

# 1

# КРАТКАЯ ИСТОРИЯ РАЗМЕТКИ

HTML — СВЯЗУЮЩИЙ ЯЗЫК Всемирной паутины. С помощью простых тегов, которые содержит этот язык, род человеческий сумел создать ошеломительно разнообразную сеть документов, связанных между собой гиперссылками — от Amazon, eBay и Wikipedia до личных блогов и страничек, посвященных котикам, похожим на Гитлера.

HTML5 — последняя на данный момент итерация этого лингва-франка, и хотя это и самое амбициозное изменение в нашем Всеобщем Наречии, но обновляется HTML не впервые. Язык начал развиваться с самого начала.

Как и собственно веб, гипертекстовый язык разметки (HyperText Markup Language, HTML) был детищем сэра Тима Бернерса-Ли, который в 1991 году составил документ под названием HTML Tags, предложив в нем около 20 элементов, которые можно было использовать для написания веб-страниц.

Не сэр Тим придумал использовать теги, состоящие из слов в угловых скобках; такие теги уже существовали в формате SGML (Standard Generalized Markup Language, стандартный обобщенный язык разметки). Вместо того чтобы изобретать новый стандарт, сэр Тим увидел все преимущества того, чтобы разрабатывать язык как надстройку к уже существующему стандарту, — эта тенденция заметна и сейчас, в разработке HTML5.

## ОТ IETF ДО W3C: ПУТЬ К HTML 4

Такой вещи, как HTML 1, никогда не было. Первой официальной стала спецификация HTML 2.0, опубликованная IETF (Инженерный совет Интернета, Internet Engineering Task Force). Многие из пунктов появились в этой спецификации потому, что они уже существовали на практике. Например, лидировавший на рынке веб-браузер Mosaic уже в 1994 году позволял авторам веб-страниц вставлять в документы картинки с помощью тега `<img>`. Впоследствии элемент `img` появился в спецификации HTML 2.0.

На смену IETF пришел W3C, Консорциум Всемирной паутины (World Wide Web Consortium), который публиковал последующие обновления стандарта HTML на сайте <http://www.w3.org>. Во второй половине девяностых появился целый шквал исправлений в спецификации, пока в 1999 году не была наконец опубликована спецификация HTML 4.01.

В этот момент HTML подошел к своей первой развилке.

## XHTML 1 : HTML ПО ПРАВИЛАМ XML

Следующая после HTML 4.01 версия языка называлась XHTML 1.0. «X» означало «экстремальный», и каждый веб-работчик, когда начинал произносить название языка, был строго обязан скрещивать руки в форме буквы «X».

Ладно, на самом деле нет. «X» значило eXtensible, «расширяемый», а скрещивать руки, в общем, было необязательно.

Содержимое спецификации XHTML 1.0 было совершенно идентично спецификации HTML 4.01. Не было добавлено никаких новых элементов и атрибутов. Единственная разница заключалась в синтаксисе языка. Если HTML давал авторам большую свободу в том, как писать элементы и атрибуты, то XHTML требовал следовать правилам XML, гораздо более строгого языка разметки, на основе которого W3C строил большинство своих технологий.

Введение более строгих правил было не так уж и плохо. Это способствовало тому, что авторы документов стали вынуждены придерживаться общего стиля написания. Если раньше теги и атрибуты можно было писать прописными, строчными буквами или любой их комбинацией, то для того чтобы документ XHTML 1.0 проходил валидацию, требовалось, чтобы все его теги и атрибуты были написаны в нижнем регистре.

Публикация стандарта XHTML 1.0 совпала с началом поддержки CSS в браузерах. По мере того как веб-разработчики стали принимать только что появившиеся веб-стандарты (это началось с Web Standards Project), более строгий синтаксис XHTML стал рассматриваться как передовая практика в написании разметки.

Потом W3C опубликовал спецификацию XHTML 1.1.

Если XHTML 1.0 — это был простой HTML, пересказанный средствами XML, то XHTML 1.1 стал настоящим XML, безаветно и полностью. Таким образом, сервер не мог отдавать его с MIME-типом `text/html`. Но если же авторы публиковали документ с MIME-типом XML, то самый распространенный браузер в мире на тот момент — Internet Explorer — вовсе не мог отобразить документ.

Казалось, что W3C стал утрачивать чувство реальности и того, что действительно происходит с публикацией документов в вебе.

## HTML 2: ТЕРПЕНИЮ ПРИШЕЛ КОНЕЦ

Если бы персонаж Дастина Хоффмана в фильме «Выпускник» был веб-разработчиком, W3C сказал бы ему одно слово, ровно одно: XML.

С точки зрения W3C разработка HTML закончилась на версии 4. Они начали работать над XHTML 2, который был спроектирован так, чтобы привести веб к светлому, основанному на XML будущему.

И хотя название XHTML 2 звучало достаточно похоже на XHTML 1, между ними не было ничего общего. В отличие от XHTML 1, в XHTML 2 не предусмотрено обратной совместимости с существующим веб-содержимым и даже с предыдущими версиями HTML. Это должен быть чистый язык, неотягощенный неряшливой историей предыдущих спецификаций.

Наступила полная катастрофа.

## РАСКОЛ: WHATWG TF?

Внутри W3C назрело восстание. Со стороны казалось, что консорциум формулировал теоретически чистые стандарты, никак не связанные с нуждами веб-разработчиков. Представители Opera, Apple и Mozilla были недовольны этим направлением развития. Им хотелось, чтобы большее внимание уделялось технологиям, позволяющим создавать веб-приложения.

Конфликт перешел в критическую фазу на семинаре в 2004 году. Ян Хиксон (Ian Hickson), который в то время работал в Opera Software, предложил идею расширения HTML с целью сделать возможным создание веб-приложений. Это предложение было отвергнуто.

Недовольные повстанцы организовали свою собственную группу: рабочую группу по разработке гипертекстовых приложений для веба (Web Hypertext Application Technology Working Group, или сокращенно WHATWG).

## ОТ ВЕБ-ПРИЛОЖЕНИЙ 1.0 К HTML5

С самого начала WHATWG стала работать совершенно не так, как W3C. В W3C использовался подход, основанный на согласии:

вопрос поднимается, обсуждается, затем по нему голосуют. В WHATWG вопросы тоже поднимаются и обсуждаются, а окончательное решение по тому, что войдет в спецификацию, а что нет, принимает редактор — Ян Хиксон.

В теории процесс W3C выглядит более демократическим и честным. На практике же политические распри и внутренние перебранки могут очень сильно замедлять продвижение. В WHATWG, где кто угодно может внести свое предложение или мнение, но последнее слово остается за редактором, все движется быстрее. Но и у редактора все же нет абсолютной власти: собранный по личным приглашениям управляющий комитет может запустить процедуру импичмента редактора в маловероятном случае, если он поведет себя как доктор Стрейнджлав.

Сначала основной объем работы в WHATWG был разбит на две спецификации: веб-форм (Web Forms 2.0) и веб-приложений (Web Apps 1.0). Обе спецификации должны быть расширениями для HTML. Со временем они объединились в одну спецификацию, которая называлась просто HTML5.

## ОБЪЕДИНЕНИЕ

Пока в WHATWG разрабатывали HTML5, W3C продолжала работать над спецификацией XHTML 2. Нельзя сказать, что она летела по шоссе в никуда. Она ехала в никуда очень-очень медленно.

В октябре 2006 года сэр Тим Бернерс-Ли написал пост в блоге, в котором признал, что попытка заставить веб перейти с HTML на XML не имеет шансов на успех. Несколько месяцев спустя W3C выпустил новый договор для рабочей группы HTML. Вместо того чтобы начинать с нуля, они мудро решили, что в качестве фундамента для любой будущей версии HTML нужно использовать наработки WHATWG.

Эта остановка и новый запуск привели к несколько запутанной ситуации. Получилось, что W3C одновременно работал над двумя разными, несовместимыми типами разметки:

XHTML 2 и HTML 5 (обратите внимание на пробел перед пятеркой). В то же время отдельная организация, WHATWG, работала над спецификацией под названием HTML5 (без пробела), которая должна быть использована в качестве основы для одной из спецификаций W3C!

Если вы относитесь к тем веб-разработчикам, которые пытаются в этом разобраться, то знайте, что проще расшифровать смысл фильмов «Мemento», «Детонатор» и всей фильмографии Дэвида Линча, даже если смотреть их подряд.

## ХТМЛ УМЕР: ДА ЗДРАВСТВУЕТ СИНТАКСИС ХТМЛ

Туман неразберихи начал рассеиваться в 2009 году. W3C объявил, что договор на XHTML 2 не будет продлеваться. Формат был мертвым уже несколько лет, и это объявление стало только официальным свидетельством о смерти.

Как ни странно, смерть XHTML 2 не прошла незамеченной. Напротив, противники XML отреагировали на нее злорадно и использовали это объявление для того, чтобы высмеять всех, кто когда-либо использовал XHTML 1, — даже несмотря на то, что между XHTML 1 и XHTML 2 не было практически ничего общего.

В то же время авторы, которые писали на XHTML 1 с тем, чтобы следовать более строгому стилю написания кода, стали волноваться, что HTML5 будет означать возвращение к небрежной разметке.

Как вы скоро увидите, это не обязательно так. Вы можете писать на HTML5 и небрежно, и строго — как захотите.

## РАЗВИТИЕ HTML5

Текущее состояние HTML5 не такое запутанное, как было когда-то, но все равно не до конца понятное.

Над HTML5 работают две группы. WHATWG создает спецификацию HTML5 в рамках процесса «утвердить, потом

пересмотреть». Рабочая группа по HTML W3C берет эту спецификацию и проводит ее через процесс «пересмотреть, потом утвердить». Как вы легко можете представить, это непростой политический союз. По крайней мере, кажется, наконец появилось единодушие по этому назойливому вопросу — с пробелом или без пробела? (На тот случай, если вам вдруг интересно, HTML5 решили писать без пробела.)

Пожалуй, самая сбивающая с толку проблема для тех веб-разработчиков, которые пробуют кончиком ноги воду HTML5, — ответ на вопрос «когда он будет готов?»

Ян Хиксон в интервью сказал, что ожидает, что HTML5 получит статус предложенной рекомендации<sup>1</sup> в 2022 году. За этим последовала волна общественного негодования от ряда веб-разработчиков. Они не понимали, что значит «предложенная рекомендация», но уж точно знали — на руках нет столько пальцев, чтобы пересчитать, сколько лет пройдет до 2022 года.

Для негодования не было повода. В данном случае для того чтобы получить статус «предложенной рекомендации», нужно иметь две полных реализации HTML5. Учитывая объем спецификации, эта дата невероятно амбициозна. В конце концов, у браузеров не самая лучшая репутация в плане реализации существующих стандартов. Internet Explorer потребовалось больше десятилетия, чтобы добавить поддержку — всего-то — элемента `abbr`.

Та дата, которая действительно имеет значение для HTML5, — 2012 год. В этом году спецификация должна стать кандидатом в рекомендации. Это на жаргоне стандартов значит «сделано и отшлифовано».

Но даже эта дата не особенно важна для веб-разработчиков. Имеет значение тот момент, когда браузеры начнут поддерживать функциональность HTML5. Мы начали использовать

---

<sup>1</sup> Описание различных этапов, которые проходят рекомендации W3C, см.: [http://ru.wikipedia.org/wiki/Рекомендации\\_W3C](http://ru.wikipedia.org/wiki/Рекомендации_W3C). *Прим. ред.*

части спецификации CSS 2.1, как только начали выпускаться браузеры с поддержкой этих частей. Если бы мы ждали, пока каждый браузер начнет полностью поддерживать CSS 2.1, и только потом стали его использовать, мы ждали бы до сих пор.

С HTML5 ровно то же самое. Не будет никакого четкого момента, когда мы могли бы объявить, что язык готов к использованию. Однако, мы можем начинать использовать части спецификации по мере того, как веб-браузеры начинают поддерживать эти функции.

Помните, что HTML5 — это не новый язык, созданный с нуля. Это скорее эволюционное, чем революционное изменение в продолжающейся истории разметки. Если сейчас вы создаете сайты на любой версии HTML, вы уже используете HTML5.