

Forum:[Hack Forums](#)

Thread: [How To Save Value After Specific Text](#)

Meterpreter - February 4, 2019, 7:06 pm

So I have managed to download a HTML page and save it, then save all lines containing my specific text. But say 50 lines containing the entire line with my keyword/key identifier in it, how can I take the value directly after that identifier? For example:

```
<span class="class" id="pid_*some number I want*
```

and in front of that is:

```
"><a href=
```

So it looks like:

```
<span class="class" id="pid_*some number I want*"><a href=
```

How would I go about get the *some number I want*?? I am new to C#, meaning I started today. I have spent a few hours googling, but to no avail.

Any suggestions or resources are very much appreciated :)

0xDEAD10CC - February 4, 2019, 7:45 pm

Many ways, quick and dirty is regular string parsing or regex but keep in mind if the formatting changes that your regex should try to be as abstract as possible while still being specific enough that it grabs the right value off the page. There's other libraries out there for DOM parsing too, some people convert HTML to XHTML and use an XML parser... Just keep in mind that if the HTML is ill-formed, then these methods might not work that well either. Lots of people still use HTML agility pack as well, but AngleSharp as well (which I've never used before).

Meterpreter - February 4, 2019, 8:18 pm

(Feb 4, 2019 08:45 PM)0xDEAD10CC Wrote: [\[CITE\]](#) Many ways, quick and dirty is regular string parsing or regex but keep in mind if the formatting changes that your regex should try to be as abstract as possible while still being specific enough that it grabs the right value off the page. There's other libraries out there for DOM parsing too, some people convert HTML to XHTML and use an XML parser... Just keep in mind that if the HTML is ill-formed, then these methods might not work that well either. Lots of people still use HTML agility pack as well, but AngleSharp as well (which I've never used before).

I had a look into regex but cant get it to work.

I have parsed the entire html to a local file. Then streamread the file and found all lines containing what I want and save to its own file.

So I have the lines but I literally need to say: Everything before pid_, ignore it, and everything after "><a href, ignore that too. And whatever is in between those two, save as a variable/array.

0xDEAD10CC - February 4, 2019, 8:33 pm

Well I gave you the options, how do you expect people to help you when you don't show any of your

code?

cristol - February 4, 2019, 9:15 pm

Hope this one useful for you

```
Code private void Form1_Load(object sender, EventArgs e)
{
    string pid = @"<span class=""class"" id=""pid_123456789""><a href="";
    Match m = Regex.Match(pid, @"<span class=""class"" id=""(.*)""><a href=",
RegexOptions.IgnoreCase);
    Console.WriteLine(m.Groups[1].Value);
}
Output
Code pid_123456789
If have multilines you should parse it to Reader and Loop until the last line.
```

Meterpreter - February 4, 2019, 9:26 pm

(Feb 4, 2019 09:33 PM)0xDEAD10CC Wrote: [\[CITE\]](#) Well I gave you the options, how do you expect people to help you when you don't show any of your code?

Here is my code:

```
Code private void button1_Click(object sender, EventArgs e)
{
    string postCount = textBox1.Text;
    string postSpeed = textBox2.Text;
    string messageSend = textBox3.Text;
    bool input = false;

    if (string.IsNullOrEmpty(postCount) == true || string.IsNullOrEmpty(postSpeed) == true ||
string.IsNullOrEmpty(messageSend) == true)
    {
        label4.Text = "Missing input!";
    }
    else
    {
        label4.Text = "Starting up";
        input = true;
    }

    if (input)
    {
        if (File.Exists(@"C:\Users\Az\WriteHTML.txt"))
        {
            // If file found, delete it
            File.Delete(@"C:\Users\Az\WriteHTML.txt");
            Console.WriteLine("File deleted.");
        }
        else
        {
            Console.WriteLine("No such file.");
        }

        if (File.Exists(@"C:\Users\Az\WriteLines.txt"))
```

```

{
    // If file found, delete it
    File.Delete(@"C:\Users\Az\WriteLines.txt");
    Console.WriteLine("File deleted.");
}
else
{
    Console.WriteLine("No such file.");
}

```

```

HttpWebRequest request =
(HttpWebRequest)WebRequest.Create("http://localhost/thread.php?pid=133");
request.UserAgent = "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/72.0.3626.81 Safari/537.36";
request.Accept =
"text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8";
request.Referer = "https://google.com/";
request.Headers.Add("Cookie", "cookie-info-here");

```

```

HttpWebResponse response = (HttpWebResponse)request.GetResponse();
Stream data = response.GetResponseStream();
string html = String.Empty;
using (StreamReader sr = new StreamReader(data))
{
    html = sr.ReadToEnd();
}

```

```

System.IO.File.WriteAllText(@"C:\Users\Az\WriteHTML.txt", html);

```

```

string line;
string algo = "thread\ id=\"pid_";

```

```

StreamReader file =
    new StreamReader(@"C:\Users\Az\WriteHTML.txt");
while ((line = file.ReadLine()) != null)
{
    if (line.Contains(algo))
    {
        File.AppendAllText(@"C:\Users\Az\WriteLines.txt", line + Environment.NewLine);
        counter++;
    }
}

```

```

file.Close();
System.Console.ReadLine();
}

```

} (Feb 4, 2019 10:15 PM)cristlol Wrote: [CITE](#) Hope this one useful for you

```

Code private void Form1_Load(object sender, EventArgs e)
{
    string pid = @"<span class=""class"" id=""pid_123456789""><a href="";
    Match m = Regex.Match(pid, @"<span class=""class"" id=""(.*)""><a href="";

```

```
RegexOptions.IgnoreCase);  
    Console.WriteLine(m.Groups[1].Value);  
}
```

Output

Code pid_123456789

If have multilines you should parse it to Reader and Loop until the last line.

After some tinkering it is working, this helped me a lot man, thanks for helping learn! Appreciate it.