

VisualBasic .net streamreader encoding

Reading text files that don't have default encoding (your system default) can be a pain. Here is how you do it - with full list of encodings

```
Dim SR As System.IO.StreamReader = New StreamReader("in.txt",  
Encoding.GetEncoding("ibm852"))  
Dim strtmp As String  
While Not SR.EndOfStream  
    strtmp = SR.ReadLine()  
    'Do something with this line  
End While  
SR.Close()  
SR.Dispose()  
SR = Nothing
```

The this example retrieves the different names for each encoding and compares them with the equivalent Encoding names:

```
Imports System
Imports System.Text

Public Class SamplesEncoding
    Public Shared Sub Main()
        Console.WriteLine("Info.CodePage      ")
        Console.WriteLine("Info.Name          ")
        Console.WriteLine("Info.DisplayName")
        Console.WriteLine()
        Dim ei As EncodingInfo
        For Each ei In Encoding.GetEncodings()
            Dim e As Encoding = ei.GetEncoding()

            Console.WriteLine("{0,-15}", ei.CodePage)
            If ei.CodePage = e.CodePage Then
                Console.WriteLine("      ")
            Else
                Console.WriteLine("*** ")
            End If

            Console.WriteLine("{0,-25}", ei.Name)
            If ei.CodePage = e.CodePage Then
                Console.WriteLine("      ")
            Else
                Console.WriteLine("*** ")
            End If

            Console.WriteLine("{0,-25}", ei.DisplayName)
            If ei.CodePage = e.CodePage Then
                Console.WriteLine("      ")
            Else
                Console.WriteLine("*** ")
            End If
        End For
    End Sub
End Class
```

```

        Else
            Console.Write("*** ")
        End If
        Console.WriteLine()
    Next ei

End Sub

End Class

```

'This code produces the following output.

```

'Info.CodePage      Info.Name          Info.DisplayName
'37                 IBM037             IBM EBCDIC (US-Canada)
'437                IBM437             OEM United States
'500                IBM500             IBM EBCDIC (International)
'708                ASM0-708           Arabic (ASM0 708)
'720                DOS-720           Arabic (DOS)
'737                ibm737            Greek (DOS)
'775                ibm775            Baltic (DOS)
'850                ibm850            Western European (DOS)
'852                ibm852            Central European (DOS)
'855                IBM855            OEM Cyrillic
'857                ibm857            Turkish (DOS)
'858                IBM00858          OEM Multilingual Latin I
'860                IBM860            Portuguese (DOS)
'861                ibm861            Icelandic (DOS)
'862                DOS-862           Hebrew (DOS)
'863                IBM863            French Canadian (DOS)
'864                IBM864            Arabic (864)
'865                IBM865            Nordic (DOS)
'866                cp866             Cyrillic (DOS)
'869                ibm869            Greek, Modern (DOS)
'870                IBM870            IBM EBCDIC (Multilingual
Latin-2)
'874                windows-874        Thai (Windows)
'875                cp875             IBM EBCDIC (Greek Modern)
'932                shift_jis          Japanese (Shift-JIS)
'936                gb2312             Chinese Simplified (GB2312)
'949                ks_c_5601-1987     Korean
'950                big5              Chinese Traditional (Big5)
'1026              IBM1026            IBM EBCDIC (Turkish
Latin-5)
'1047              IBM01047           IBM Latin-1
'1140              IBM01140           IBM EBCDIC (US-Canada-Euro)
'1141              IBM01141           IBM EBCDIC (Germany-Euro)
'1142              IBM01142           IBM EBCDIC (Denmark-Norway-
Euro)
'1143              IBM01143           IBM EBCDIC (Finland-Sweden-
Euro)

```

'1144	IBM01144	IBM EBCDIC (Italy-Euro)
'1145	IBM01145	IBM EBCDIC (Spain-Euro)
'1146	IBM01146	IBM EBCDIC (UK-Euro)
'1147	IBM01147	IBM EBCDIC (France-Euro)
'1148	IBM01148	IBM EBCDIC (International-Euro)
'1149	IBM01149	IBM EBCDIC (Icelandic-Euro)
'1200	utf-16	Unicode
'1201	unicodeFFFE	Unicode (Big-Endian)
'1250	windows-1250	Central European (Windows)
'1251	windows-1251	Cyrillic (Windows)
'1252	Windows-1252	Western European (Windows)
'1253	windows-1253	Greek (Windows)
'1254	windows-1254	Turkish (Windows)
'1255	windows-1255	Hebrew (Windows)
'1256	windows-1256	Arabic (Windows)
'1257	windows-1257	Baltic (Windows)
'1258	windows-1258	Vietnamese (Windows)
'1361	Johab	Korean (Johab)
'10000	macintosh	Western European (Mac)
'10001	x-mac-japanese	Japanese (Mac)
'10002	x-mac-chinesetrad	Chinese Traditional (Mac)
'10003	x-mac-korean	Korean (Mac)
'10004	x-mac-arabic	Arabic (Mac)
'10005	x-mac-hebrew	Hebrew (Mac)
'10006	x-mac-greek	Greek (Mac)
'10007	x-mac-cyrillic	Cyrillic (Mac)
'10008	x-mac-chinesesimp	Chinese Simplified (Mac)
'10010	x-mac-romanian	Romanian (Mac)
'10017	x-mac-ukrainian	Ukrainian (Mac)
'10021	x-mac-thai	Thai (Mac)
'10029	x-mac-ce	Central European (Mac)
'10079	x-mac-icelandic	Icelandic (Mac)
'10081	x-mac-turkish	Turkish (Mac)
'10082	x-mac-croatian	Croatian (Mac)
'12000	utf-32	Unicode (UTF-32)
'12001	utf-32BE	Unicode (UTF-32 Big-Endian)
'20000	x-Chinese-CNS	Chinese Traditional (CNS)
'20001	x-cp20001	TCA Taiwan
'20002	x-Chinese-Eten	Chinese Traditional (Eten)
'20003	x-cp20003	IBM5550 Taiwan
'20004	x-cp20004	TeleText Taiwan
'20005	x-cp20005	Wang Taiwan
'20105	x-IA5	Western European (IA5)
'20106	x-IA5-German	German (IA5)
'20107	x-IA5-Swedish	Swedish (IA5)
'20108	x-IA5-Norwegian	Norwegian (IA5)
'20127	us-ascii	US-ASCII
'20261	x-cp20261	T.61
'20269	x-cp20269	ISO-6937
'20273	IBM273	IBM EBCDIC (Germany)

'20277	IBM277	IBM EBCDIC (Denmark-Norway)
'20278	IBM278	IBM EBCDIC (Finland-Sweden)
'20280	IBM280	IBM EBCDIC (Italy)
'20284	IBM284	IBM EBCDIC (Spain)
'20285	IBM285	IBM EBCDIC (UK)
'20290	IBM290	IBM EBCDIC (Japanese
katakana)		
'20297	IBM297	IBM EBCDIC (France)
'20420	IBM420	IBM EBCDIC (Arabic)
'20423	IBM423	IBM EBCDIC (Greek)
'20424	IBM424	IBM EBCDIC (Hebrew)
'20833	x-EBCDIC-KoreanExtended	IBM EBCDIC (Korean
Extended)		
'20838	IBM-Thai	IBM EBCDIC (Thai)
'20866	koi8-r	Cyrillic (KOI8-R)
'20871	IBM871	IBM EBCDIC (Icelandic)
'20880	IBM880	IBM EBCDIC (Cyrillic
Russian)		
'20905	IBM905	IBM EBCDIC (Turkish)
'20924	IBM00924	IBM Latin-1
'20932	EUC-JP	Japanese (JIS 0208-1990 and
'0212-1990)		
'20936	x-cp20936	Chinese Simplified (GB2312-
'80)		
'20949	x-cp20949	Korean Wansung
'21025	cp1025	IBM EBCDIC (Cyrillic
Serbian-Bulgarian)		
'21866	koi8-u	Cyrillic (KOI8-U)
'28591	iso-8859-1	Western European (ISO)
'28592	iso-8859-2	Central European (ISO)
'28593	iso-8859-3	Latin 3 (ISO)
'28594	iso-8859-4	Baltic (ISO)
'28595	iso-8859-5	Cyrillic (ISO)
'28596	iso-8859-6	Arabic (ISO)
'28597	iso-8859-7	Greek (ISO)
'28598	iso-8859-8	Hebrew (ISO-Visual)
'28599	iso-8859-9	Turkish (ISO)
'28603	iso-8859-13	Estonian (ISO)
'28605	iso-8859-15	Latin 9 (ISO)
'29001	x-Europa	Europa
'38598	iso-8859-8-i	Hebrew (ISO-Logical)
'50220	iso-2022-jp	Japanese (JIS)
'50221	csISO2022JP	Japanese (JIS-Allow 1 byte
Kana)		
'50222	iso-2022-jp	Japanese (JIS-Allow 1 byte
Kana - S0/SI)		
'50225	iso-2022-kr	Korean (ISO)
'50227	x-cp50227	Chinese Simplified
(ISO-2022)		
'51932	euc-jp	Japanese (EUC)
'51936	EUC-CN	Chinese Simplified (EUC)

'51949	euc-kr	Korean (EUC)
'52936	hz-gb-2312	Chinese Simplified (HZ)
'54936 (GB18030)	GB18030	Chinese Simplified
'57002	x-iscii-de	ISCII Devanagari
'57003	x-iscii-be	ISCII Bengali
'57004	x-iscii-ta	ISCII Tamil
'57005	x-iscii-te	ISCII Telugu
'57006	x-iscii-as	ISCII Assamese
'57007	x-iscii-or	ISCII Oriya
'57008	x-iscii-ka	ISCII Kannada
'57009	x-iscii-ma	ISCII Malayalam
'57010	x-iscii-gu	ISCII Gujarati
'57011	x-iscii-pa	ISCII Punjabi
'65000	utf-7	Unicode (UTF-7)
'65001	utf-8	Unicode (UTF-8)

From:

<https://wiki.plecko.hr/> - **Eureka Moment**

Permanent link:

https://wiki.plecko.hr/doku.php?id=development:vbs:sr_enc

Last update: **2019/10/31 09:05**

