

Re: How to judge whether content type is truly "text/html"?

## Re: How to judge whether content type is truly "text/html"?

---

*Source:*

<http://www.tech-archive.net/Archive/DotNet/microsoft.public.dotnet.languages.csharp/2006-09/msg03932.html>

---

- *From:* "Morgan Cheng" <[morgan.chengmo@xxxxxxxx](mailto:morgan.chengmo@xxxxxxxx)>
  - *Date:* 24 Sep 2006 22:30:50 -0700
- 

Kevin Spencer wrote:

Hi Morgan,

Your experience underscores my point. While it is possible to manually (or, perhaps unintentionally) change the ContentType header, any web site that did would find out about it very quickly, because there are many different browsers in use out there, and they would hear about the problem and fix it.

It isn't productive to imagine the most remote of possibilities and handle them gracefully. If one did, one would never finish much of anything. Sometimes the most graceful thing to do is to handle the error as an error and move on. My guess is that you would never run into the issue at all.

I happen to find one function FindMimeFromData in UrlMon.dll. It works.

[http://msdn.microsoft.com/workshop/networking/moniker/overview/appendix\\_a.asp](http://msdn.microsoft.com/workshop/networking/moniker/overview/appendix_a.asp)

--  
HTH,

Kevin Spencer  
Microsoft MVP  
Software Composer

A watched clock never boils.

"Morgan Cheng" <[morgan.chengmo@xxxxxxxx](mailto:morgan.chengmo@xxxxxxxx)> wrote in message  
<news:1158854086.157720.267730@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>

Kevin Spencer ??:

Re: How to judge whether content type is truly "text/html"?

Re: How to judge whether content type is truly "text/html"?

The property is not decided by the HTTP Response Header. It is decided by the web server and/or the developer who created the web site. The problem here is, the reason for the ContentType header is to tell the client what is stored in the stream of bits it is sending. Since a stream of bits is just 1's and 0's there's no way to tell without it.

Yes, web server can config the response MIME type, which turns to be in HTTP response header. That is my understanding.

However, I have never heard of what you describe happening. If it did, browsers would not be able to view the content, and whomever created the web site would know about it very shortly (from the response of the users).

I tried to manually set one html to be "image/jpeg" type in IIS6. Then access the page from another machine and ambush the http package with Fiddle. It shows that the response header has "ContentType: image/jpeg". Interestingly, IE still show the html page, while Firefox cannot show it up. It looks that IE does further job.

--  
HTH,

Kevin Spencer  
Microsoft MVP  
Software Composer

A watched clock never boils.

"Morgan Cheng" <morgan.chengmo@xxxxxxxx> wrote in message  
[news:1158847522.975327.265330@xx](mailto:news:1158847522.975327.265330@xx)

I know that `HttpRequest.GetResponse()` generates a `HttpWebResponse`. The response has one `ContentType` property. But the property is just decided by http response header. It is possible that the content is actually HTML, while the `ContentType` is "image/jpeg".

Is there any effective way to judge whether the response type is truly "text"?

Re: How to judge whether content type is truly "text/html"?

I have a idea to read the first several bytes of the response stream;  
and check whether they are real displayable characters. But, they can be any kind of Encoding. Should I try all kinds of Encoding?