

Re: Convert DOS Cyrillic text to Unicode

Source:

<http://www.tech-archive.net/Archive/DotNet/microsoft.public.dotnet.languages.vb/2004-07/6251.html>

From: Nikolay Petrov (*johntup2_at_mail.bg*)

Date: 07/27/04

Date: Tue, 27 Jul 2004 12:23:21 +0300

That was very helpfull.

But I have some problems. Let me first tell you exactly what I want to achieve.

I've made a simple ASP .NET page with two text boxes and a button. What I need is, that a user paste DOS cyrillic text (taken from Notepad) in left text box, and when he clicks the button, the Converted to Unicode text to appear at the right box.

So I get the DOS text as String, not as bytes. How should I proceed in this case?

"Jay B. Harlow [MVP – Outlook]" <Jay_Harlow_MVP@msn.com> wrote in message news:utwflAycEHA.2268@TK2MSFTNGP12.phx.gbl...

> *Nikolay,*

> *In addition to the other comments*

>

> *What is the Code Page for DOS Cyrillic? Quickly checking MSDN I think its*

> *866, but you need to double check!*

>

> *You would use Encoding.GetEncoding to get the DOS Cyrillic Encoding object.*

>

> *Imports System.Text*

>

> *Dim cyrillic As Encoding = Encoding.GetEncoding(866)*

>

> *Given an array of Bytes with DOS Cyrillic in it, you would use*

> *Encoding.GetString to convert to a Unicode String.*

>

> *Dim bytes() As Byte*

> *Dim s As String = cyrillic.GetString(bytes)*

>

> *Given a Unicode String, you would us Encoding.GetBytes to get an array of*

> *Bytes with DOS cyrillic.*

>

> *bytes = cyrillic.GetBytes(s)*

>

microsoft.public.dotnet.languages.vb: Re: Convert DOS Cyrillic text to Unicode

- > *If your DOS cyrillic is in a Text File you pass the Encoding object to your*
- > *System.IO reader & writer classes*
- >
- > *Dim input As New StreamReader("myCyrillic.txt", cyrillic)*
- >
- > *Dim output As New StreamWriter("myCyrillic.txt", False, cyrillic)*
- >
- > *For information on Unicode, Encoding, and code pages (such as DOS Cyrillic)*
- > *see:*
- >
- > *<http://www.yoda.arachsys.com/csharp/unicode.html>*
- >
- > *One last thing: Once you have a String it is Unicode! Only Byte arrays &*
- > *Streams contain DOS Cyrillic and other character encodings.*
- >
- > *Hope this helps*
- > *Jay*
- >
- > *"Nikolay Petrov" <johntup2@mail.bg> wrote in message*
- > *news:u8aV5UvcEHA.3632@TK2MSFTNGP09.phx.gbl...*
- > > *How can I convert DOS cyrillic text to Unicode*
- > >
- > >
- >
- >