

Re: Strip HTML

Source:

<http://www.tech-archive.net/Archive/Access/microsoft.public.access.formscoding/2006-05/msg00033.html>

- *From:* State Troopers <StateTroopers@xxxxxxxxxxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Mon, 1 May 2006 08:50:02 -0700
-

replace < > with
& lt; and & gt;

was just for testing purposes.

"State Troopers" wrote:

Hey.
I want to remove everything BUT

and I was replacing < with < just to test out ".< and .>"

Thanks.

"Douglas J Steele" wrote:

If all you want to do is remove the
, you can use

```
strOutput = Replace(strOutput, "<BR>", " ")
```

BTW, why are you using Replace to change < to < and > to >? That does absolutely nothing, other than taking machine cycles!

--
Doug Steele, Microsoft Access MVP
<http://I.Am/DougSteele>
(no e-mails, please!)

"State Troopers" <StateTroopers@xxxxxxxxxxxxxxxxxxxxxxxxxxxxx> wrote in message
news:C77C1BDF-CCF1-4CFE-9D07-A0ABF9A106E4@xxxxxxxxxxxxxxxxxxxxx

Hi there.

I am currently working on a database/program that will take

Re: Strip HTML

a certain html
page and store the strings on the page into a table, which will
then be

used

in reports/queries.

I have started with stripping the HTML tags off the page
first. The

function

works very well.

But, my problem is that I am not sure how I would eliminate
the function
from removing "
" .

Here is the function:

```
'Ensure that strHTML contains something
If Len(strHTML) = 0 Then
stripHTML = strHTML
Exit Function
End If
```

```
Dim aysplit, i, j, strOutput
```

```
aysplit = Split(strHTML, "<")
```

```
'Assuming strHTML is nonempty, we want to start iterating
'from the 2nd array position
If Len(aysplit(0)) > 0 Then j = 1 Else j = 0
```

```
'Loop through each instance of the array
For i = j To UBound(aysplit)
'Do we find a matching > sign?
If InStr(aysplit(i), ">") Then
'If so, snip out all the text between the start of the string
'and the > sign
'IF statement to NOT remove <br> tags.
```

```
aysplit(i) = Mid(aysplit(i), InStr(aysplit(i), ">") + 1)
Else
'Ah, the < was was nonmatching
aysplit(i) = "<" & aysplit(i)
End If
Next
```

```
'Rejoin the array into a single string
strOutput = Join(aysplit, "")
```

Re: Strip HTML

```
'Snip out the first <  
strOutput = Mid(strOutput, 2 - j)
```

```
'Convert < and > to < and >  
strOutput = Replace(strOutput, ">", ">")  
strOutput = Replace(strOutput, "<", "<")  
strOutput = Replace(strOutput, "-", "<")  
stripHTML = strOutput
```

Thanks.
-State