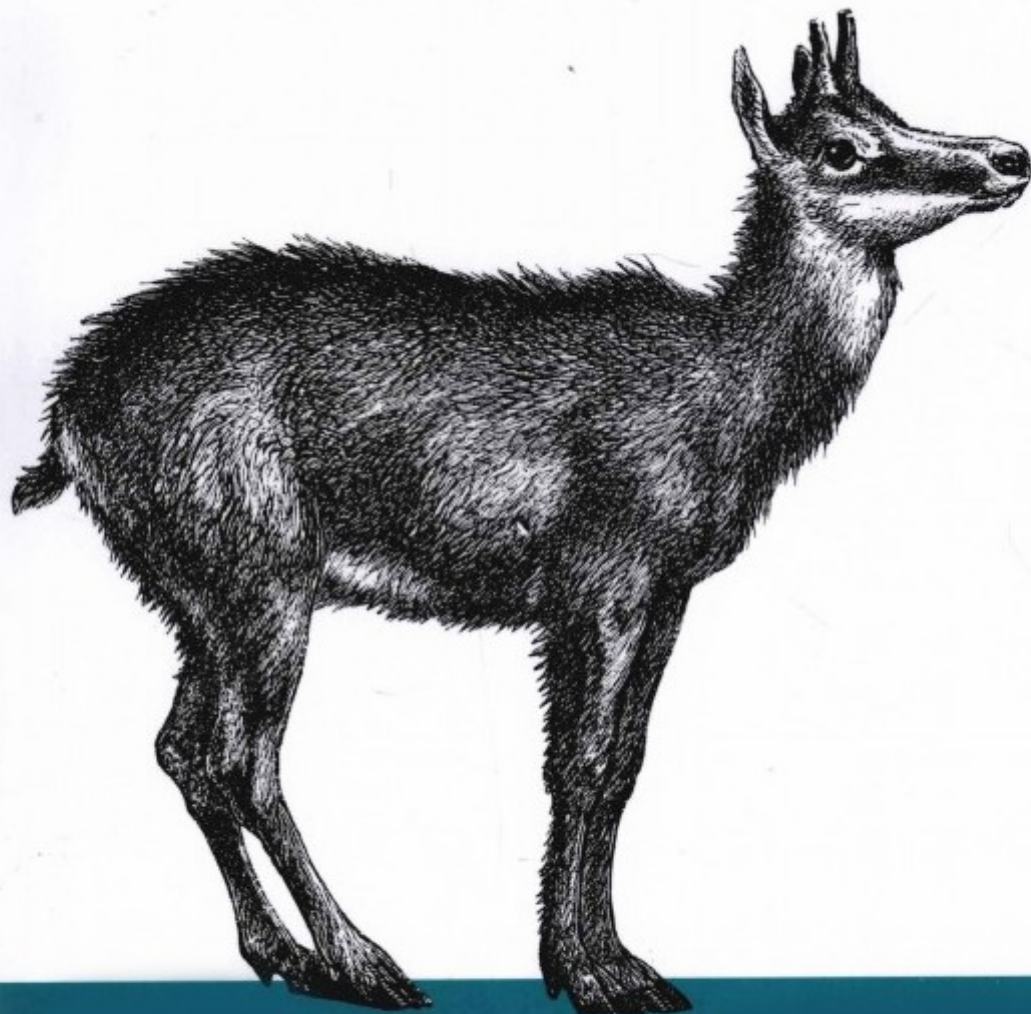


HTML5: Up and Running
Dive into the Future of Web Development



HTML5

揭秘

Mark Pilgrim 著
常可 胡金埔 赵静 译 赵泽欣 审校

O'REILLY® | Google™ PRESS



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

O'REILLY®

HTML5 揭秘

HTML5: Up and Running

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

PDG

内容简介

本书全面而深入地対 HTML5 相关的技术进行详细介绍和剖析。“从开始到现在”道出 HTML5 的坎坷发展史；“HTML5 特性检测”介绍了多种针对不同特性的检测方法；“本地存储”揭开了如何把“数据库”搬到客户端的神秘面纱；“离线应用”展示了脱机状态下依然能让 Web 应用完好无损的神奇技艺；“疯狂的表单”大秀了一把炫酷无比的下一代 Web 表单……它几乎涵盖了 HTML5 标准中描述的所有新特性。本书在以诙谐的文字结合生动的实例介绍 HTML5 特性的同时，还深入剖析其内部原理。让读者不仅知其然，而且知其所以然。无论是刚接触 Web 前端技术的新人，还是经验丰富的老手，只要是 HTML5 技术爱好者，都会从本书中受益。

978-0-596-80602-6 HTML5: Up and Running © 2010 by O'Reilly Media, Inc. Simplified Chinese edition, jointly published by O'Reilly Media, Inc. and Publishing House of Electronics Industry, 2010. Authorized translation of the English edition, 2010 O'Reilly Media, Inc., the owner of all rights to publish and sell the same. All rights reserved including the rights of reproduction in whole or in part in any form.

本书中文简体版专有出版权由 O'Reilly Media, Inc. 授予电子工业出版社，未经许可，不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号 图字：01-2010-6579

图书在版编目 (CIP) 数据

HTML5 揭秘 / 皮尔格林 (Pilgrim, M.) 著；常可，胡金埔，赵静译。

—北京：电子工业出版社，2010.12

书名原文：HTML5: Up and Running

ISBN 978-7-121-12408-2

I. ①H… II. ①皮… ②常… ③胡… ④赵… III. ①超文本标记语言，HTML—程序设计 IV. ①TP312

中国版本图书馆 CIP 数据核字 (2010) 第 227311 号

策划编辑：卢鹏翔

责任编辑：杨绣国

项目管理：杨绣国

印刷：北京市天竺颖华印刷厂

装订：三河市鑫金马印装有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱

邮编：100036

开本：787×980 1/16 印张：14 字数：280 千字

印次：2010 年 12 月第 1 次印刷

定价：45.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。服务热线：(010) 88258888。

O'Reilly Media, Inc.介绍

为了满足读者对网络和软件技术知识的迫切需求，世界著名计算机图书出版机构 O'Reilly Media, Inc. 授权电子工业出版社，翻译出版一批该公司久负盛名的英文经典技术专著。

O'Reilly Media, Inc. 是世界上在 Unix、X、Internet 和其他开放系统图书领域具有领导地位的出版公司，同时也是在线出版的先锋。

从最畅销的“The Whole Internet User's Guide & Catalog”（被纽约公共图书馆评为20世纪最重要的50本书之一）到GNN（最早的Internet 门户和商业网站），再到 WebSite（第一个桌面 PC 的Web服务器软件），O'Reilly Media, Inc. 一直处于Internet发展的最前沿。

许多书店的反馈表明，O'Reilly Media, Inc. 是最稳定的计算机图书出版商——每一本书都一版再版。与大多数计算机图书出版商相比，O'Reilly Media, Inc. 具有深厚的计算机专业背景，这使得O'Reilly Media, Inc. 形成了一个非常不同于其他出版商的出版方针。O'Reilly Media, Inc. 所有的编辑人员以前都是程序员，或者是顶尖级的技术专家。O'Reilly Media, Inc. 还有许多固定的作者群体——他们本身是相关领域的技术专家、咨询专家，而现在编写著作，O'Reilly Media, Inc. 依靠他们及时地推出图书。因为 O'Reilly Media, Inc. 紧密地与计算机业界联系着，所以O'Reilly Media, Inc. 知道市场上真正需要什么图书。

知识
舟
PDG

目录

Table of Contents

前言.....	1
第 1 章 从开始到现在.....	1
1.1 引言.....	1
1.2 MIME 类型.....	1
1.3 很长的题外话：一份标准是如何诞生的？.....	2
1.4 未曾间断的路线.....	7
1.5 HTML 发展史：从 1997 到 2004 年.....	9
1.6 你所知道的关于 XHTML 的一切都是错误的.....	10
1.7 一个竞争愿景.....	11
1.8 WHAT 工作小组？.....	12
1.9 回到 W3C.....	13
1.10 后记.....	14
1.11 扩展阅读.....	14
第 2 章 HTML5 特性检测.....	15
2.1 引言.....	15
2.2 检测技术.....	15
2.3 Modernizr：一个 HTML5 特性检测库.....	16
2.4 画布.....	16
2.5 画布文本.....	17
2.6 视频.....	18
2.7 视频格式.....	19
2.8 本地存储.....	21
2.9 Web Workers.....	23
2.10 离线 Web 应用.....	23
2.11 地理位置.....	24
2.12 输入框类型.....	25
2.13 占位文本.....	27
2.14 表单自动聚焦.....	27
2.15 微数据.....	28

2.16 扩展阅读	29
第 3 章 从这一切的含义	31
3.1 引言	31
3.2 文档类型	31
3.3 根元素	33
3.4 <head>元素	34
3.4.1 字符编码	35
3.4.2 朋友和 (链接) 关系	36
3.5 HTML5 中新增的语义元素	41
3.6 题外话: 浏览器如何处理未知元素	42
3.7 页头	45
3.8 文章	47
3.9 日期和时间	49
3.10 导航	51
3.11 页脚	52
3.12 扩展阅读	55
第 4 章 Canvas 绘图	57
4.1 引言	57
4.2 简单的图形	58
4.3 Canvas 坐标系	60
4.4 路径	61
4.5 文本	63
4.6 颜色渐变	67
4.7 图片	70
4.8 IE 怎么办?	73
4.9 一个完整的例子	75
4.10 扩展阅读	79
第 5 章 网络上的视频	81
5.1 前言	81
5.2 视频容器	81
5.3 视频编解码器	83
5.3.1 H.264	84
5.3.2 Theora	84
5.3.3 VP8	85
5.4 音频编解码器	85
5.4.1 MPEG-1 音频层 3	86
5.4.2 高级音频编码	87
5.4.3 Vorbis	87
5.5 在网页中怎么工作	88
5.6 H.264 视频的授权问题	90

5.7	使用 Firefogg 编码 Ogg 视频.....	91
5.8	使用 ffmpegtheora 批量编码 Ogg 视频.....	98
5.9	使用 HandBrake 编码 H.264 视频.....	100
5.10	使用 HandBrake 批量编码 H.264 视频.....	107
5.11	使用 ffmpeg 编码 WebM 视频.....	108
5.12	最后, 标记.....	110
5.12.1	MIME 类型很重要.....	113
5.13	IE 怎么办?	114
5.14	完整的例子.....	114
5.14	扩展阅读.....	115
第 6 章	地理位置.....	117
6.1	引言.....	117
6.2	地理位置 API.....	117
6.3	代码展示.....	118
6.4	容错处理.....	120
6.5	方案! 我要方案!	121
6.6	那 IE 怎么办?	123
6.7	geo.js 来拯救.....	123
6.8	一个完整的例子.....	125
6.9	扩展阅读.....	126
第 7 章	Web 应用本地存储的过去、现在和未来.....	127
7.1	引言.....	127
7.2	HTML5 之前的伪本地存储简史.....	128
7.3	HTML5 存储介绍.....	129
7.4	使用 HTML5 存储.....	130
7.4.1	跟踪 HTML5 存储区的改动.....	131
7.4.2	现有浏览器的局限性.....	132
7.5	HTML5 存储实践.....	132
7.6	超越键值对的存储形式.....	134
7.7	扩展阅读.....	135
第 8 章	离线 Web 应用.....	137
8.1	引言.....	137
8.2	缓存清单.....	138
8.2.1	“网络”段.....	139
8.2.2	“默认”部分.....	140
8.3	事件流.....	141
8.4	调试的艺术——杀了我! 现在就杀了我!	142
8.5	让我们来构建一个离线 Web 应用!	145
8.6	扩展阅读.....	146

第9章 疯狂的表单	147
91 引言	147
92 占位文本	147
93 自动聚焦	148
94 Email 地址	150
95 Web 地址	151
96 数字类型输入框：数字选择器	153
97 数字类型输入框：滑块	155
98 日期选择器	156
99 搜索框	158
910 颜色选择器	160
911 还有一点	160
912 扩展阅读	161
第10章 “分布式”、“可扩展性”及其他华丽词藻	163
101 引言	163
102 什么是微数据？	164
103 微数据的数据模型	165
104 标注“人”	168
104.1 Google Rich Snippets 介绍	174
105 标注“组织”	176
106 标注“事件”	180
106.1 Google Rich Snippets 的回归	184
107 标注“点评”	186
108 扩展阅读	190
附录 A 全方位特性检测指南	191
元素列表	191
扩展阅读	199
索引	201

1.1 引言

Diving In

前不久，我偶然读到某 Mozilla 开发人员的一个观点，内容是关于在创建一种标准的过程中存在的冲突 (<http://lists.w3.org/Archives/Public/public-html/2010Jan/0107.html>)：

一份技术规范和他的具体实现必须要做到步调一致。实现先于规范完成不是什么好事情，因为人们会开始依赖这些已实现的细节，这样会对规范形成制约。然而，你也不希望在规范已经完成时还没有任何相关的具体实现和实践经验，因为这样规范就得不到任何反馈。这里面如果存在着无法避免的冲突，而我们也只能硬着头皮去克服了。

请你先牢记这番话，接下来我开始讲述 HTML5 是如何横空出世的。

1.2 MIME 类型

MIME Types

本书的内容是关于 HTML5，不是旧版本的 HTML，也不是任何版本的 XHTML。但要理解 HTML5 的历史及其背后的发展动力，你需要先了解一些技术细节。具体来说，就是 MIME 类型。

每当 Web 浏览器请求一个页面时，Web 服务器会在发送实际页面内容之前，先发送一些头信息 (header)。这些头信息通常是不可见的，当然如果你有兴趣，也可以借助许多 Web 开发工具来查看到这些信息。看不到不代表不重要，事实上，浏览器需要这些信息来决定如何解析随后的页面内容。最重要的头信息是 Content-Type，比如：

```
Content-Type: text/html
```

“text/html”即是这个页面的“内容类型 (content type)”，或者称作“MIME 类型”。这条头信息将唯一确定某个资源本质上是什么，也因而决定了它应该如何被呈现。图片文件有自己的 MIME 类型 (JPEG 图像的类型是 image/jpeg, PNG 图像则是 image/png, 等等)。JavaScript 文件有它自己的 MIME 类型。CSS 样式表也有自己的 MIME 类型。任何

资源都有自己的 MIME 类型。整个 Web 都依靠 MIME 类型运作。

当然，现实情况比这还要复杂得多。非常早期的 Web 服务器（我指的是 1993 年）并不会发送 Content-Type 头信息，因为它当时压根还不存在。（1994 年以前还没有“内容类型”这一概念。）为了兼容到这样的情况，某些流行的浏览器还会在一些特定情况下忽略 Content-Type 头信息。（这一行为被称为“内容嗅探”。）但作为一个一般性的经验法则，凡是在一个网页上能看到的东​​西，比如 HTML，图片，脚本，视频，PDF 文件，所有这些可以用一个 URL 地址定位的东西，都具有特定的 MIME 类型——在 content-type 头信息中指明。

把上面这些内容先塞到脑子里吧。在后面还会复习到。

1.3 很长的题外话：一份标准是如何诞生的？

A Long Digression into How Standards Are Made

为什么会存在

“公开”地制定标准有一大好处，你可以重返过去，寻找上述问题的答案。与之相关的讨论都会出现在邮件列表里，这些邮件通常留有存档，并且可以被公开搜索。所以，我决定做一下对电子邮件的考古，来尝试回答这个关于

在 1993 年 2 月 25 日，马克·安德雷森（Marc Andreessen）写道：¹

我想提议一个新的可选 HTML 标签：

IMG

其必需参数是 SRC="url"。

它指明一个位图或像素图文件，让浏览器尝试通过网络获取并解析为一幅图像，显示于页面上标签所在的位置。

例如：

```
<IMG SRC="file:///foobar.com/foo/bar/blargh.xbm">
```

（没有结束标记，它是一个独立标签。）

¹译注：<http://1997.webhistory.org/www.lists/www-talk.1993q1/0182.html>。点击“下一条消息”和“上一条消息”的链接可以跟踪阅读这份长达数页的讨论。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：《HTML5揭秘》（中文版）Mark Pilgrim.pdf

请登录 <https://shgis.com/post/1816.html> 下载完整文档。

手机端请扫码查看：

