META теги

META теги используются для описания свойств HTML документа и должны находится в рамках тега HEAD. Если в документе используется тег TITLE, то META теги рекомендуется вставлять в документ после него.

META теги имеют очень широкую функциональную направленность, но многое еще не стандартизировано или находится на стадии разработки.

META теги могут идентифицировать авторство HTML документа, его адрес и как часто он обновляется. Поисковые системы используют META теги для индексации и формирования заголовков HTML документов.

META теги могут влиять на режим отображения HTML документов, хотя сами на экран не выводятся.

Есть два типа META тегов: **NAME** и **HTTP-EQUIV**.

- META теги типа NAME содержат текстовую информацию о документе, его авторе и некоторые рекомендации для поисковых машин. Например: Robots, Description, Keywords, Author, Copyright.
- META теги типа HTTP-EQUIV влияют на формирование заголовка документа и определяют режим его обработки.

Примечание – подчеркнуты снизу чертой те мета теги, которые наиболее часто применяются при создании сайта и имеют практическую пользу.

Описание тегов типа NAME

Author, Copyright

Эти теги не отражаются броузерами, но обрабатываются некоторыми поисковыми машинами. Как правило, нет необходимости использовать эти теги одновременно. В большинстве случаев достаточно одного информационного тега, позволяющего идентифицировать автора или принадлежность документа.

Тег **Author** должен содержать имя автора, если сайт принадлежит организации, имеет смысл использовать тег **Copyright**. Информация обязательно должна заключатся в кавычки, формат произвольный, например:

```
<meta name="Author" content="Смирнов Сергей">
```

Дополнительно информационные теги могут содержать атрибут «**Lang**», указывающий язык, на котором написано значение свойства.

Пример: `<meta name="Author" Lang="ru" content="Смирнов Сергей">`.

Description

Meta тег **Description** – служит для краткого описания странички. Он используется поисковыми машинами для индексации и в качестве краткой аннотации при ссылке в ответе на запрос. И именно по содержанию этого тега пользователь поисковой системы будет оценивать, соответствует ваш сайт его ожиданиям или нет.

Если META тег **Description** не используется, то в качестве описания берется первая строка или отрывок из текста с найденным ключевым словом. Желательно чтобы длина описания не превосходила 80 символов.

Пример: `<meta name="Description" content= ". . .">`

Document-state

Document-state – предназначен для управления индексацией поисковых роботов. Может иметь два значения:

- **Static** – Нет необходимости индексировать эту страницу в будущем.
- **Dynamic** – Индексировать эту страницу регулярно (по умолчанию).

Пример: `<meta name="Document-state" content ="Dynamic">`

Generator

Meta тег **Generator** не несет полезной нагрузки. Используется для указания имени программы, с помощью которой создан html-код страницы.

Keywords

Meta тег **Keywords** используется поисковыми машинами для оценки релевантности.

Слова, содержащиеся в Meta теге Keywords, рассматриваются поисковыми машинами как рекомендация, но иногда именно этой капли не хватает, чтоб перевесить чашу.

При формировании списка ключевых слов для Meta тега Keywords необходимо использовать слова, содержащиеся в тексте документа. Слова, не содержащиеся в тексте, можно использовать в очень небольших количествах, два – три слова. Вставлять их надо в конец списка. Ключевые слова не должны повторятся в Meta теге Keywords, в крайнем случае, не более двух раз. В Meta теге Keywords имеет смысл использовать не более 10 слов, большее количество релевантности не увеличит.

Пример: `<meta name="Keywords" content="пример, meta, тега">`

Если документ написан на нескольких языках, можно использовать добавочный атрибут **lang** для выбора кодировки (смотри описание Meta тега Content-Language).

Пример:
```
<meta name="Keywords" lang="ru" content="пример, meta, тега">
<meta name="Keywords" lang="en" content="meta tegs, example">
```

Но более предпочтительно сделать отдельные страницы, каждая на своем языке с переходами с одной на другую, с помощью тегов <A> и <Link>.

Resource-type

Meta тег **Resource-type** – описывает состояние данного документа. Если его значение отлично от «Document», то поисковые системы не будут его индексировать. Предназначен для использования в крупных проектах, с множеством документов разного типа.

Некоторые возможные значения: Build, Classification, Creation, Document (принимается по умолчанию), Formatter, Host, Operator, Random text, Rating, Site-languages, Subject, Template, Version.

Пример: `<meta name="Resource-type" content ="Document">`

Revisit-after

Meta тег **Revisit-after** – Указывает поисковому роботу, через сколько дней ему нужно вернуться и переиндексировать данный документ.

Пример (приходить раз в неделю): `<meta name="REVISIT-AFTER" content="7">`

Примечание. В настоящее время поисковые системы этот мета тег игнорируют (http://help.yandex.ru/webmaster/?id=995296).

Robots

Meta тег **Robots** – содержит указания для роботов поисковых машин, собирающих информацию о HTML документах в сети.

Значение свойства Robots может состоять из следующих директив, разделенных запятыми:

- Index – эта страница должна быть индексирована.
- Noindex – эта страница не должна индексироваться.
- Follow – прослеживать гиперссылки на странице.
- Nofollow – не прослеживать гиперссылки на странице.
- All (эквивалентно index, follow) – принято по умолчанию.
- None (эквивалентно noindex, nofollow).

Тег `<meta name="Robots" content="all">` ничего не меняет в работе робота поисковой машины, так как значение «all» принято по умолчанию.

Но если эта страница создается динамически или очень часто обновляется, то нет смысла ее индексировать, так как информация о странице в поисковой машине и ее истинное содержание будут разными. В этом случае можно рекомендовать поисковой машине не индексировать эту страницу, а только отслеживать на ней гиперссылки, чтобы проиндексировать остальную часть сайта:

`<meta name="Robots" content="noindex,follow">.`

При использовании тега, запрещающего отслеживание гиперссылок `<meta name=Robots content="index,nofollow">` часть сайта может остаться не проиндексированной, так как внутри сайта роботы перемещаются по гиперсылкам. И те страницы, переход на которые заблокирован, не будут индексированы.

META-тег Robots имеет приоритет над директивами управления, заданными в файле **robots.txt**. Подробнее о запрете индексации сайта (или его отдельных частей) можно посмотреть здесь http://help.yandex.ru/webmaster/?id=996567.

Subject

Meta тег **Subject** используется поисковыми машинами для определения тематики документа. Но до тех пор, пока поисковые системы не согласуют классификационные таблицы, использование этого тега не всегда оправдано.

URL

Meta тег URL – служит для исключения из индексации генерируемых страниц и зеркал. Встретив этот тег, робот поисковой машины должен прекратить индексацию текущего документа и перейти по указанной ссылке.

Пример: `<meta name="URL" content= "http://www.Main-Site.com">`

Описание тегов типа HTTP-EQUIV

Cache-Control

Cache-Control – Контроль кэширования. Может принимать следующие значения:

- **no-cache** – Запрет на кэширование вообще (документ не будет кэшироваться ни proxy-сервером, ни браузером)
- **private** – Документ будет кэшироваться браузером, но не будет кэшироваться proxy-сервером
- **public** – Документ будет кэшироваться, даже если и не должен, вроде бы, при обычных обстоятельствах
- **no-store** – Документ кэшируется, но не сохраняется в архиве
- **max-age=*time*, must-revalidate** – Этот параметр указывает броузеру сколько времени хранить страницу в кэше. (В параметре max-age=***time*** указано, на сколько секунд кэшируется документ).
- **max-age= *time*, proxy-revalidate** – Этот параметр указывает proxy-серверу сколько времени хранить страницу в кэше.

Пример: `<meta http-equiv="Cache-Control" content="max-age=3600, must-revalidate">`

Content-Language

Meta тег **Content-Language** – это указание языка документа. Используется поисковыми машинами при индексировании. Хотя большинство из них умеют различать язык по тексту.

Некоторые возможные значения:

- de – Немецкий.
- en – Английский.
- fr – Французский.
- ru – Русский.

Пример: `<meta http-equiv="Content-language" content ="ru">`

В спецификации HTML 4.0 есть альтернативная возможность явного указания языка – `<html lang="en">`

Content-Type

Content-Type – Указание типа документа и кодировки символов.

Использовать Meta тег Content-Type надо только с учетом некоторых нюансов.

Во первых – кодировка символов текста должна соответствовать кодировки указанной в теге.

Во вторых – сервер не должен менять кодировку текста при обработке запроса броузера.

В третьих – если сервер меняет кодировку текста, он должен скорректировать или удалить meta тег Content-Type.

Несоблюдение этих требований может привести к следующему: web-сервер автоматически определит кодировку запроса клиента и отдаст страничку web-броузеру перекодированной. Броузер, в свою очередь, будет читать документ в соответствии с META-тегом Content-Type. И если кодировки не совпадут, то прочитать документ можно будет только после ряда замысловатых манипуляций. Особенно это характерно для старых броузеров IE 3.x – 4.x и Netscape 4.x.

Некоторые из возможных типов кодировки (стандарт [ISO10646]):

- ISO-8859-1 – Latin-1, для большинства западноевропейских языков.
- Windows -1251- Кириллица (Windows).
- KOI8-r – Кириллица (КОИ8-Р).
- cp866 – Кириллица (DOS).
- Windows-1252 – Западная Европа (Windows).
- Windows-1250 – Центральная Европа (Windows).

Пример: `<meta http-equiv="Content-Type" content="text/html; charset=windows-1251">`

Expires

Expires – Управление кэшированием. Если указанная в теге дата прошла, то броузер должен сделать повторный запрос, а не использовать копию из кэша. Если изначально указать прошедшую дату, то документ не будет кэшироваться. Дата должна указываться в стандарте [RFC850].

Некоторые поисковые роботы могут отказаться индексировать документ с устаревшей датой.

Пример: `<meta http-equiv="Expires" content="Wed, 26 Feb 1999 08:21:57 GMT">`

Refresh

Refresh – Определение задержки времени в секундах, после которой броузер автоматически обновляет документ. Дополнительная возможность – автоматическая загрузка другого документа.

Пример: `<meta http-equiv="Refresh" content ="4; URL=http://www.name.com/">`

Если нужно просто обновить документ, то URL указывать не обязательно.

Set-Cookie

Set-Cookie – Настройка cookie броузера.

Пример:

```
<meta http-equiv="Set-Cookie" content="NAME=value; EXPIRES=date; DOMAIN=domain_name;
PATH=path; SECURE">
```

Атрибуты EXPIRES, DOMAIN, PATH и SECURE при необходимости можно опустить.

Пример: `<meta http-equiv="Set-Cookie" content="NAME=value;">`

- **NAME** Имя cookie, не может содержать символы перевода строки, пробелов, точки с запятой (;) и табуляции.
- **EXPIRES** Время хранения cookie. После указанной даты (в формате «Wdy, DD-Mon-YYYY HH:MM:SS GMT») истекает время хранения cookie.
  Если этот атрибут не указан, то cookie хранится в течение одного сеанса, до закрытия броузера.
- **DOMAIN** Домен, для которого задается значение cookie. Для доменов COM, EDU, NET, ORG, GOV, MIL, INT значение можно задавать сокращенно «MY.COM», оно будет распространяться и на домен «WWW.MY.COM». Для всех остальных (в том числе и RU) значение надо задавать полностью «WWW.MY.RU».
  Если этот атрибут опущен, то по умолчанию используется доменное имя сервера, с которого было выставлено значение cookie.
- **PATH** Устанавливает подмножество документов, на которые распространяется действие cookie. При значение «/doc» действие cookie будет распространено на все файлы и каталоги в этой директории начинающееся на «doc» (/doc/, /document/, /doc2/, docs.html, doc-test.htm).
  Если этот атрибут опущен, то значение cookie распространяется только на документы директории, в которой расположен текущий документ.
- **SECURE** Указывает, что информация о cookie должна пересылается через HTTPS (HTTP с использованием SSL). В противном случае информация о cookie будет пересылается по протоколу HTTP.

Перед запросом к серверу, броузер проверяет cookie. И если атрибуты NAME, DOMAIN и PATH совпадают, броузер посылает cookie серверу.

Если cookie принимает новое значение, старое значение удаляется.

Броузер имеет ряд ограничений по работе с cookie:

- Единовременно может храниться не более 300 значений cookie.
- Cookie не может превышать 4 Кбайт.
- От одного сервера (домена) не может быть более 20 cookie.
- Эти ограничения могут быть несколько иными, в зависимости от настройки и типа броузера.
- Если происходит превышение лимита (общего или по домену) удаляется первая по времени запись. Удаление происходит, не зависимо от значения атрибута EXPIRES.
- При превышении 4 Кбайт, cookie усекается.

Пример: `<meta http-equiv="Set-Cookie" content="NAME=cookexample; EXPIRES=Monday, 22-FEB-02 21:11:51 GMT; DOMAIN=WWW.MY.RU; PATH=/; SECURE">`

Window-target

Window-target – Определяет окно текущей страницы. Приведенный пример может быть использован для прекращения появления новых окон броузера при применении фреймовых структур. Действует для многих (но не для всех) броузеров.

Пример: `<meta http-equiv="Window-target" content ="_top">`

Другие теги

Base

Этот тег определяет для данного документа базовый адрес URL, который затем будет использоваться для переопределения относительных адресов. Таг BASE полезен, если документ может быть скопирован в другое место или запрошен по другому пути с использованием символической ссылки, или редиректа с imagemap, и т.д.

Link

Тег Link предоставляет документу независимый от среды метод определения отношения данного документа к другим документам и ресурсам Сети. Используется с аргументами REL и REV. С помощью тега Link можно:

- создавать в документе специальные навигационные кнопки или меню
- управлять процессом отображения набора HTML файлов в печатные документы
- привязывать такие ассоциированные ресурсы, как таблицы стилей и скрипты
- предоставлять альтернативные формы для данного документа

Пример: `<link rel="stylesheet" type="text/css" href="style.css">`