04 | HTML语义:如何运用语义类标签来呈现Wiki网页?

2019-01-24 winter



你好,我是winter,今天我们继续来聊聊HTML模块的语义类标签。

在上一篇文章中,我花了大量的篇幅和你解释了正确使用语义类标签的好处和一些场景。那么,哪些场景适合用到语义类标签呢,又如何运用语义类标签呢?

不知道你还记不记得在大学时代,你被导师逼着改毕业论文格式的情景,如果你回想一下,你在 论文中使用的那些格式,你会发现其实它们都是可以用**HTML**里的语义标签来表示的。

这正是因为HTML最初的设计场景就是"超文本",早期HTML工作组的专家都是出版界书籍排版的专家。

所以,在这一部分,我们找了个跟论文很像的案例:wikipedia文章,这种跟论文相似的网站比较适合用来学习语义类标签。通过分析一篇wiki的文章用到的语义类标签,来进一步帮你理解语义的概念。

你可以在电脑上,打开这个页面: https://en.wikipedia.org/wiki/World_Wide_Web

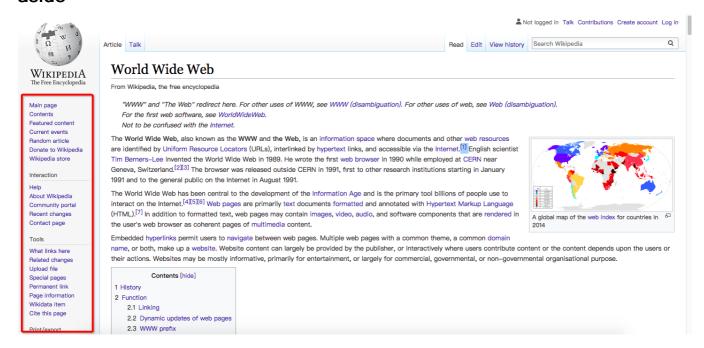
(为了防止这个页面被修改,我们保存了一个副本:

http://static001.geekbang.org/static/time/quote/World_Wide_Web-Wikipedia.html

这是一篇我们选择的wiki文章,虽然在原本的wikipedia网站中,也是大量使用了div和span来完成

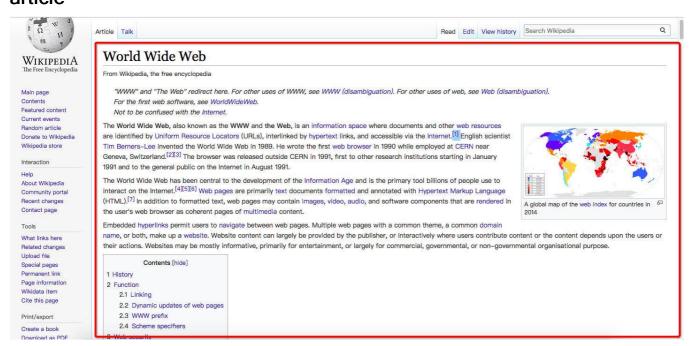
功能。在这里,我们来尝试分析一下,应该如何用语义类标签来呈现这样的一个页面/文章。 我们看一下这个页面。

aside



首先我们来看下,左侧侧边栏,根据上一篇文章中提到的语义定义,这里属于**aside**内容。是导航性质的工具内容。

article



我们来到文章主体部分,因为主体部分具有明确的独立性,所以可以用article来包裹。

hgroup, h1, h2

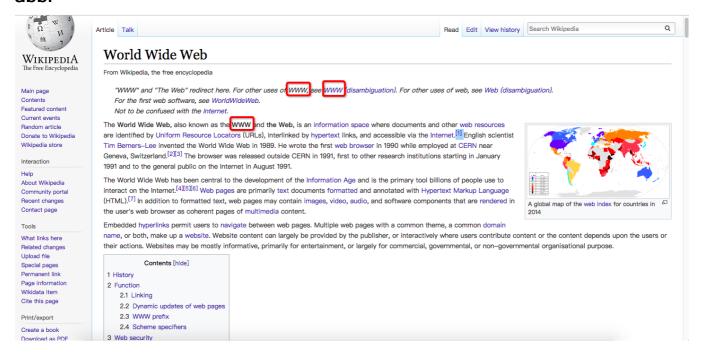


在语义的上一篇文章中,我们介绍过hgroup和h1-h6的作用,hgroup是标题组,h1是一级标题,h2是二级标题。这里,World Wide Web 是文章的大标题,适合h1元素。

接下来出现了一个副标题。From Wikipedia, the free encyclopedia。这个地方适合使用h2,跟h1 组成一个hgroup,所以代码可能是类似这样的:

<hgroup>
<h1>World Wide Web </h1>
<h2>From Wikipedia, the free encyclopedia</h2>
</hgroup>

abbr



<abbr title="World Wide Web">WWW</abbr>.

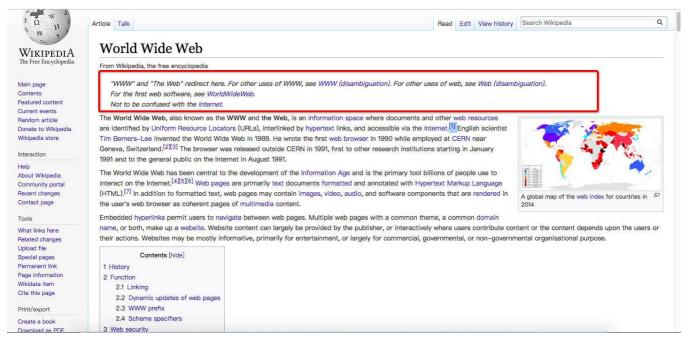
hr



细心的同学会发现,在wiki的界面中,出现了一条很长的横线,大家都知道hr标签表示横向分隔线,那么这个地方是不是应该用hr呢?

答案是不用。我们读一下标准的定义就知道了,hr表示故事走向的转变或者话题的转变,显然此处两个标题并非这种关系,所以我们应该使用CSS的border来把它当作纯视觉效果来实现,所以这里是不需要用hr的。

p



接下来一段,我们看到了三段"note",也就是注记。它在文章中用作额外注释。

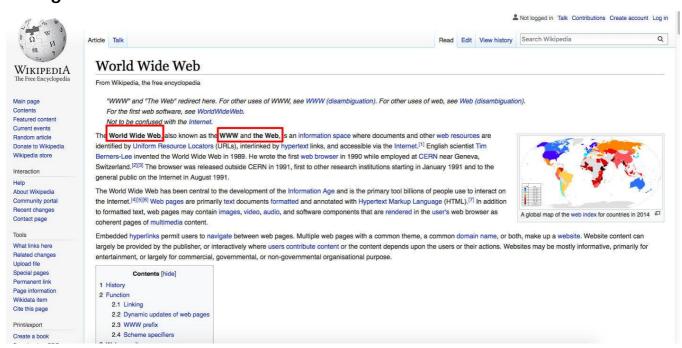
"WWW" and "The Web" redirect here. For other uses of WWW, see WWW (disambiguation). For other uses of web, see Web (disambiguation).

For the first web software, see WorldWideWeb.

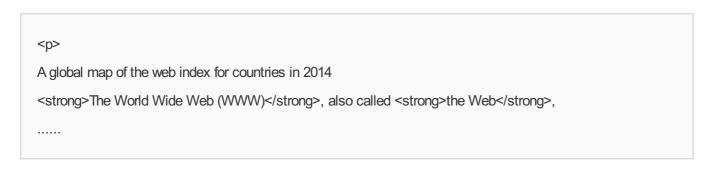
Not to be confused with the Internet.

HTML中并没有note相关的语义,所以,我们用普通的p标签,加上class="note"来实现。后面的多数自然段都是普通的段落,我们用p标签来实现。

strong



注意,这里 "World Wide Web (WWW)" 和 "the Web" 使用了黑体呈现,从上下文来看,这里表示这个词很重要,所以我们使用strong标签。



blockquote, q, cite

World Wide Web



接下来我们看到了一个论文中很常见的用法"引述"。

interlinked by hypertext links, and accessible via the Internet.[1]

注意看这里的[1], 当我们把鼠标放上去的时候, 出现了引述的相关信息:

"What is the difference between the Web and the Internet?". W3C Help and FAQ. W3C. 2009. Archived from the original on 9 July 2015. Retrieved 16 July 2015.

在HTML中,有三个跟引述相关的标签blockquote表示段落级引述内容,q表示行内的引述内容,cite表示引述的作品名。

这里的作品名称 "What is the difference between the Web and the Internet?", 应当使用cite标签。

<cite>"What is the difference between the Web and the Internet?"</cite>. W3C Help and FAQ. W3C. 2009. Archiv

在文章的结尾处,有对应的 References 一节,这一节中所有的作品名称也应该加入cite标签。

References [edit]

- A "What is the difference between the Web and the Internet?" 당. W3C Help and FAQ. W3C. 2009. Archived 당 from the original on 9 July 2015. Retrieved 16 July 2015.
- ^ McPherson, Stephanie Sammartino (2009). Tim Berners-Lee: Inventor of the World Wide Web
 @. Twenty-First Century Books. ISBN 978-0-8225-7273-2. Archived Ø from the original on 15 April 2016.
- 3. ^ Quittner, Joshua (29 March 1999). "Network Designer Tim Berners-Lee" . Time Magazine. Archived . from the original on 15 August 2007. Retrieved 17 May 2010. (Subscription required (help)). "He wove the World Wide Web and created a mass medium for the 21st century. The World Wide Web is Berners-Lee's alone. He designed it. He loosed it on the world. And he more than anyone else has fought to keep it open, nonproprietary and free. "[Jage needed]
- 4. ^ "World Wide Web Timeline" ©. Pew Research Center. 11 March 2014. Archived © from the original on 29 July 2015. Retrieved 1 August 2015.
- Dewey, Caitlin (12 March 2014). "36 Ways the Web Has Changed Us" @. The Washington Post. Archived @ from the original on 9 September 2015. Retrieved 1 August 2015.
- 6. ^ "Internet Live Stats" ☑. Archived ☑ from the original on 2

- 30. ^ "WWW (World Wide Web) Definition" ₽. TechTerms.

 Archived ₽ from the original on 11 May 2009. Retrieved
 19 February 2010.
- 31. A Jacobs, Ian; Walsh, Norman (15 December 2004).
 "Architecture of the World Wide Web, Volume One" & Introduction: W3C. Archived & from the original on 9 February 2015. Retrieved 11 February 2015.
- 32. A "Supplement no.1, Diplomatic and Overseas List, K.B.E." (PDF) thegazette.co.uk. The Gazette. 31 December 2003. Archived (PDF) from the original on 3 February 2016. Retrieved 7 February 2016.
- 33. ^ "Web's inventor gets a knighthood" 량. BBC. 31 December 2003. Archived량 from the original on 23 December 2007. Retrieved 25 May 2008.
- 34. A "What is the difference between the Web and the Internet?" ₽. World Wide Web Consortium. Archived ₱ from the original on 22 April 2016. Retrieved 18 April 2016.
- 35. ^ Muylle, Steve; Rudy Moenaert; Marc Despont (1999). "A grounded theory of World Wide Web search behaviour" £. Journal of Marketing Communications. 5 (3): 143. doi:10.1080/135272899345644 £. Archived & from the original on 20 October 2014.
- 55. ^ O'Reilly, Tim (30 September 2005). "What Is Web 2.0" @. O'Reilly Media. pp. 4–5. Archived @ from the original on 15 April 2013. Retrieved 4 June 2008. and AJAX web applications can introduce security vulnerabilities like "client-side security controls, increased attack surfaces, and new possibilities for Cross-Site Scripting (XSS)", in Ritchie, Paul (March 2007). "The security risks of AJAX/web 2.0 applications" (APPF). Infosecurity. Elsevier. Archived from the original (APPF) on 25 June 2008. Retrieved 6 June 2008. which cites Hayre, Jaswinder S. & Kelath, Jayasankar (22 June 2008). "Ajax Security Basics" @. SecurityFocus. Archived @ from the original on 15 May 2008. Retrieved 6 June 2008.
- 56. ^ Berinato, Scott (1 January 2007). "Software Vulnerability Disclosure: The Chilling Effect" ⊕. CSO. CXO Media. p. 7. Archived from the original ⊕ on 18 April 2008. Retrieved 7 June 2008.
- 57. ^ "2012 Global Losses From phishing Estimated At \$1.5 Bn*d. FirstPost. 20 February 2013. Archived d from the original on 21 December 2014. Retrieved 21 December 2014.
- 58. ^ Prince, Brian (9 April 2008). "McAfee Governance, Risk and Compliance Business Linit" © MEEK 7iff Davis

这里我们看看引用的原文就可以知道,**wiki**文章中的信息并非直接引用,如果是直接引用的内容,那么,我们还应该加上**blockquote**或者**q**标签。

time

World Wide Web



这里除了引用的文章外,还出现了日期,为了让机器阅读更加方便,可以加上time标签:



figure, figcaption



我们注意一下文章的右侧,出现了几张图片,这种出现在文中的图片,不仅仅是一个**img**标签,它和下面的文字组成了一个**figure**的语法现象,**figure**也是我们的一种标签(用于表示与主文章相关的图像、照片等流内容)。

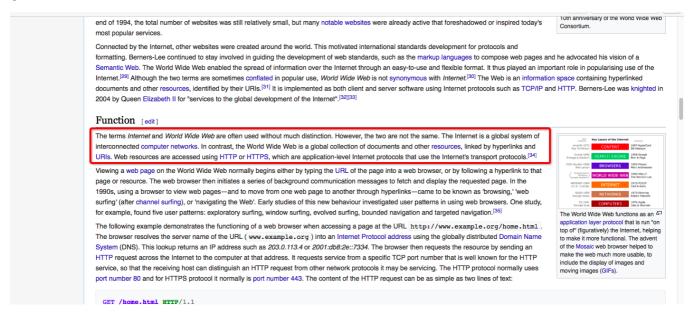
<figure>

<figcaption>The NeXT Computer used by Tim Berners-Lee at CERN.</figcaption>
</figure>

这种插入文章中的内容,不仅限图片,代码、表格等,只要是具有一定自包含性(类似独立句

子)的内容,都可以用**figure**。这里面,我们用**figcaption**表示内容的标题,当然,也可以没有标题。

dfn



然后我们继续往下看,来注意这一句:

The terms Internet and World Wide Web are often used without much distinction. However, the two are not the same. The Internet is a global system of interconnected computer networks. In contrast, the World Wide Web is a global collection of documents and other resources, linked by hyperlinks and URIs.

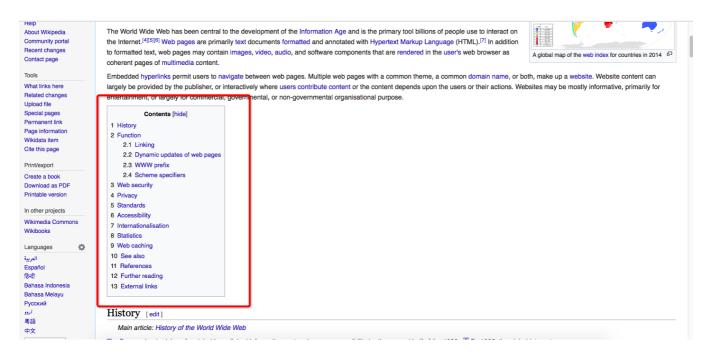
这里分别定义了Internet和World Wide Web,我们应该使用dfn标签。

The terms Internet and World Wide Web are often used without much distinction. However, the two are not the sar The <dfn>Internet</dfn> is a global system of interconnected computer networks.

In contrast, the <dfn>World Wide Web</dfn> is a global collection of documents and other resources, linked by hypersonal collection of documents.

代码中你可以看见,你需要在你要定义的词前后放上**dfn**标签,所以我们知道了,**dfn**标签是用来包裹被定义的名词。

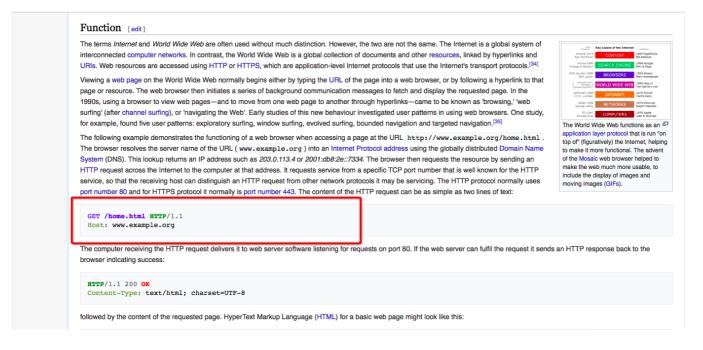
nav, ol, ul



接下来,几个普通的段落之后,我们看到了文章的目录。这里的目录链接到文章的各个章节,我们可以使用nav标签。因为这里的目录顺序不可随意变化,所以我们这里使用多级的ol结构。

我们这里必须要指出,ol和ul的区分是内容是否有顺序关系,每一项的前面不论是数字还是点,都不会影响语义的判断。所以,你可以注意一下这里,不要因为视觉表现效果,而改变语义的使用。

pre, samp, code



继续往下,我们来到了这里,我们看见这篇文章有一个很重要的特色,文章中嵌入了一些代码和一些预先编写好的段落。我们看到在"Function"小节中有一段背景色是灰色的文字。

GET/home.html HTTP/1.1

Host: www.example.org

这是一段HTTP协议的内容描述,因为这段内容的换行是非常严格的,所以我们不需要浏览器帮我们做自动换行,因此我们使用了pre标签,表示这部分内容是预先排版过的,不需要浏览器进行排版。

又因为这是一段计算机程序的示例输出,所以我们可以使用samp标签:

<samp>
GET /home.html HTTP/1.1
Host: www.example.org
</samp>

接下来Wiki中的内容出现了一段HTML代码,我们同样不希望浏览器做自动换行。

```
<html>
<head>
<title>Example.org – The World Wide Web</title>
</head>
<body>
The World Wide Web, abbreviated as WWW and commonly known ...
</body>
</html>
```

因为同时是代码,我们还需要加上**code**标签。最后的代码是**pre**标签包裹了**code**标签,**code**标签包裹了**HTML**代码。

```
<code>
&lt;html&gt;
&lt;head&gt;
&lt;title&gt;Example.org — The World Wide Web&lt;/title&gt;
&lt;/head&gt;
&lt;/body&gt;
&lt;p&gt;The World Wide Web, abbreviated as WWW and commonly known ...&lt;/p&gt;
&lt;/body&gt;
&lt;/html&gt;
</code>
```

在后面的代码中,还有一些在行内的**code**,比如 **title**和 **p**括起来的内容,这些也都应该使用**code** 标签。

总结

在这一篇**Wiki**文章中,已经涉及了大部分语义标签,可见**HTML**工作组对语义标签的增加是非常谨慎和保守的。

当然了,我们选择的案例不可能刚巧覆盖所有的标签,还有些没讲到的标签,我们这里稍微做一下简要的补充说明。

标签	说明
small	之前表示字体缩小的废弃标签,HTML5救回来表示补充评论。
s	之前表示划线的废弃标签,HTML5救回来表示错误的内容,经常用于电商领域表示打折前的价格。
i	之前表示斜体的废弃标签,HTML5救回来表示读的时候变调。
b	之前表示黑体的废弃标签,HTML5救回来表示关键字。
u	之前表示下划线的废弃标签,HTML5救回来表示避免歧义的注记。
data	跟time标签类似,给机器阅读的内容,意义广泛,可以自由定义。
var	变量,多用于计算机和数学领域。
kbd	用户输入,表示键盘按键居多。
sub	下标,多用于化学/物理/数学领域。
sup	上标, 多用于化学/物理/数学领域。
bdi, bdo	用于多语言混合时指定语言或者书写方向(左到右或者右到左)。
mark	表示高亮,这里并非指显示为高亮,而是从读者角度希望的高亮(注意与 strong的区分)。
wbr	表示可以换行的位置,主要是英文等文字不允许单词中间换行,这个标签一般 在把多个单词粘成很长的单词时候用。
menu	ul的变体,用于功能菜单时使用。
dl, dd, dt	一般出现较为严肃的文章,对一些术语进行定义,dt和dd其实并不总是成对出现,两者是多对多的关系。
main	整个页面只出现一个,表示页面的主要内容,可以理解为特殊的div。

(长按点击大图查看)

实际上,HTML这种语言,并不像严谨的编程语言一样,有一条非此即彼的线。一些语义的使用 其实会带来争议,所以我的建议是:你可以尽量只用自己熟悉的语义标签,并且只在有把握的场 景引入语义标签。这样,我们才能保证语义标签不被滥用,造成更多的问题。

你最擅长使用哪些语义标签,会把它们用在哪些场景里呢?欢迎留言告诉我,我们一起讨论。



重学前端

每天10分钟, 重构你的前端知识体系

winter 程劭非前手机淘宝前端负责人



精选留言



连

6 43

就擅长div span a标签I

2019-01-24



Carson

凸 14

由于项目的关系,大部分时候是制作品牌网站和软件界面的场景,因此我最擅长的语义标签是「作为整体结构」的语义类标签。

在这两个场景中,一般不会有大段的阅读文字(即使是 FAQ,也避免大段文字,毕竟没人喜欢读字)。

深度了解语义类标签使用的一个技巧,就是去了解爬虫是如何理解自己的网站,从而逆向理解标签是否使用得当。

这篇文章最大的收获,就是认识到自己对「自然语言延伸」和「标题摘要」的语义类标签理解不足。比如 pre 和 samp,让我有"居然是这样使用"的感觉。

现在知道,如果遇到内容网站的场景(比如博客),我就应该复习一次「自然语言延伸」和「标题摘要」的语义类标签文档,确保正确使用。

谢谢 winter 老师。

2019-01-24





2019-01-24



Brigand

凸 10

微信小程序的wiki适配性很好,能简单说一下怎么实现的吗?

2019-01-24



咖飞的白

化 8

看懂之后,果断敲出一个wiki页面来,才是最好得学习。

2019-01-24



小烦·张

ம் 5

只想说, 好多标签都没见过, 更别说用过, 惭愧惭愧

2019-01-24



mfist

凸 5

1通过wiki的例子发现真实的网页中还是有使用语义化的,之前以为现实应用中很少。

2看来有必要优化下我们网页的布局模板了。

3hr标签语义是话题的转向,学习了。之前以为就是一条横线,学习了。

2019-01-24



鱼膩

企4

进入wiki点开控制台,发现很多地方wiki本身并没有严格地按照win大说的来, 大部分也是did, s pan一把梭,有顺序的nav直接用的ul, 文中说的code, sample这些也都没有用。。是不是可以理解成其实很多时候我们为了实现样式的完全控制,降低了对HTML本身语义化的要求; 或者说在大部分条件下,快速还原设计稿比语义化本身更重要

2019-01-26



小妖精

企3

果真需要重学前端

2019-01-24



allen-黄河水

企 3

用css样式类名来进行语义化标签可行么?所以tagName还是className,这是个问题。

2019-01-24



木子烁東岸

企2

简历上写精通html的, 脸疼了

2019-01-25



叫我大胖就好了

凸 2

2019-01-24



水煮饺

凸 2

我只会div套div。还在努力学习。

2019-01-24



Damon

凸 2

好多都没见过。。。

2019-01-24



Sylvia

ഥ 1

不求多,只求精,除了上一篇的标签,此篇文章的abbr pre code figcaption site time dfn,我能说这些标签我几年前就用过,但后来见几乎没人用,也舍弃了,有些东西还是需要多坚持呀 Σ (′ Δ ')?!

2019-02-09



Dylan-Tseng

ഥ 1

大部分都见过,但都没有严格的去考究什么时候该用什么标签,不了解它的语义 2019-01-28



jim

ഥ 1

我最会用section...

2019-01-28



7777777

ഥ 1

为什么很多网站使用i...等标签作为小图标的背景??

2019-01-26



Lisongyi

ഥ 1

曾经做过whatwg文档的中文翻译工作,那里面的语义化在我看来是做到了极致



远恒之义

2019-01-25

凸 1

用 wiki 的网页例子来说明 HTML 语义标签的具体使用场景,浅显易懂,加强理解。

这种结合实际的网页来介绍不同的语义标签,非常适合前期入门学习。

2019-01-24