

ХНТМЛ

ХНТМЛ (англ. *Extensible Hypertext Markup Language* — расширяемый язык разметки гипертекста) — семейство языков разметки веб-страниц на основе XML, повторяющих и расширяющих возможности HTML 4. Спецификации ХНТМЛ 1.0 и ХНТМЛ 1.1 являются рекомендациями консорциума Всемирной паутины.

Главное отличие ХНТМЛ от HTML заключается в обработке документа. Документы ХНТМЛ обрабатываются своим модулем (парсером) аналогично документам XML. В процессе этой обработки ошибки, допущенные разработчиками, не исправляются.

ХНТМЛ соответствует спецификации SGML, поскольку XML является её подмножеством. HTML обладает множеством особенностей в процессе обработки и фактически перестал относиться к семейству SGML, что и закреплено в черновике спецификации HTML 5.

Браузер выбирает парсер для обработки документа на основании заголовка `content-type`, полученного от сервера:

- HTML — `text/html`
- ХНТМЛ — `application/xhtml+xml`
- Для локального просмотра на клиенте выбор основывается на расширении файла.
- В Internet Explorer вплоть до 8-й версии парсер обработки ХНТМЛ-документов отсутствует.

Вопреки мнению начинающих разработчиков, декларация `DOCTYPE` не оказывает никакого влияния на определение парсера для обработки.

Различия между ХНТМЛ и HTML

- Все элементы должны быть закрыты. Теги, которые не имеют закрывающего тега (например, `` или `
`) должны иметь на конце `/` (например, `
`).
- Булевы атрибуты записываются в развёрнутой форме. Например, следует писать `<option selected="selected">` или `<td nowrap="nowrap">`.
- Имена тегов и атрибутов должны быть записаны строчными буквами (например, `` вместо ``).
- ХНТМЛ гораздо строже относится к ошибкам в коде; `<` и `&` везде, даже в URL, должны замещаться `<` и `&` соответственно. По рекомендации W3C браузеры, встретив ошибку в ХНТМЛ, должны сообщить о ней и не обрабатывать документ. Для HTML браузеры должны были попытаться понять, что хотел сказать автор.
- Кодировкой по умолчанию является UTF-8 (в отличие от HTML, где кодировкой по умолчанию является ISO 8859-1).

Для ХНТМЛ страниц рекомендуется задавать MIME-тип — `application/xhtml+xml`, но это не является обязательным, более того — браузер Internet Explorer 8 и младшие версии, не смогут обрабатывать страницу, поэтому с ХНТМЛ 1.0 традиционно используется MIME-тип для HTML — `text/html`.

Также стандарт рекомендует указание `<?xml version="1.0" encoding="utf-8"?>` перед DTD, но это не обязательно, более того — браузер Internet Explorer воспринимает такое указание (как и любой другой текст перед `<!DOCTYPE>`), как признак того, что данную страницу необходимо отображать в режиме обратной совместимости, а не согласно стандарту. Существует три типа документов ХНТМЛ: `strict`, `transitional` и `frameset`. Наиболее употребительной и универсальной из версий ХНТМЛ является переходная (англ. *transitional*), поскольку она позволяет использовать `iframe` (включение содержимого одной веб-страницы в другую) и атрибут `target` у ссылок (для указания того, например, что ссылке необходимо открываться в новом окне). Фреймовая версия

(англ. *frameset*) представляет собой расширенный вариант *transitional* и добавляет к нему, как следует из названия, возможность установки *frameset* вместо *body*. DTD строгой версии XHTML (англ. *strict*) не содержит многих тегов и атрибутов, описанных в DTD *transitional*, и признанных устаревшими.

Версии XHTML

XHTML 1.0

Спецификация XHTML 1.0 одобрена в качестве рекомендации консорциума Всемирной паутины 26 января 2000 года. 1 августа 2002 года была опубликована вторая редакция спецификации.

XHTML 1.0 определяет три типа документов (по аналогии с HTML 4):

- *XHTML-1.0-Strict*: тип, исключающий использование всех элементов и атрибутов, задающих графическое отображение (например, элемент `font` или атрибут `bgsColor`). Это наиболее предпочтительный тип документов.
- *XHTML-1.0-Transitional*: тип, допускающий использование элементов и атрибутов, задающих графическое отображение. Указанные элементы рекомендуется использовать только при необходимости, например, для облегчения миграции документов с HTML 3.2 на XHTML.
- *XHTML-1.0-Frameset*: тип, добавляющий к возможностям *XHTML-1.0-Transitional* использование фреймов.

XHTML 1.1

- *XHTML 1.1 Модульный* (Module-based): авторы могут импортировать дополнительные свойства в их разметку.

XHTML Basic и XHTML MP

- *XHTML Основной* (Basic): специальная облегчённая версия XHTML для устройств, которые не могут использовать полный набор элементов XHTML — в основном используется в миниатюрных устройствах, таких как мобильные телефоны. Предполагается, что он заменит WML и С-HTML.
- *XHTML мобильного профиля* (Mobile Profile): основанный на XHTML Basic, добавляет специфические элементы для мобильных телефонов.

XHTML 2.0

Основная статья: XHTML 2.0

XHTML 2.0 является последней версией XHTML. Эта версия не достигла статуса рекомендации: деятельность рабочей группы XHTML 2.0 была прекращена в конце 2010 года, а все ресурсы перенесены в рабочую группу HTML 5^[1]. Достигнутые результаты работы над XHTML 2.0 были опубликованы 16 декабря 2010 года в виде отчета рабочей группы (англ. *Working Group Note*).

Валидация XHTML документов

Валидным (т. е. отвечающим всем правилам) XHTML-документом считается документ, удовлетворяющий технической спецификации. В идеале, все браузеры должны следовать веб-стандартам и, в соответствии с ними, валидные документы должны отображаться во всех браузерах на всех платформах. Валидация XHTML-документа рекомендована даже несмотря на то, что она не гарантирует кросс-браузерности. Документ может быть проверен на соответствие спецификации с помощью онлайн-Службы валидации разметки W3C. Валидация обнаружит и разъяснит ошибки в XHTML-разметке.

Валидный документ должен содержать определение типа документа (DTD). DTD должен быть расположен до всех других элементов документа. Вот наиболее распространённые типы DTD для XHTML:

XHTML 1.0 Strict

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

XHTML 1.0 Transitional

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

XHTML 1.0 Frameset

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

XHTML 1.0 Mobile

```
<!DOCTYPE html PUBLIC "-//WAPFORUM//DTD XHTML Mobile 1.0//EN"
"http://www.wapforum.org/DTD/xhtml-mobile10.dtd">
```

XHTML 1.1

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

Пространство имён тегов `xhtml` должно быть обозначено в корневом теге как `"http://www.w3.org/1999/xhtml"` [2]

Самыми распространёнными ошибками в XHTML-разметке являются:

- Незакрытые элементы (XHTML, в отличие от HTML, требует закрытия всех элементов, в том числе не имеющих закрывающего тега, как, например, `
`).
- Отсутствие альтернативных текстов для изображений (достигающийся применением атрибута `alt`, который помогает сделать документы доступнее для устройств, которые не в состоянии отображать изображения, или предназначенных для слабовидящих людей).
- Присутствие текста непосредственно в теге `<body>` документа (должен быть объявлен блочный элемент, внутрь которого следует помещать содержимое).
- Вложение блочных элементов внутрь инлайновых (внутристрочных) (например, блочные элементы `<div>` или `<p>` не могут быть вложены внутрь инлайновых элементов `<a>`, ``, `` и так далее).
- Пренебрежение заключением значений атрибутов в кавычки (`` вместо ``).
- Неправильное вложение элементов (конструкции вида ` `).
- Неправильное использование ссылок-мнемоник (например, `&` вместо `&`);
- Написание тегов и/или атрибутов прописными буквами (`<DIV style="...">` вместо `<div style="...">`).
- Задание в теге `<!DOCTYPE ...>` относительного пути к DTD-файлу, скопированного с сайта `www.w3.org` ("`DTD/xhtml11.dtd`" вместо `"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd"`).

Это неполный список, но дающий общее представление о часто допускаемых ошибках при составлении XHTML-документов.

Валидный XHTML-документ можно снабжать специальным баннером W3C, подтверждающим валидность XHTML-разметки