

HTML & XHTML: The Definitive Guide

第六版



HTML & XHTML

权威指南

Chuck Musciano & Bill Kennedy 著

张洪涛 邢璐 译

清华大学出版社

O'REILLY®



第六版

HTML & XHTML权威指南

Chuck Musciano & Bill Kennedy 著

张洪涛 邢璐 译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

O'Reilly Media, Inc. 授权清华大学出版社出版

清华大学出版社

Copyright ©2007 by O'Reilly Media, Inc.

Authorized Simplified Chinese translation edition, by O'Reilly Media, Inc., is published by Tsinghua University Press, 2007. Authorized translation of the original English edition, 2007 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.

本书之英文原版由 O'Reilly Media, Inc. 于 2007 年出版。

本中文简体翻译版由 O'Reilly Media, Inc. 授权清华大学出版社于 2007 年出版。此翻译版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有，未经书面许可，本书的任何部分和全部不得以任何形式复制。

北京市版权局著作权合同登记

图字：01-2006-7219 号

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。

版权所有，侵权必究。侵权举报电话：010-62782989 13501256678 13801310933

图书在版编目 (CIP) 数据

HTML & XHTML 权威指南：第 6 版 / (美) 穆西亚诺 (Musciano, C.), (美) 肯尼迪 (Kennedy, B.) 著；张洪涛，邢璐译. —北京：清华大学出版社，2007. 4

书名原文：HTML & XHTML: The Definitive Guide, Sixth Edition

ISBN 978-7-302-14693-3

I. H… II. ①穆… ②肯… ③张… ④邢… III. 超文本标记语言, HTML、XHTML—程序设计 IV. TP312

中国版本图书馆 CIP 数据核字 (2007) 第 021298 号

责任编辑：常晓波

封面设计：Edie Freedman, 张健

责任校对：张 剑

责任印制：王秀菊

出版发行：清华大学出版社

<http://www.tup.com.cn>

c-service@tup.tsinghua.edu.cn

社总机：010-62770175

投稿咨询：010-62772015

地址：北京清华大学学研大厦 A 座

邮编：100084

邮购热线：010-62786544

客户服务：010-62776969

印刷者：北京市世界知识印刷厂

装订者：三河市李旗庄少明装订厂

经销：全国新华书店

开本：178 毫米×233 毫米 42.25 印张 920 千字

版次：2007 年 4 月第 1 版

印次：2007 年 4 月第 1 次印刷

印数：1~3000(册)

定 价：75.00 元(册)

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。

联系电话：010-62770177 转 3103

产品编号：024468-01

封面介绍

本书封面上的动物是考拉。考拉是澳大利亚的一种有袋动物，是 *Phascolarctidae* 科的唯一成员。

考拉在刚刚生下来时很小，大约只有 0.5 克。幼年的考拉会在母亲的育儿袋里待上大约 7 个月的时间。与大多数有袋动物不同，考拉的育儿袋的开口朝向臀部，而不是朝向头部。考拉的死亡率很高，而且由于 1887 年—1889 年和 1900 年—1903 年间的流行病，以及整个 20 世纪不受限制的捕猎行动，考拉已经接近灭绝了。它们是一种受保护的物种，并且正在重建它们的种群，但到目前为止它们还仅仅只在澳大利亚东部存活。

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. 授权清华大学出版社，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，同时也是联机出版的先锋。

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。

作者简介

Chuck Musciano (cmusciano@aol.com) 在1982年从佐治亚理工学院计算机科学系获得了学士学位。在接下来的15年里,他受雇于佛罗里达州墨尔本市的Harris公司,刚开始是编译器设计师,后来成为Harris先进技术组的一员。他对基于Unix和Internet技术方面的关注使他在Harris公司得以支持早期的Web和因特网的初创。在1997年,他荣升为 American Kennel Club 的首席信息官,在这里工作时他将前沿技术与激情结合了起来。在Chuck的职业生涯中,他已经为各种商业性出版物编写了很多文章,不管是在报纸上还是作为在线的专栏作家,其中包括作为Sunworld的Webmaster的专栏作家,以及WebReview的Tag of the Week的专栏作家。在业余时间,他和他的妻子Cindy、女儿Courtney和儿子Cole在北卡罗莱纳州享受着家庭生活的乐趣。

Bill Kennedy (bkennedy@activmedia.com) 目前是ActivMedia Robotics有限公司的总裁和首席技术官。他积极投身于人工智能机器人方面的研究,这为他的事业带来了许多迂回。Bill在芝加哥的Loyola大学获得了生物化学和生物物理学的博士学位。从20世纪70年代以来,他一直在进行生物化学方面的研究;80年代初期,他开办了一家软件公司,专门开发计算机游戏和教育软件(他有很多古老的PC机用来测试);他在80年代和90年代还任职于各种因特网数据组杂志的技术编辑、编审和总编,其中包括inCider、Sun Technology Journal、Sun World和A+Publishing/Mac Computing。自从80年代以来,Bill与他的妻子Jeanne都是热心的因特网用户,他们开办了ActivMedia Research,这是第一家正式研究因特网上商业进化的一家市场调查公司。Bill同他的合作伙伴总是准备着要采用新兴技术,因此他们在1995年成立了ActivMedia Robotics (<http://www.activrobots.com>),这家公司为全世界研究人工智能机器、智能机器人的研究人员装配了更为智能的移动机器人。下一个会是什么?

目录

前言	1
第 1 章 HTML, XHTML 和万维网	9
1.1 因特网	9
1.2 遵循因特网规则	12
1.3 HTML 和 XHTML 的历史	16
1.4 HTML 和 XHTML 的局限性	17
1.5 标准和扩展	18
1.6 Web 设计工具	20
第 2 章 快速启动	23
2.1 编写工具	23
2.2 一个最简单的 HTML 文档	24
2.3 内嵌的标签	25
2.4 HTML 的“骨架”	26
2.5 HTML 和 XHTML 文档的“皮肉”	27
2.6 文本	28
2.7 超链接	33

2.8 特殊的图像	36
2.9 列表、可搜索文档和表单	38
2.10 表格	40
2.11 框架	40
2.12 样式表和 JavaScript	42
2.13 继续学习	43
第 3 章 剖析 HTML 文档	44
3.1 外观的欺骗性	44
3.2 HTML 文档的结构	45
3.3 标签和属性	46
3.4 格式良好的文档和 XHTML	50
3.5 文档的内容	51
3.6 HTML/XHTML 文档的元素	52
3.7 文档的头部	56
3.8 文档的主体	59
3.9 编辑标记	61
3.10 <bdo> 标签	64
第 4 章 文本基础	66
4.1 部分和段落	66
4.2 标题	74
4.3 改变文本的外观和含义	79
4.4 基于内容的样式标签	81
4.5 物理样式标签	88
4.6 精确的间距和布局	93
4.7 块引用	106
4.8 地址	110
4.9 特殊字符的编码	112
4.10 HTML 中已废弃的扩展字体处理	113

第 5 章 分隔线、图像和多媒体	120
5.1 水平分隔线	120
5.2 把图像插入文档	128
5.3 文档颜色和背景图像	155
5.4 背景音乐	162
5.5 活动文字	164
5.6 其他多媒体内容	168
 第 6 章 链接和 Web	 171
6.1 超文本基础	171
6.2 引用文档：URL	172
6.3 创建超链接	189
6.4 创建有效的链接	198
6.5 对鼠标敏感的图像	202
6.6 创建可搜索文档	214
6.7 关系	217
6.8 支持文档自动化	222
 第 7 章 格式化列表	 225
7.1 无序列表	225
7.2 有序列表	228
7.3 标签	232
7.4 嵌套列表	235
7.5 定义列表	237
7.6 正确的列表用法	241
7.7 目录列表	241
7.8 菜单列表	243
 第 8 章 级联样式表 (CSS)	 245
8.1 样式元素	246

8.2 样式语法	255
8.3 样式类	260
8.4 样式属性	267
8.5 标签内样式: <code></code> 标签	322
8.6 对文档应用样式	323
第 9 章 表单	327
9.1 表单简介	328
9.2 <code><form></code> 标签	329
9.3 一个简单的表单示例	337
9.4 利用电子邮件收集表单数据	338
9.5 <code><input></code> 标签	340
9.6 <code><button></code> 标签	351
9.7 多行文本区	353
9.8 多选元素	355
9.9 一般表单控件属性	360
9.10 给表单元素标记和分组	363
9.11 创建有效的表单	366
9.12 表单编程	371
第 10 章 表格	377
10.1 标准表格模型	377
10.2 基本表格标签	379
10.3 高级表格标签	397
10.4 普通表格之外的内容	407
第 11 章 框架	409
11.1 框架概述	409
11.2 框架标签	410
11.3 框架布局	412

11.4 框架内容	417
11.5 <noframes> 标签	420
11.6 内联框架	422
11.7 命名的框架或窗口目标	424
11.8 XFrames	428
第 12 章 可执行的内容	431
12.1 applet 和对象	431
12.2 嵌入式内容	434
12.3 JavaScript	450
12.4 JSS (过时的)	458
第 13 章 动态文档	465
13.1 动态文档概述	465
13.2 客户端拉文档	466
13.3 服务器端推文档	471
第 14 章 移动设备	475
14.1 移动 Web	475
14.2 设备考虑事项	477
14.3 XHTML Basic	479
14.4 高效移动 Web 设计	482
第 15 章 XML	489
15.1 语言和元语言	490
15.2 文档和 DTD	492
15.3 理解 XML DTD	494
15.4 元素语法	498

15.5 元素属性	502
15.6 条件段	505
15.7 构建 XML DTD	506
15.8 使用 XML	507
第 16 章 XHTML	510
16.1 为什么要用 XHTML	511
16.2 创建 XHTML 文档	512
16.3 HTML 与 XHTML	516
16.4 XHTML 1.1	521
16.5 应该使用 XHTML 吗	523
第 17 章 技巧、诀窍和奥妙	526
17.1 重要的技巧	526
17.2 使用 HTML 编辑器之后进行清理	528
17.3 有关表格的诀窍	533
17.4 有关窗口和框架的诀窍	538
附录 A HTML 语法	543
附录 B HTML/XHTML 标签快速参考	555
附录 C CSS 属性快速参考	580
附录 D HTML 4.01 DTD	590
附录 E XHTML 1.0 DTD	609
附录 F 字符实体	628
附录 G 颜色名和值	635
附录 H Netscape 布局扩展	639

前言

学习超文本标记语言（Hypertext Markup Language, HTML）和可扩展超文本标记语言（Extensible Hypertext Markup Language, XHTML）与学习其他计算机语言或人类语言一样。大多数学生一开始会沉浸于大量的示例当中，因为学习别人是一种很自然的学习方式，这样可以使学习变得轻松而有趣。对于任何希望学习HTML及XHTML的人来说，我们的建议是到网上去，选择一个合适的浏览器，去看看什么东西不错，什么东西非常有效，什么东西是你可以拿来用的。研究别人的文档并观察哪些是可以实现的，哪些是不可实现的。现在许多网站管理员学习这种方法就是模仿。

但是，模仿也只能带你走这么远。示例既有优点，也存在缺点。通过示例来进行学习，只可以学到表面的东西，但却不能彻底全面地掌握。要想做到真正熟练，必须学会在各种不同的情况下使用这种语言。当然，如果时间足够多的话，也可以通过示例来学习所有这些不同情况。

另外还要记住，计算机语言比人类语言精确得多。必须正确使用语法，否则它就不会起到应有的作用。所以我们就会遇到“标准”的问题。学术委员会和业界的专家制定了类似HTML等计算机语言的语法和使用方法。问题是你与你的受众所使用的显示文档的浏览器技术可能是基于不同的标准。他们甚至自己又定义了某些部分，把标准完全不当一回事。

标准本身也在不断地改变。HTML正在经历向XHTML的转换，这种转换使它成为XML（Extensible Markup Language，可扩展标记语言）的一种应用。HTML和XHTML非常相似，以至于我们经常把它们当做同一种语言来对待。但它们却有着非常重要的区别，我们将在后面更多地阐述这些区别。

出于安全方面的考虑，要想熟练地掌握HTML和XHTML，就要学习全面的、最新的语

言参考书，这些资料要详细介绍该语言的语法、语义以及各种变化，来帮助读者区分好的和不好的用法。

要想熟练使用语言还需要另外一个步骤。要想成为该语言真正的专家就必须形成自己的风格。这就意味着不仅要知道什么是合适的，更要知道什么是有效的。还要考虑布局方面的很多问题，以及文档内、文档间和文档集之间内容的表示顺序。

本书的目的正是为了帮助读者熟练地掌握 HTML 和 XHTML，完全精通它们的语法、语义和样式方面的要素。我们采用自然的学习方法，也就是使用好的示例。本书包括了现在最常用的标准语言版本（HTML 4.01 和 XHTML 1.0）方方面面的内容，以及现在所有流行浏览器接受的语言扩展。我们将解释这些要素的工作方式，以及它们如何与其他要素相互作用。

而且，本书为读者提出了样式和组织方面的建议，以帮助读者决定怎样才能最好地利用 HTML 和 XHTML 来完成各种各样的任务：从简单的在线文档到复杂的市场营销文案。我们将告诉读者哪些是可行的，哪些则不行，哪些对网页浏览者来说是清楚的，而哪些可能会让他们变得糊涂。

简而言之，本书是一本全面讲述使用 HTML 和 XHTML 创建文档的指南。它深入浅出，从最基本的语法和语义入手，以丰富的样式准则结束，这些内容可以帮助你创建内容漂亮、丰富而且可访问性良好的文档，当读者在浏览器中浏览这些文档时，你将会为这些文档感到自豪。

我们的读者对象

本书针对任何对学习 Web 语言感兴趣的读者，包括一般的使用者和专业网页设计人员。在拿起这本书之前，你不需要具备任何 HTML 或者 XHTML 方面的知识。实际上我们甚至都不指望你有过上网浏览的经验，虽然这有些让人惊讶。本书并不要求你的计算机一定要和因特网相连，不过，如果确实没有这种连接的话，本书就有点像一本旅游指南了。

我们只要求读者有一台计算机，一个能编辑纯文本文件的编辑器和一个最新的 Web 浏览器。我们使用了 Internet Explorer、Netscape Navigator、Mozilla Firefox 和 Opera Software ASA 的 Opera 浏览器。因为 HTML 和 XHTML 文档是以一种被广泛接受的格式（纯文本）存储的，并且因为这种语言和使用的计算机完全无关，因此我们毫不关心你使用的是什么样的计算机。然而浏览器却会因平台和操作系统的不同而有着很大的差异，这就意味着 HTML 或 XHTML 文档可能会因为计算机或者浏览器版本的不同而在显示效果方面相去甚远。我们将解释各种浏览器如何使用一定的语言特性，其中特别要注意它们之间的不同所在。

如果读者对 HTML、Web 或者超文本文档都不熟悉，那么应该从第 1 章开始学习。第 1 章将描述所有这些技术如何相互结合使用，共同创建由互相关联的文档构成的网站。

如果读者对 Web 很熟悉，但是并不熟悉 HTML 和 XHTML，那么请从第 2 章开始。这一章将简要地介绍这种语言的重要特性，指引读者如何在本书后面学习这门语言。

其他章节全面介绍了 HTML 和 XHTML 语言的各种特性。读者可以按照章节的顺序从头到尾阅读，这样可以对本语言有一个全面的了解，也可以选择性地阅读，寻找自己感兴趣的内容。

排版约定

本书使用等宽字体 (constant-width) 来突出表示所有 HTML/XHTML 标准、标签和属性的文字部分。使用小写字体表示标签 (注 1)。在表明定义文件名和新的概念时，则用斜体 (italic) 表示。创建文档时，在代码中用等宽斜体 (constant-width italic) 表示需要提供的元素，例如标签属性或者由用户定义的字符串等。

本书讨论了语言的各种元素，你会发现每一个元素都在简写形式的快速参考定义框中有详细深入的介绍，就像下面的 <title> 定义框一样。该框中的第一行是元素的名称，然后是其功能的简述。再后，如果元素有各种不同的属性，我们将把它们都列出来：这些属性是指元素的一部分，可能需要也可能必须进行指定。

<title>	
功能	定义文档标题
属性	dirlang
结束标签	</title>; 从不省略
包含	plain_text
使用范围	head_content

图标说明此标签和属性不属于 HTML 4.01 和 XHTML 1.0 标准，并且在各种流行浏览器中处理这些是非常困难的。

注 1: HTML 对于标签和属性名是不区分大小写的，而 XHTML 则区分。但 HTML 中的某些项目也是区分大小写的，比如源文件的名称

如果有结束标签的话,“说明”中还会包括元素的结束标签,以及在 HTML 的一般应用中省略这个结束标签是否安全的说明。少数几个标签在 HTML 中没有结束标签,但在 XHTML 中却要求必须有,这时可以用一个正斜杠 (/) 来结束这个标签,如 `<br />`。在这种情况下,标签还可以包含很多属性,可用省略号来表示,如 `<br ... />`。

其中的“包含”部分列出了 HTML 的语法规则,这些语法规则定义了在本标签中使用的元素。类似地,“使用范围”列出了其他可以使用本标签的语法规则。这些规则将在附录 A 中进行定义。

最后要说的是,HTML 和 XHTML 是两种关系紧密的语言。有时根据内容的不同,某些元素的使用也会不同,但很多元素都具有共同的特性。所以在所有可能的地方,我们都会放置交叉参考,这样读者可以很容易地找到本书中其他相关的讨论。这些交叉参考和本段末尾一样,就像超文本文档的纸质版本。如果本书是以电子版形式发布的话,就肯定要使用超文本链接来代替它们。[请参见第 3.3.1 小节“标签语法”]

我们鼓励读者在任何可能的时候使用这些参考。很多时候,我们在一个地方只是很简要地描述一种属性,因为我们希望读者能够在其他参考中找到足够详细的描述。而在另外一些时候,使用交叉参考可以了解到同一种元素的不同使用方法,或者其他与本元素相关的使用和样式方面的建议。

版本和语义

HTML 最新的版本是 4.01,但语言主要的更新和变动是在 4.0 版加进去的。所以在本书中我们通常都说 HTML 4,即包括 4.0 及之后的版本。只有在必要时才会提及“点版本号”。

XHTML 标准现在还是它的第一个版本:1.0。万维网联盟(W3C)已经提出了第二个版本(XHTML 2.0)的草案,但还没有正式建立该标准。很大程度上来说,XHTML 1.0 和 HTML 4.01 是相同的。在第 16 章中我们将详细介绍它们的区别。在本书中,如果 XHTML 在处理某特性或元素时与 HTML 有什么不同的话,我们会明确指出。

HTML 和 XHTML 标准明确地区分了文档中的“元素类型(element type)”和界定这些元素的“标签(tag)”。例如,标准所指的段落元素类型,与段落标签 `<p>` 是不同的。段落元素包括在段落开始标签(`<p>`)和段落结束标签(`</p>`)中间可以接受的元素类型名和插入的内容。标签 `<p>` 是段落元素的起始标签,而段落的内容,即属性,才最终决定了段落元素类型的内容。

虽然这些区别十分关键,但毕竟创作者在文档中使用的是标记标签,而这些标签直接影

响着内容，所以本书并没有特意去区分元素类型和标签。最经常讨论的是标签及其相关的内容，但并不一定使用“元素类型”这个术语，虽然从技术角度讲，进行区分可能更合适。请原谅我们违反了规则，不过我们确实是为了要说明得更清楚才这样做的。

HTML 与 XHTML

实际上，就版本 4.01 而言，HTML 这门语言正值其成熟期。W3C 打算不再开发另一个版本，并且已经正式宣布过此类说法。而且，HTML 已经模块化并被包含进可扩展标记语言（Extensible Markup Language, XML）当中了。它的新名称是 XHTML，即可扩展超文本标记语言（Extensible Hypertext Markup Language）。

XHTML 的出现只不过是 HTML 和 Web 轰轰烈烈的历史中的又一崭新的篇章。在这一历史进程中，使用者感到困惑并不新鲜，而且简直就是司空见惯的事情。在最糟糕的时候，万维网联盟（W3C）中负责已经接受和可以接受使用的语言（也就是标准）的老前辈们，在 Netscape 公司和 Microsoft 公司的浏览器“大战”中失去了对语言的控制。已流产的 HTML+ 标准从来都没有成型过，而 HTML 3.0 之争也陷入了困境，以至于 W3C 只能把整个标准的草稿束之高阁。除了在某些机会主义分子的文章中，HTML 3.0 从来都没有出现过。直到 1996 年底，浏览器厂商才说服 W3C 发布 HTML 标准的 3.2 版本，这个版本的目的只不过是把 Netscape 的浏览器对 HTML 的大部分扩展进行了标准化。

Netscape 在浏览器市场以及 Web 技术中的主导地位于上世纪末结束。之后，Microsoft 有效地将 Internet Explorer 捆绑到 Windows 操作系统中，不仅仅作为安装的应用程序，而是作为 GUI 桌面系统的主要特征。而且，Internet Explorer 还具有几个特征（当时还不是标准）受到了 Internet 界和市场的普遍赞赏。

对于我们当中那些喜欢并且强烈支持标准的人来说，幸运的是，到 HTML 4.0 版本时，W3C 又恢复了它的重要地位。这个版本就是今天我们所说的版本 4.01，它于 1999 年 12 月发布。这个版本吸取了 Netscape 和 Internet Explorer 革新中的许多方面，比以前的任何一个版本都要更清晰和整洁，它为不同浏览器和不同平台之间的一致性建立了坚实的实现模型，它支持并鼓励使用基于 HTML 显示的 CSS 标准，同时还制定了对其他（非图形化）用户代理程序（user-agent）的规定，并提供了更加广泛的语言支持。

除了更加清晰和整洁之外，W3C 还认识到，HTML 永远也不可能跟上 Web 对发布、处理和显示文档方法的更多要求。HTML 只是提供了一套最基本的文档创建原语（primitive），当遇到像化学公式、音乐标注或者数学表达式之类的非传统文档类型时，它就无能为力。而且它也不能支持其他一些显示设备，像手持计算机、智能蜂窝电话等。