

# Re: Unicode to UTF8

---

*Source:*

<http://www.tech-archive.net/Archive/Scripting/microsoft.public.scripting.vbscript/2007-03/msg00413.html>

---

- *From:* "Anthony Jones" <[Ant@xxxxxxxxxxxxxxxxxxxx](mailto:Ant@xxxxxxxxxxxxxxxxxxxx)>
  - *Date:* Wed, 14 Mar 2007 17:25:54 -0000
- 

"Paul Randall" <[paulr901@xxxxxxxxxxxx](mailto:paulr901@xxxxxxxxxxxx)> wrote in message  
[news:OfC6txkZHHA.1260@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](mailto:news:OfC6txkZHHA.1260@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

"Anthony Jones" <[Ant@xxxxxxxxxxxxxxxxxxxx](mailto:Ant@xxxxxxxxxxxxxxxxxxxx)> wrote in message  
[news:%23LPSqKiZHHA.2320@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](mailto:news:%23LPSqKiZHHA.2320@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

"Saku Ruokolahti" <[no-spam@xxxxxxxxxxxx](mailto:no-spam@xxxxxxxxxxxx)> wrote in message  
[news:u7EHjchZHHA.4832@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](mailto:news:u7EHjchZHHA.4832@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

Sorry for my English.  
I search for suitable solution.  
My HTA opens XML base (As MSXML2. DOMDocument),  
creates HTML (UTF8)

document and saves HTML file.

XML contains unicode strings. I need to convert strings from  
Unicode to

UTF8. How to do it?

Thanks

Your question doesn't make sense. Unicode and UTF-8 represent the same  
set  
of characters. Why would you want to convert Unicode to UTF-8?

Perhaps because in this application so few characters need more than one  
byte of UTF-8 encoding that it saves significant storage and web traffic  
versus the two bytes required for each 16-bit Unicode character?

## Re: Unicode to UTF8

Yes but the OP already stated that the HTML is being saved as UTF-8 so file size is not an issue. There is no practical way to avoid using Unicode in the script itself. There doesn't seem to be an issue here to address. Place unicode string in an XML DOM and save the DOM the default is to save UTF-8

Anthony – might this be a good application of OlePrn.OleCvt? I can't find any good documentation for this object at [msdn.microsoft.com](http://msdn.microsoft.com), but it was updated as recently as Win Server 2003 SP1, and it has a Property ToUtf8(ByVal bstrUnicode As String) As String [Get/o], according to my object browser.

It's a bit weird that something that performs such a conversion returns a string. I suspect it returns a string since it is a handy lightweight way to store a series of bytes. I'm not sure what you could do with it in VBScript once you had it.

My own implementation returns a byte array. It's only a small amount of VB6 code around the WideCharToMultiByte API function.

In pure VBScript I would use the ADODB.Stream object.

–Paul Randall

–Paul Randall