt;элемент страницы&lt;/тег&gt;</code>

Введём тег-контейнер `<b>`.

Содержимое файла sample22-3.html:  
Hello `<b>`world`</b>`!

В результате слово "world" отображается на жирным шрифтом:



В нашем случае все, что идет после символов `</b>` перестанет отображаться жирным. В примере восклицательный знак отображается уже обычным шрифтом.

Теги могут вкладываться друг в друга иерархически, но без пересечений.

Синтаксис
-----------

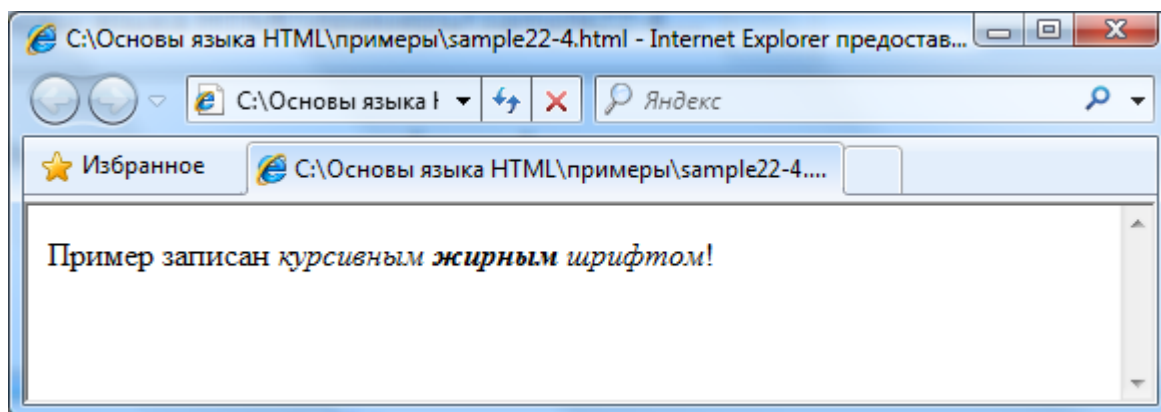
<code>&lt;тег1&gt;&lt;тег2&gt;элемент страницы&lt;/тег2&gt;&lt;/тег1&gt;</code>
---

Например, чтобы пометить текст курсивом, а внутри него сделать часть жирным шрифтом:

Содержимое файла sample22-4.html:

Пример записан `<i>`курсивным `<b>`жирным`</b>` шрифтом`</i>`!

В результате получается:



В данном примере тег `<b>` вложен в тег `<i>`.

## 2.4. Мета-теги

Тег `<meta>` - необязательный. Располагается в теле заголовка страницы. Этот тег содержит описание страницы, ключевые слова, информацию об авторе и прочую служебную информацию. Другими словами можно сказать, что это визитная карточка страницы для различных служб Интернета. Существуют два способа применения мета-тегов.

Первый способ позволяет определить некоторую переменную путем указания ее имени (name) и значения (content).

Синтаксис
<code>&lt;meta name="атрибут" content="значение"&gt;</code>

Второй способ позволяет определить эквиваленты команд протокола передачи гипертекстов.

Синтаксис
<code>&lt;meta http-equiv="атрибут" content="значение"&gt;</code>

Мета-тегов существует очень много и далеко не все они нужны на конкретной странице.

Наиболее часто используемые значения параметра **name**:

Атрибут	Пояснение
Author	Определяет имя автора страницы
Copyright	Указывает наличие авторских прав на документ
Title	Заголовок страницы
Description	Описание страницы. Если страница не содержит этого тэга, то поисковая система в качестве аннотации возьмет начальный фрагмент текста страницы, что зачастую нежелательно, т.к. в верхней части страницы может располагаться, например, реклама
Keywords	Список ключевых слов страницы через запятую. Часто используется поисковыми системами. Ключевые слова должны быть простыми и описывать содержимое страницы
Document-state	"Dynamic". Указывает, что страница динамичная, т.е. ее содержание периодически меняется. "Static". Указывает, что страница статичная, т.е. ее содержание не меняется
Generator	Содержит имя HTML-редактора, в котором создана страница
Robots	"All". Указывает поисковому роботу занести страницу в каталог поисковой системы. "Noindex". Запрет поисковому роботу занести страницу в каталог поисковой системы.

Таблица 1. Значения параметра name в мета-тегах.

Наиболее часто используемые значения параметра **http-equiv**:

Атрибут	Пояснение
Reply-to	Содержит адрес электронной почты автора

Content-Type	Указывает на тип документа и его кодировку. Например, "text/html; charset=windows-1251"
Refresh	Указывает браузеру через сколько секунд после завершения загрузки текущего документа пользователь будет перенаправлен на другую страницу, адрес которой записан в теге. Если указано только время, то будет обновляться только содержимое текущей страницы. Например, "[время в секундах];URL=[адрес страницы]"

Таблица 2. Значения параметра http-equiv в мета-тегах.

Пример мета-тегов сайта <http://www.rsreu.ru>:

```
<meta name="description" content="Официальный сайт Рязанского государственного радиотехнического университета (РГРТУ)" />
<meta name="keywords" content="радиотехнический университет, РГРТУ, Рязань, высшее образование" />
<meta name="robots" content="index, follow" />
```

Пример мета-тегов сайта <http://www.gismeteo.ru/>:

```
<meta http-equiv="content-type" content="text/html; charset=utf-8">
<meta name="description" content="Подробный прогноз погоды по 1300+ городам России и по 13000+ городам Мира на ближайшие 72 часа на GISMETEO.RU. Долгосрочные прогнозы (на две недели). Погодные карты России и Мира.">
<meta name="keywords" content="прогноз, погода, карты, осадки, ветер, влажность, давление">
```

## 2.5. Описание спецификации HTML-документа

Для того, чтобы текстовый файл превратился в HTML-файл, поменять его расширение с "txt" на "html" недостаточно.

Каждый HTML-документ, отвечающий спецификации HTML какой-либо версии, обязан начинаться со строки декларации версии HTML `<!DOCTYPE>`, например:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
```

Эта строка поможет браузеру определить, как правильно интерпретировать данный документ. В данном случае мы говорим браузеру, что HTML соответствует международной спецификации версии 4.01.

В целом HTML-документ состоит из начала, заголовка и тела.

## 2.6. Начало HTML-документа

После объявления версии и типа документа необходимо обозначить его начало и конец. Это делается с помощью тэга-контейнера `<html>`.

Синтаксис
<pre>&lt;html&gt;     текст страницы &lt;/html&gt;</pre>

## 2.7. Заголовок HTML-документа

Заголовок определяется с помощью тэга-контейнера `<head>`, который вложен в тег-контейнер `<html>`.

Синтаксис
<pre>&lt;head&gt;   тело заголовка &lt;/head&gt;</pre>

Одним из элементов тела заголовка служит специальный тег-контейнер **<title>**. Помещенный внутри него текст выводится в строке заголовка окна браузера.

Синтаксис
<pre>&lt;title&gt;   заголовок страницы &lt;/title&gt;</pre>

## 2.8. Тело HTML-документа

Именно в теле находится все то, что отображается на странице: текст, картинки, таблицы. Тело документа размещается внутри тэга-контейнера **<body>** и вложено в тег-контейнер **<html>**.

Синтаксис упрощенного варианта тега
<pre>&lt;body&gt;   тело страницы &lt;/body&gt;</pre>

Дополнительные параметры данного тега будут рассмотрены позднее.

## 2.9. Комментарий в HTML-документах

Комментарий в HTML заключаются в теги **<!--** и **-->**, который не отображается браузером.

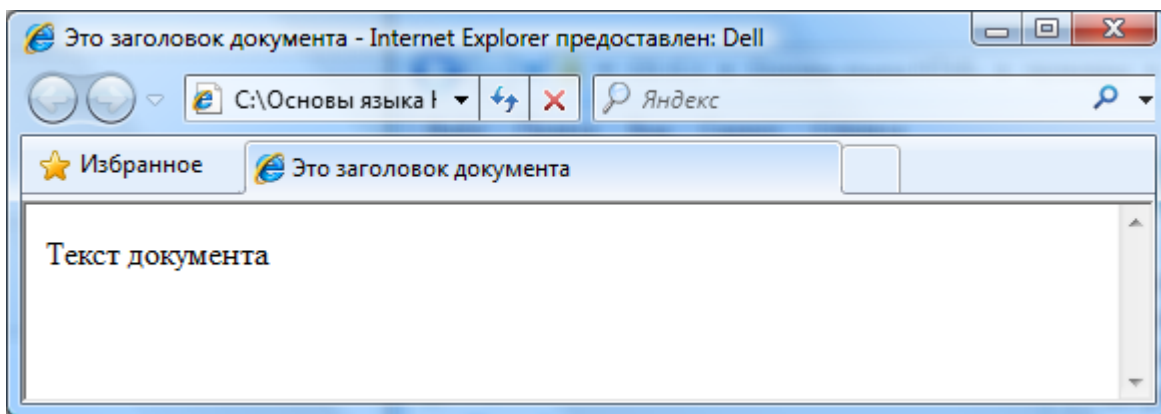
Синтаксис
<pre>&lt;!-- текст комментария --&gt;</pre>

## 2.10. Примеры

Пример короткого HTML-документа:

```
Содержимое файла sample22-5.html:
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Final//RU">
<html>
  <head>
    <title>Это заголовок документа</title>
  </head>
  <body>
    Текст документа
  </body>
</html>
```

Результат на экране:

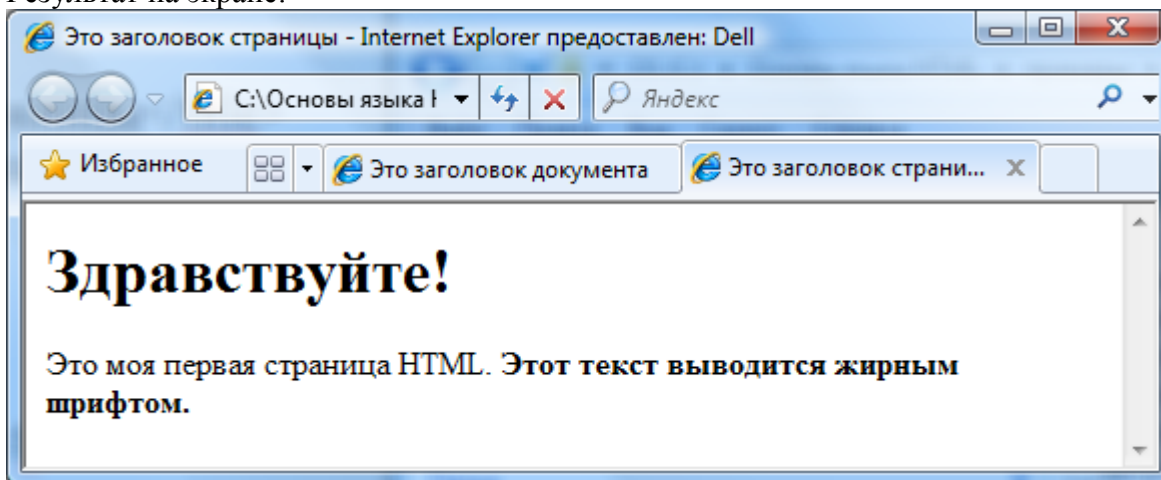


Пример:

Содержимое файла sample22-6.html:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Final//RU">
<html>
  <head>
    <title>Это заголовок страницы</title>
    <meta http-equiv="Content-Type" content="text/html;
charset=windows-1251">
  </head>
  <body>
    <h1>Здравствуйте!</h1>
    <!-- Это комментарий -->
    <p>Это страница HTML. <b>Этот текст выводится жирным шрифтом.</b></p>
  </body>
</html>
```

Результат на экране:



Разберем пример:

- Мета-тег указывает на то, что документ содержит HTML-текст в кодировке windows-1251 (кириллица).
- Весь текст внутри тега-контейнера `<body>` отображается в окне браузера.
- Текст «Здравствуйте!» в теге-контейнере `<h1>` отображается стилем заголовка первого уровня (обычно жирным шрифтом большего размера).
- Комментарий не отображается браузером, хотя присутствует в исходном коде страницы.
- Тег `<p>` означает, что начинается новый параграф, тег `</p>` означает конец параграфа.