Subject: Qtf and non UTF8 strings
Posted by koldo on Fri, 16 Oct 2015 07:07:34 GMT
View Forum Message <> Reply to Message

Hello all

Sometimes for example handling file names, it is necessary to show in QTF strings that have some characters that are not UTF8. Unfortunately the QTF output gets broken with an ugly bold non UTF8 error message.

Is it possible to transform a dubious string into a right UTF8 string by, for example, substituting wrong chars with a symbol (a square or other)?

Subject: Re: Qtf and non UTF8 strings
Posted by forlano on Fri, 16 Oct 2015 08:50:41 GMT
View Forum Message <> Reply to Message

koldo wrote on Fri, 16 October 2015 09:07

Is it possible to transform a dubious string into a right UTF8 string by, for example, substituting wrong chars with a symbol (a square or other)?

Hello,

this is an annoying problem even with the file content. The proposed solution would be very nice and easily recover the error in the file.

Luigi

Subject: Re: Qtf and non UTF8 strings
Posted by koldo on Fri, 16 Oct 2015 21:23:47 GMT
View Forum Message <> Reply to Message

File Attachments
1) window.jpg, downloaded 624 times

Subject: Re: Qtf and non UTF8 strings
Posted by koldo on Sun, 18 Oct 2015 09:20:33 GMT
View Forum Message <> Reply to Message

I mean that instead of output this ugly error message and truncate string, it could be preferable to output all string replacing strange characters with a symbol (a square or other). This could be

done by an U++ function called ConvertToUf8() (or a better name :))

Subject: Re: Qtf and non UTF8 strings
Posted by mirek on Sun, 18 Oct 2015 14:14:53 GMT
View Forum Message <> Reply to Message

koldo wrote on Fri, 16 October 2015 09:07Hello all

Sometimes for example handling file names, it is necessary to show in QTF strings that have some characters that are not UTF8. Unfortunately the QTF output gets broken with an ugly bold non UTF8 error message.

I really doubt the "necessary" part here...

Quote:
Is it possible to transform a dubious string into a right UTF8 string by, for example, substituting wrong chars with a symbol (a square or other)?

Anyway, this is perhaps possible. Still, I would say error should be displayed....

Mirek

Subject: Re: Qtf and non UTF8 strings
Posted by koldo on Mon, 19 Oct 2015 06:31:11 GMT
View Forum Message <> Reply to Message

Quote:Anyway, this is perhaps possible. Still, I would say error should be displayed....Probably the final user wants to read a string as clear as possible, and s/he does not know about UTF8 or internal character codifications that are more devoted to programmer and debugging.
Anyway if it is necessary to have this UTF8 message, I would ask for a function to convert a string to a fully readable one so it would be as complete as possible for an end user.

Subject: Re: Qtf and non UTF8 strings
Posted by forlano on Mon, 19 Oct 2015 06:44:40 GMT
View Forum Message <> Reply to Message

koldo wrote on Mon, 19 October 2015 08:31Probably the final user wants to read a string as clear as possible, and s/he does not know about UTF8 or internal character codifications that are more devoted to programmer and debugging.

I can support this point of view. Most of the users know nothing about UTF8. If the error appear in the content of a qtf file the visual effect is terrible (part of the file is shown and a part is missing or

with table bad formatted). He simply go in panic and is not able to recover the problem (just change the bad character).

Luigi

Subject: Re: Qtf and non UTF8 strings
Posted by mirek on Mon, 19 Oct 2015 11:08:25 GMT
View Forum Message <> Reply to Message

koldo wrote on Mon, 19 October 2015 08:31Quote:Anyway, this is perhaps possible. Still, I would say error should be displayed....Probably the final user wants to read a string as clear as possible, and s/he does not know about UTF8 or internal character codifications that are more devoted to programmer and debugging.
Anyway if it is necessary to have this UTF8 message, I would ask for a function to convert a string to a fully readable one so it would be as complete as possible for an end user.

User does not now about UTF8, but programmer does. IMO is it programmer's error to put non-UTF8 into QTF, something that should be fixed. Hence the error to alert the programmer.

Still, maybe I am missing context. Where that non-UTF8 string came from?

Mirek

Subject: Re: Qtf and non UTF8 strings
Posted by forlano on Mon, 19 Oct 2015 13:48:54 GMT
View Forum Message <> Reply to Message

mirek wrote on Mon, 19 October 2015 13:08
Still, maybe I am missing context. Where that non-UTF8 string came from?
Mirek

Good question. Even I do not understand how the problem is originated :) .

Anyway, for my program, the user import data from his national database or from the file he got somewhere (for example the German database has some special characters that do not appear in the standard latin alphabet). It seems that the import went OK because I can see the character in an ArrayCtrl. However when the output is sent to a QTF file something goes wrong. Usually are accented characters.
Perhaps the problem is the internal conversion of the file where the data is taken. The problem do not occur if the special character is typed within the application and sent to qtf file.

If exists the possibility to intercept the error and substitute the unknown code with something else I would do it myself.

Luigi

Subject: Re: Qtf and non UTF8 strings
Posted by koldo on Tue, 20 Oct 2015 08:42:38 GMT
View Forum Message <> Reply to Message

In my case the problem comes when reading file names in Windows. These files have accented names, and were copied first from Windows to Linux and then copied back to Windows. Windows does not complain and can open or copy these files without problems, but I cannot show its names using QTF.

Subject: Re: Qtf and non UTF8 strings
Posted by mirek on Tue, 20 Oct 2015 12:48:50 GMT
View Forum Message <> Reply to Message

koldo wrote on Tue, 20 October 2015 10:42In my case the problem comes when reading file names in Windows. These files have accented names, and were copied first from Windows to Linux and then copied back to Windows. Windows does not complain and can open or copy these files without problems, but I cannot show its names using QTF.

Now, there might be error in U++... If you are using U++ to read those file names, e.g. FindFile, and you have default charset (UTF-8), then you should already get them in UTF-8 (from FindFile).

Is this the case?

Subject: Re: Qtf and non UTF8 strings
Posted by koldo on Thu, 29 Oct 2015 10:01:06 GMT
View Forum Message <> Reply to Message

Hello Mirek

Problem found. Simplest code:

sequence"ToSystemCharset() is used as the program takes the file to compress it with 7z using Sys("7z.exe " + ToSystemCharset(fileName)).

Now in case of error the program would show in Exclamation() fileName instead of ToSystemCharset(fileName).

Thank you.

Subject: Re: Qtf and non UTF8 strings
Posted by sergeynikitin on Sat, 31 Oct 2015 14:06:50 GMT
View Forum Message <> Reply to Message

Try DeQtf() function before call.
for ex.:
instead

put

Subject: Re: Qtf and non UTF8 strings
Posted by mirek on Sat, 31 Oct 2015 14:41:13 GMT
View Forum Message <> Reply to Message

sergeynikitin wrote on Sat, 31 October 2015 15:06Try DeQtf() function before call.
for ex.:
instead

put

Well, that was not a cause here. However, speaking of DeQtf, in most cases you can use "\1" escape intead.

Exclamation("\1" + filename + "\1");

(In most case -> as long as you can be reasonably sure there is no \1 in the string... :)

Subject: Re: Qtf and non UTF8 strings
Posted by koldo on Sat, 31 Oct 2015 18:07:47 GMT
View Forum Message <> Reply to Message

sergeynikitin wrote on Sat, 31 October 2015 15:06Try DeQtf() function before call.
for ex.:
instead

put

es.pdf") works very well.

Subject: Re: Qtf and non UTF8 strings

Posted by koldo on Sat, 31 Oct 2015 18:10:37 GMT
View Forum Message <> Reply to Message

mirek wrote on Sat, 31 October 2015 15:41sergeynikitin wrote on Sat, 31 October 2015 15:06Try DeQtf() function before call.
for ex.:
instead

put

Well, that was not a cause here. However, speaking of DeQtf, in most cases you can use "\1" escape intead.

Exclamation("\1" + filename + "\1");

(In most case -> as long as you can be reasonably sure there is no \1 in the string... :)
Great ! :)

"\1" + filename + "\1" is bullet proof.