استخراج اطلاعات از صفحات وب با کمک HtmlAgilityPack

نویسنده: ایمان عبیدی

تاریخ: ۳۰/۱/۰۶/۱۳۹ ۸۵:۲۳

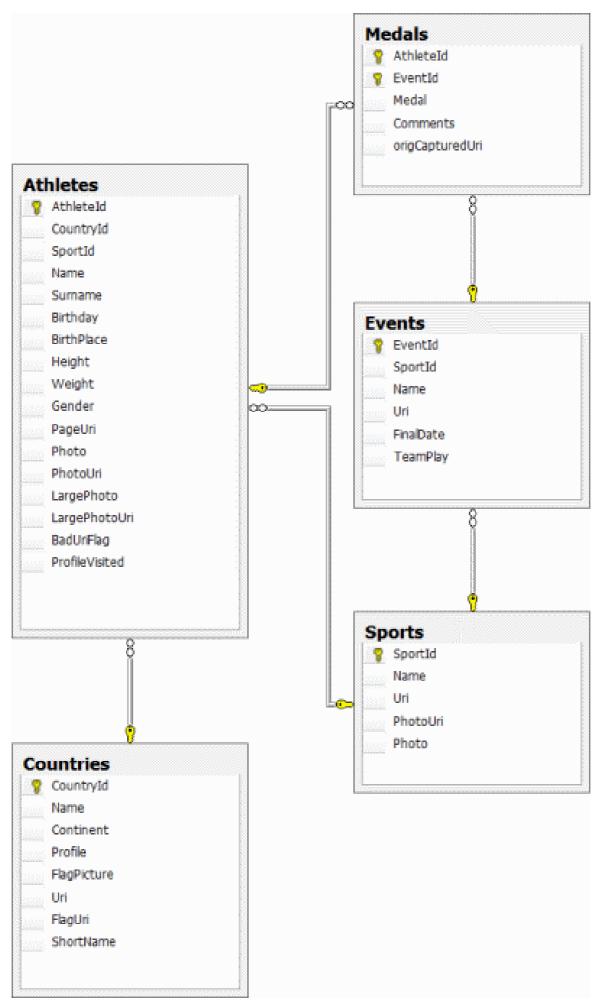
عنوان:

آدرس: www.dotnettips.info

گروهها: XPATH, xml, Html Agility Pack, FireBug

چند مدت پیش موقعی که تب المپیک بود و جدول http://www.london2012.com/medals/medal-count/ رو زیاد نگاه می کردم به نظرم رسید که کاشکی به اطلاعاتی مثل اینکه چند نفر از مدال آورها خانم و یا آقا هستند و یا اینکه در روزهای مختلف تعداد مدالها چطور توزیع می شند و بشه با یک jQuery UI Slider روزهای مختلف رو انتخاب کرد و جدول رو دید.

برای این کار اولین چیزی که لازم بود دریافت و ذخیره اطلاعات بود که من برای این کار از -Entity framework 4.1 Database first و کتابخانه HAP htmlagilitypack - HAP استفاده کردم . طراحی دیتابیس نهایی به این صورت شد



خوب در تلاش اول و مبتدیانه و بدون استفاده از این کتابخانه مفید چون اکثر صفحات وب XHTML نیستند و بالاخره چند تگ درست بسته نشده دارند و شما اگر بخواهید در آبجکت XmlDocument این htmlهای به ظاهر سالم رو لود کنید فورا با استثنای زیر مواجه میشوید

XmlException Was unhandeled
The 'img' start tag on line 1 position 1604 does not match the end tag of 'a'. Line 1, position 1766

راه حل ساده اینه که این کتابخونه رو با کمک NuGet نصب کنید

PM> Install-Package HtmlAgilityPack

و از اینجا به بعد با کدی مثل این میتونید از کلاس HtmlDocument و مشابه XmlDocument ولی بدون ارور استفاده کنید. مثلا با کد زیر میشه تاریخ تولد یک ورزشکار رو بدست آورد .توابع دیگه ای که خیلی جاها میتونه بدرد خورد GetAttributeValue و ChildNodes هست که یک نمونه نحوه استفادشو در ادامه میبینید

```
HtmlDocument xhtml = Crawler.GetXHtmlFromUri("http://www.london2012.com/athlete/hadadi-ehsan-
1077408/");
HtmlNode tempNode = xhtml.DocumentNode.SelectSingleNode("//table[@class='athleteBio']/tbody/tr[4]");

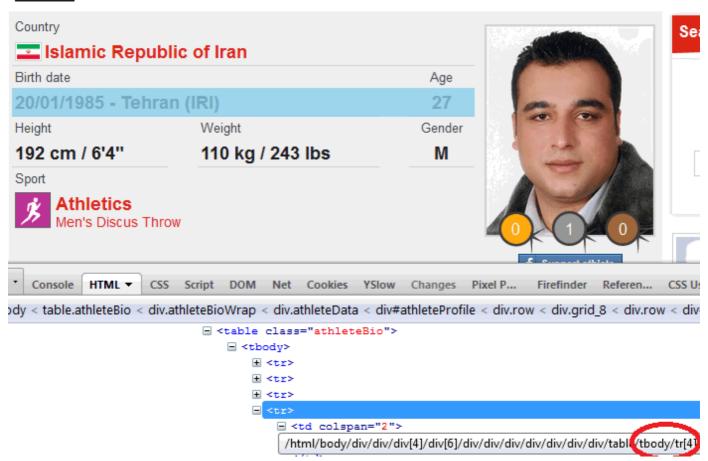
string temp = tempNode.FirstChild.FirstChild.InnerText.Replace(" ", "").Trim();
athlete.Birthday = DateTime.Parse(temp.Substring(0, 10), new CultureInfo("en-GB"));

tempNode = xhtml.DocumentNode.SelectSingleNode("//div[@class='athletePhotoMedals']/div/div/img");
athlete.LargePhotoUri = tempNode.GetAttributeValue("src", "");
```

البته تابع GetXHtmlFromUri رو جدا باید با کمک HttpWebRequest بنویسید و توی خوده HAP متاسفانه چنین تابعی توکار نشده نکته اصلی هم پیدا کردن محل دقیق اطلاعاته که با ابزاری مثل Firebug خیلی راحتتر میشه این کارو انجام داد. کافیه روی تاریخ تولد راست کلیک و inspect element by Firebug رو بزنید و حالا اگر تویه dom روی هر المنت html نگه دارید بهتون XPath کامل رو میده که میتونید تویه تابع DocumentNode.SelectSingleNode ازش استفاده کنید.

Ehsan Hadadi

Overview | Events and results |



برای درک بهتر XPath هم این 2 تا صفحه xpath_syntax و xpath_examples خیلی میتونه کمکتون بکنه.

نظرات خوانندگان

```
نویسنده: مهدی پایروند
تاریخ: ۴/۰۶/۱۳۹۱ ۱۲:۵۳
```

این روش برای استخراج مطالب همین سایت خیلی مفیده!

```
نویسنده: محسن کریمی
تاریخ: ۴۰/۶۰/۱۳۹۷ ۱۳:۵۷
```

با تشكر از مطلبتون

نکته: این Htmlagilitypack متاسفانه با صفحات فارسی و UTF8 مشکلات زیادی داره و واقعا نمیشه ازش استفاده کرد.

```
نویسنده: جمال
تاریخ: ۱۴:۱۹ ۱۳۹۱/۰۶/۰۴
```

Crawler.GetXHtmlFromUri

شى Crawler ازچه نوعيه؟موقع اجرا خطا ميگيره؟

```
نویسنده: وحید نصیری
تاریخ: ۴۰/۶۰/۱۳۹۱ ۱۴:۲۸
```

شروع کار به این صورت هم میتواند باشد:

OptionDefaultStreamEncoding رو به UTF8 تنظیم کنید.

```
نویسنده: محسن کریمی
تاریخ: ۱۵:۴۳ ۱۳۹۱/۰۶/۰۴
```

با تشکر

ولی طبق تجربه خود من کد بالا هم کمک نمیکنه و با تنظیم OptionDefaultStreamEncoding به UTF8 مشکل حل نمیشه ولی یه راه که قبلا من پیدا کرده بودم تو خوندن کد صفحات اینه که شما صفحات رو به صورت استریم دریافت کرده بعد توسط متد LoadHtml بخونید به این صورت مشکل برطرف میشه!(البته این تو سایت فارسی که من قصد خوندشو داشتم، بود با سایتهای دیگه تست نکردم)

```
var request = (HttpWebRequest)WebRequest.Create("آدرس سایت");
request.Method = "GET";
using (var response = (HttpWebResponse)request.GetResponse())
{
    using (var stream = response.GetResponseStream())
    {
        htmlDoc.Load(stream, Encoding.UTF8);
    }
}
```

```
نویسنده: وحید نصیری
تاریخ: ۴۶:۴۹ ۱۳۹۱/۰۶/۰۴
```

بستگی داره content نظر قبلی رو به چه فرمتی (چه Encoding ایی) از وب دریافت کردید. مابقی آن توسط این کتابخانه بدون مشکل یردازش میشود.

```
using System.Net;
//...
var content = new WebClient { Encoding = Encoding.UTF8 }.DownloadString(url);
```

```
نویسنده: ایمان عبیدی
تاریخ: ۴۰/۲۰۶۱ ۲۰:۴۶
```

Crawler همونطور که در متن هم نوشته شده دست سازه و مهم نیست و تابع GetXHtmlFromUri میتونه مثل نمونه زیر باشه و دقت کنید خالی نبودن UseAgent خیله مهمه وگرنه ارور The remote server returned an error: (409) Conflict رو میده. من با همین تابع یک سایت فارسی رو چک کردم و اروری نداد و متن فارسی قابل کوئری گرفتن بود. کامل تر و با ارور هندلینگ بهترش رو میتونید در برنامه مفید plrip آقای وحید نصیری ببینید

```
private static HtmlDocument GetXHtmlFromUri(string uri)
            HtmlDocument htmlDoc = new HtmlDocument()
                OptionCheckSyntax = true,
                OptionFixNestedTags = true,
                OptionAutoCloseOnEnd = true,
                OptionDefaultStreamEncoding = Encoding.UTF8
            };
            var request = (HttpWebRequest)WebRequest.Create(uri);
            request.Method = "GET";
            //important
             request.UserAgent = "Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64;
Trident/5.0)";
             request.Accept = "text/html";
            request.AutomaticDecompression = DecompressionMethods.GZip | DecompressionMethods.Deflate;
            using (var response = request.GetResponse() as HttpWebResponse)
                using (var stream = response.GetResponseStream())
                    htmlDoc.Load(stream, Encoding.UTF8);
            return htmlDoc;
```

اینم روش دوم که بازم UserAgent باید اضافه بشه

نویسنده: افشار محبی

تاریخ: ۵۰/۶/۱۳۹۱ ۹:۳۷

من هم از این ابزار در کارهایم استفاده کردهام. خیلی خوب جواب میدهد.

نویسنده: ایمان عبیدی تاریخ: ۲۳۹۱/۰۶/۰۶:۰

برای مشاهده نتایج بدست آمده رده بندی المپیک 2012 لندن به همراه اطلاعات جنسیت مدال گیرها و همچنین وضعیت جدول در روز هایه مختلف میتونید به لینک هایه زیر مراجعه کنید.

تویه این صفحات از پلاگین tableSorter و یکم جاوا اسکریپت هم در لینک اول برای کش کردن اطلاعات json استفاده کردم .

 $\underline{\text{http://olympics2012.iabidi.ir/Home/DailyMedals}}$

http://olympics2012.iabidi.ir/Home/RankingsFull

نویسنده: پریسا تاریخ: ۲۰:۵۸ ۱۳۹۲/۰۲/۰۲

چرا وقتی XPath از Firebug استفاده میکنم با استفاده از SelectSingleNode جواب دقیقی نمیده یا هیچی نمیاره

> نویسنده: وحید نصیری تاریخ: ۳۸/۲۰۲۰ ۱۴:۱۴

علت این است که html ایی که در فایرباگ بررسی میشه عموما به دلیل یک سری از نرمال سازیها توسط موتور فایرفاکس و همچنین خودش، با html اصلی یک سایت متفاوت است. به همین جهت XPath استخراجی از آن روی سایت اصلی کار نخواهد کرد. یک برنامه کمکی برای یافتن XPathها به همان نحوی که هستند.

> نویسنده: mahsan تاریخ: ۳۸:۳۸ ۱۳۹۲/۰۳/۲۳

این متد رو در چه صفحه ای باید بنویسم؟ آیا باید در همون صفحه ای که میخوام اطلاعات لود بشه بنویسم؟ uri مقدارش را باید داخل تابع بگیره؟ در page load فرمم چی بنویسم؟

نویسنده: وحید نصیری تاریخ: ۳/۲۳ /۱۳۹۲ ۱۱:۲۹

- تفاوتی نمی کنه کجا فراخوانی بشه ؛ در page load یا در یک روال رخداد گردان کلیک و یا در یک سرویس مستقل.
 - بهتره نتیجه رو بعد از فراخوانی برای مدتی کش کنی، تا هربار اطلاعات از وب درخواست نشود.

نویسنده: mahsan

تاریخ: ۲/۲۳ ۱۲:۱۰ ۱۲:۱۰

ببخشین که دوباره مزاحمتون شدم: وقتی من کد بالا رو در page load کپی میکنم زیر بعضی کلمات مانند , Crawler xhtml,HtmlNode,Parse, Birthday, LargePhotoUri, GetAttributeValueخط قرمز میکشه ! چرا؟ آیا HtmlDocument یک سند html که من باید ایجادش میکردم؟ فقط همین کدها رو بنویسم؟ جواب میگیرم با باید چیز دیگه ای هم اضافه بشه؟

نویسنده: وحید نصیری تاریخ: ۳۲/۴۱ ۱۳۹۲/۰۳۲

- شما باید از طریق نیوگت با دستور Install-Package HtmlAgilityPack این بسته رو نصب کنید. یا اینکه فایل DLL اون رو <u>از</u> سایتش دریافت و به ارجاعات پروژه اضافه کنید.
 - كدهاى كلاس Crawler چند كامنت بالاتر ارسال شدن. تابع GetXHtmlFromUri كه ملاحظه مىكنيد .
 - مواردی مانند Birthday, LargePhotoUri یک سری متغیر هستند که از طرف نویسنده مقاله تعریف شدن. مهم نیستند. حذفشون کنید.
 - یک مثال دیگر در مورد استفاده از کتابخانه HtmlAgilityPack با کد قابل دریافت .

نویسنده: mahsan

تاریخ: ۲۵/۳۹۲/۰۳/۲۵ ۹:۶

سلام ممنون که جوابم رو دادین

من کلاسی بنام Crawler ایجاد کردم ولی حالا زیر

OptionDefaultStreamEncoding

OptionAutoCloseOnEnd OptionFixNestedTags OptionCheckSyntax

,

LoadHtml

خط قرمز میکشه: (لطف میکنید بگید دلیلش چیه و باید چیکار کنم؟

و در فرمم هم زیر این کلمه GetXHtmlFromUri

مينويسه

generate metod stub for 'GetXHtmlFromUri in "crawel

وقتی من متد GetXHtmlFromUri را در کلاس crawel تعریف کردم جرا باید این پیغام رو بده ؟ آیا باید این گزینه رو انتخاب کنم؟

نویسنده: وحید نصیری

تاریخ: ۲۵/۳۹/۱۳۹۲ ۴۴:۹

پیشنهاد من این است که یک دوره سی شارپ مقدماتی رو بگذرونید. با مباحثی مانند نحوه تعریف فضای نام و روش فراخوانی یک متد استاتیک از کلاس متناظر با آن آشنا شوید.

یک دوره خوب مقدماتی سیشارپ

نویسنده: ایمان توکلی

تاریخ: ۲۲/۱۱/۲۶ ۱۷:۱۰

با سیاس

در صورتی که مقداری در querystring مربوط به صفحه اضافه شود و در هر درخواست این مقدار تغییر کند چطور میتوان صفحه را خواند مثال: http://website.com?log=1731004

سایت برای هر بازدید یک لاگ جدید مینویسد یعنی هر درخواست جدید در صورتی که لاگ معتبر نباشد به صفحه دیگر ارسال میکند.حال اطلاعات این صفحه را چطور میتوان خواند ؟

نویسنده: محسن خان

تاریخ: ۱۷:۲۱ ۱۳۹۲/۱۱/۲۶

این نوع مسایل رو نمیتونید با HtmlAgilityPack حل کنید. HtmlAgilityPack فقط کارش آنالیز اطلاعات ثابتی هست که بهش میدید و نهایتا خواندن نودهای HTML نهایی. موردی که عنوان کردید نیاز به آنالیز همزمان هدرهای دریافتی به همراه جاوا اسکریپت سایت مورد نظر داره.

عنوان: آموزش فايرباگ - #5 - HTML Panel

نویسنده: احمد احمدی

تاریخ: ۲۸:۱۵ ۱۳۹۱ ۱۸:۱۵

آدرس: www.dotnettips.info

برچسبها: JavaScript, FireBug, CSS, XPATH, HTML

در این سری از مقالات آموزش FireBug ، به صورت ترتیبی پیش رفتیم و ابتدا توضیحات تقریبا مفصلی در مورد پنل Console دادیم.

اکنون به یرکاربردترین بخش آن ، یعنی ینل HTML میرسیم.

محتویاتی که در این پنل نمایش داده میشود ، کدهای صفحهی جاری ، بصورت زنده است و با چیزی که از سمت سرور به مرورگر ارسال میشود متفاوت است ، تگهای HTML بصورت درختی مرتب شده اند و رنگی نمایش داده میشوند و امکان ویرایش محتوا ، ویرایش استایل ، مشاهدهی آرایش تگها بصورت بصری و ... وجود دارد.

نگاهی به کدهای ارسال شده به مرورگر در آدرس $\frac{Google.com}{}$ و نحوهی نمایش آن در فایرباگ را ملاحظه بفرمایید. (کدهای ارسال شده سمت چپ - نمایش در فایرباگ سمت راست)

```
Source of: http://www.google.com/ - Mozilla Firefox
                                                                                                                                                                                                                                                                                                                                                                                                                       🊧 🚏 < > | ≡ 🔻 Console HTML ▼ CSS Script DOM Net Cookies
<u>File Edit View Help</u>
                                                                                                                                                                                                                                                                                                                                                                                                                    Edit script < div#xjsi < body#gsr < html
                              <!doctype html><html itemscope="itemscope" itemtype="http://schema.org/We</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                              <!DOCTYPE html>
                            ml:function(){},pageState:"#",kHL:"en",time:function(){return(new Date).ge
                                                                                                                                                                                                                                                                                                                                                                                                                  ☐ <html itemtype="http://schema.org/WebPage" itemscope="itemscope
                            var k=/^http:/i;if(k.test(g)&&google.https()){google.ml(new Error("GLMM")
                                                                                                                                                                                                                                                                                                                                                                                                                                   # <head>
                            b: ! \ (!location.hash) \ \&\&! \ (!location.hash.match ("[\#\&] \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ ( \ (q|fp) = |tbs = simg|tbs = sbs)) \ | \ (q|fp) = |tbs = simg|tbs = sbs)
                                                                                                                                                                                                                                                                                                                                                                                                                                   (function() {var a=google.j; window.onpopstate=function() {a.psc=1}; for (var }
                                                                                                                                                                                                                                                                                                                                                                                                                                             document.f.g.focus();document.gbgf&&document.gbgf.g.focus();}}ca
                            {}; window.chrome.searchBox.onsubmit=function() {google.x({id:"psyapi"},function()
                                                                                                                                                                                                                                                                                                                                                                                                                                                   <a style="left:-1000em; position:absolute" href="/setprefs?pre
                            window.google.pt=window.gtbExternal&&window.gtbExternal.pageT();}catch(u)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            reader users, click here to turn off Google Instant. </a>
                              (function(){'use strict';var f=null,j=this;var l="undefined"!=typeof navig
                                                                                                                                                                                                                                                                                                                                                                                                                                                              <textarea id="csi" style="display:none"></textarea>
                           \label{lem:window.rwt=function(a,f,g,l,m,h,c,n,i)} $$ \left( \sup_{t \in \mathbb{R}^n} \frac{1}{t} \exp(a_t,b_t,c_t) \right) \left( \sup_{t \in \mathbb{R}^n} \frac{1}{t} \exp(a_t,b_t,c_t) \right) \left( \frac{1}{t} \exp(a_t,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    # <script>
                                                                                                                                                                                                                                                                                                                                                                                                                                                   □ <div id="mngb">
                            function ca(a,b,c) {var d="on"+b;if(a.addEventListener)a.addEventListener(k
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   document.body).appendChild(c)},F=function(a){for(var b=0,c;(c=y[b])&&!(c[0])
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  # <script>
                            I.l=[];H||y.push(["gl",{url:"//ssl.gstatic.com/gb/js/abc/glm_e7bb39a7e1a2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 = <div id="gbw">
                  15 function mlToken(a,b){trv{if(1>xa){xa++;var c,d=a,g=b||{},f=encodeURICom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 # <div id="gbzw">
                            \texttt{m.push} (\texttt{f} (\texttt{n})), \texttt{m.push} (\texttt{"="}), \texttt{m.push} (\texttt{f} (\texttt{g} [\texttt{n}])); \texttt{m.push} (\texttt{"\&emsg="}); \texttt{m.push} (\texttt{f} (\texttt{d.name})) 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 # <div id="gbg">
                  17 var Ka=function(){for(var a=[],b=0,c;c=Da[b];++b)(c=document.getElementByI
18 L(f,"gbto");else{if(Q){var k=document.getElementById(Q);if(k&&k.getAttribu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 # <div id="gbu">
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           </div>
                            a. \verb|currentStyle[c]|: a. \verb|style[c]|; \verb|return"| \verb|rtl"==b||, \verb|Wa=function(a,b,c)| | if (a) try (variable of the context o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <div id="gbx1" class="gbqfh"></div>
                            1<f.childNodes.length) {var Ca=f.childNodes[d+1]: !J(Ca.firstChild."gbmh") &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <div id="gbx3"></div>
                           \texttt{b[d]))} \texttt{return e; return j}, Xa = \texttt{function(a,b,c)} \\ \{ \texttt{Wa(a,b,c)} \}, Ya = \texttt{function(a,b)} \\ \{ \texttt{Wa(a,b,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ± <script>
                22 e},fb=h,Qa=function(a,b){R();if(a){gb(a,"Opening…");S(a,e);var c="u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           </div>
```

این پنل به سه بخش اصلی تقسیم میشود:

بخش اصلی یا NodeView که محتوای صفحه را بصورت درختی و مرتب و رنگی نمایش میدهد.

Panel Toolbar که در بالای پنل قرار دارد.

Side Panels که شامل پنلهای Style , Computed , Layout , DOM میشود.

که به ترتیب برای نمایش و ویرایش استایلها ، مشاهده استایلهای محاسبه شده ، مشاهده Layout یا آرایش و نمایش اطلاعات DOM تگ انتخاب شده در NodeView است.

در این مقاله با دو بخش NodeView و Panel Toolbar ، و در مقالهی بعد با پنلهای سمت راست یعنی Side Panels آشنا میشویم.

ويرايش HTML

هر تگ HTML از یک سری Attribute تشکیل میشود که در فایرباگ نام ویژگی بصورت آبی تیره و مقدار آن با رنگ قرمز مشخص شده است. برای ویرایش هریک از آنها کافیست برویش کلیک کنید تا آن مقدار در یک باکس ویرایش به نمایش دربیاید. با ویرایش مقدار ، این تغییر در لحظه بروی صفحه اعمال میشود.



برای اضافه کردن یک Attribute جدید به تگ هم بروی تگ مورد نظر راست کلیک میکنید و سپس گزینهی New Attribute را انتخاب میکنید. ابتدا نام ویژگی ، یک Enter ، سپس مقدار را وارد میکنید. با زدن کلیدهای Enter متوالی ، میتوانید به وارد کردن ویژگیها ادامه دهید.

برای ویرایش کردن یک تگ و تگهای فرزندش ، بروی تگ موردنظر کلیک کنید تا به حالت انتخاب دربیاید ، سپس بروی دکمهی Edit در بالای ینل کلیک کنید.

Node View

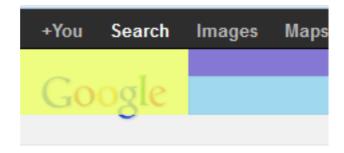
NodeView نام بخش اصلی پنل HTML است که محتویات صفحه را بصورت مرتب شده و درختی نمایش میدهد.

تگ (Node) هایی که دارای فرزند میباشند ، یک علامت + یا - در کنارشان وجود دارد که با کلیک بروی آن میتوانید آن تگ را باز/بسته کنید. همچنین این کار با کلیدهای + و - یا Right Arrow و Left Arrow از روی کیبورد هم قابل انجام است.

برای باز کردن یک لینک در این قسمت هم از ترکیب Ctrl و کلیک موس کمک بگیرید.

در نهایت زمانی که موس را بروی یک تگ قرار میدهید ، محیطی که توسط آن تگ در صفحه اشغال شده است بصورت رنگی مشخص میشود.

که رنگ زرد به معنی محیط margin ، رنگ آبی تیره به معنی محیط padding و رنگ آبی روشن هم به معنی محیط محتوا است.



Panel Toolbar

این قسمت در بالای پنل HTML قرار دارد که گزینههای زیر را دارا میباشد:



Break On Mutate

این دکمه امکان توقف کد JavaScript ای که سعی در ویرایش محتوای صفحه را دارد ، میدهد.

زمانی که FireBug تشخیص دهد که کدی سعی در ویرایش محتوا دارد ، شما را به خط مورد نظر از کد ، در پنل Script منتقل میکند.

Edit

این دکمه برای ویرایش مستقیم محتوای یک تگ بکار میرود

نکتهی جالب در ویرایش محتوا در فایرباگ این است که تغییرات در لحظه اعمال میشوند و نیاز به عمل بروزرسانی جداگانه نیست. برای مثال در همین قسمت Edit ، با هر ویرایش محتوا ، تغییرات در لحظه اعمال میشوند.

Element Path

زمانی که یک تگ را در FireBug انتخاب میکنید ، لیستی از تگها که از تگ جاری شروع و به تگ ریشه ختم میشود ، نمایش داده میشود که با کلیک بروی هرکدام ، همان به عنوان تگ فعلی یا انتخاب شده تعیین میشود.

نتیجهی عملیاتی که بروی این تگهای انجام میدهید (حرکت موس و راست کلیک کردن) معادل همان عملیات بروی تگهای نمایش داده شده در قسمت اصلی (NodeView) است.

Options Menu

هر تب یا پنل در فایرباگ دارای یک سری تنظیمات است که Options Menu نام دارد. تب HTML هم دارای یک سری تنظیمات است که دانشتن آنها بسیار به شما کمک خواهد کرد.

این منو با کلیک کردن بروی فلش تب (



) یا راست کلیک کردن بروی تب ظاهر میشود.

Show Full Text

در صورت فعال بودن ، متون بصورت کامل نمایش داده میشوند ، در غیراینصورت بصورت خلاصه شده نمایش داده میشوند.

Show White Space

در صورت فعال بودن ، فضاهای خالی ، کرکترهای خط جدید و ... را نمایش میدهد.

 ${\bf Show}\;{\bf Comments}$

در صورت فعال بودن ، كامنتها را هم نمايش مىدهد

```
div id="jump-to-nav">
  <!-- start content -->

div class="thumb tright">
  The HTML panel displays the generated HTML/XML of the currently opened page.
```

سه گزینه ی Show Entities As Symbols ، Show Entities As Names و Show Entities As Names ، نوع نمایش کرکترهای ویژه را تعیین میکنند.

Highlight Changes

در صورت فعال بودن ، تگ تغییر یافته توسط JavaScript (یا تگ والدی که قابل مشاهده باشد) Highlight میشود

Expand Changes

در صورت فعال بودن ، زمان تغییر دادن یک تگ توسط JavaScript ، والدهای آن تگ باز (Expand) میشوند

Scroll Changes Into View

در صورت فعال بودن این گزینه ، هنگام تغییر یک تگ در صفحه توسط JavaScript ، قسمت NodeView به قسمت تغییر بافته حرکت میکند.

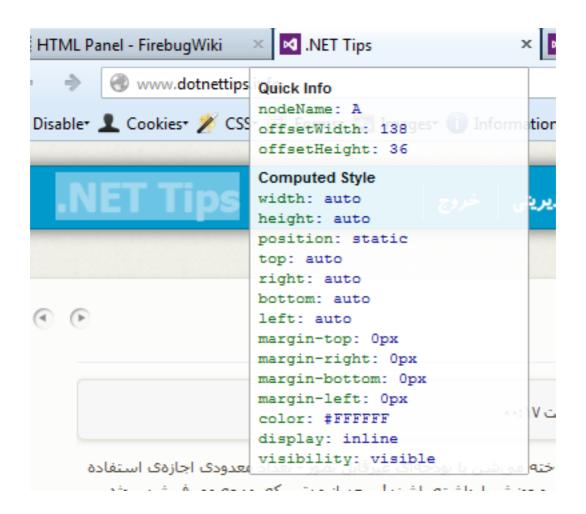
(فعال بودن این گزینه هنگام بررسی سایت هایی که اسلایدر یا سیستم هایی مشابه دارند ، باعث میشه که نتوانید بروی قسمتهای سایت تمرکز کنید و مدام اسکرول به قسمت تغییرات منتقل میشود.)

Shade Box Model

در صورت فعال بودن ، فضای margin , padding و content را به شکلی که در بالا گفته شد نمایش میدهد ، در غیر اینصورت فقط یک خط دور تگ نشان میدهد.

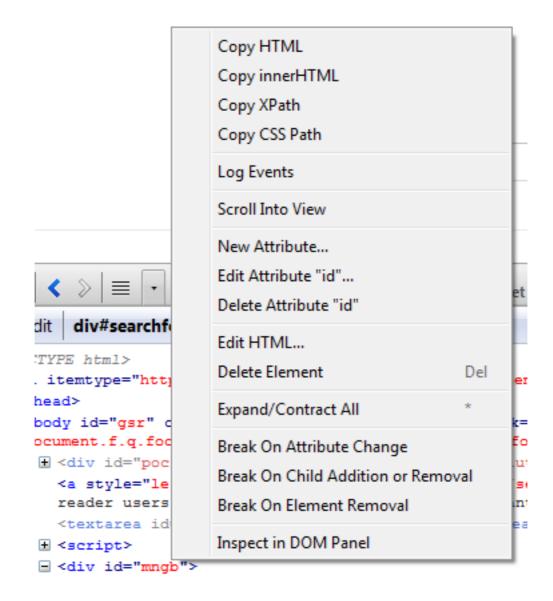
Show Quick Info Box

در صورت فعال بودن ، یک پاپآپ به همراه اطلاعات مختصری از تگ در صفحه نمایش میدهد.



Context Menu

اگر بروی یک تگ راست کلیک کنید ، یک منو نمایش داده میشود ، در این قسمت با گزینههای این منو آشنا میشویم.



Copy HTML

خود تگ و محتوایش را بصورت HTML در حافظه کپی میکند.

Copy innerHTML

محتوای تگ را در حافظه کیی میکند.

Copy XPath

آدرس XPath تگ را در حافظه کپی میکند.

Copy CSS Path

CSS Selector تگ را در حافظه کپی میکند.

Log Events

رویدادهای تگ را در پنل Console ثبت می کند. (کلیک موس ، فشردن کلید ، ...) برای لغو لاگ کردن ، مجددا بروی تگ راست کلیک کرده و این گزینه را از حالت انتخاب خارج کنید.

Scroll Into View

صفحه را به جایی که تگ قابل نمایش است منتقل میکند.

...New Attribute

یک attribute جدید به تگ اضافه میکند. برای لغو عملیات ، دکمهی Esc را بزنید.

..."<Edit Attribute "<attribute name

اگر بروی یک attribute راست کلیک کرده باشید ، این گزینه و گزینهی بعدی را مشاهده خواهید کرد. معادل کلیک بروی نام attribute است ، نام را به حالت ویرایش درمی آورد.

"<Delete Attribute "<attribute name</pre>

attribute را حذف میکند.

...Edit HTML

تگ را به حالت ویرایش میبرد. معادل انتخاب تگ ، زدن کلید Edit است.

Delete Element

تگ را حذف میکند.

راه دیگر حذف یک تگ ، انتخاب تگ و فشردن کلید Del از کیبورد است.

Expand/Contract All

تگ و Childهایش را باز/بسته میکند. (بجز تگ های <script> , <style> , <link>) میتوان با ترکیب کلید **Shift + *** هم این کار را انجام داد که در این حالت تگهای فوق هم شامل باز/بسته شدن میشوند.

Break On Attribute Change

مانع اجرای کد JavaScript ای که attribute تگ را تغییر میدهد میشود و فایرباگ به خطی که باعث این عمل شده است در پنل Script منتقل میشود.

به عبارتی دیگر یک Break Point در خط JavaScript ای که باعث ویرایش attribute میشود قرار میدهد.

Break On Child Addition or Removal

مشابه توضيح قبل ، Break Point را در خطى كه باعث اضافه/حذف شدن تكِ Child شده است قرار مىدهد.

Break On Element Removal

مشابه توضیح قبل ، Break Point را در خطی که باعث حذف شدن تگ شده است قرار میدهد.

Inspect in DOM Tab

تگ فعلی را در ینل DOM ، برای بررسی باز میکند.

در قسمت بعدی با ینلهای سمت راست (Side Panels) آشنا میشویم.

نظرات خوانندگان

نویسنده: احمد احمدی

تاریخ: ۵۲/۱۲/۲۵ ۰۵:۰

در ورژن 1.11 از این افزونه ، امکانی برای paste کردن Html به صفحه ، به Context منوی پنل HTML اضافه شده که کار ویرایش محتویات Html را سادهتر کرده.

در حالی که قبلا برای این کار میبایست تگ را به حالت ویرایش در آورد و محتویات را اضافه کرد.

