

## Bugs & Wish list

Unicode characters that make life of translators harder | Tiki Wiki CMS Groupware :: Development

Unicode characters that make life of translators harder

### Status

● Open

### Subject

Unicode characters that make life of translators harder

### Version

18.x

### Category

- Usability
- Less than 30-minutes fix
- Easy for Newbie Dev

### Feature

Admin Interface (UI)

### Resolution status

New

### Submitted by

hman

### Lastmod by

hman

### Rating

★★ (0) ?

### Description

It took me some time to find out why I could not translate a particular string in /lib/prefs/freetags.php, and this is it:



```
&#039;freetags_normalized_valid_chars&#039; =&gt; [
    &#039;name&#039; =&gt; tra(&#039;Valid characters
pattern&#039;),
    &#039;description&#039; =&gt; tra(&#039;Click on the links
below to set or clear a pattern to limit characters accepted in tags. &#039;),
    &#039;type&#039; =&gt; &#039;text&#039;,
    &#039;size&#039; =&gt; &#039;30&#039;,
    &#039;hint&#039; =&gt; &#039;Useful to eliminate characters
such as &ldquo;&,&rdquo; which users can enter by mistake instead of a space.&#039;,
    &#039;default&#039; =&gt; &#039;&#039;,
],
```

This looks decent, it also carries calls to the tra() subroutine. But I could not catch it with this entry in custom.php:



```
&quot;Useful to eliminate characters such as &ldquo;&,&rdquo; which users can enter
by mistake instead of a space.&quot; =&gt; &quot;N&uuml;ml;tzlich, um Zeichen zu
```

eliminieren wie Kommata, die Benutzer anstelle von Leerzeichen eingeben  
&ouml;n",

The string for pattern matching is exactly identical to its declaration in the source code, but it was not caught. Until finally I pumped the string into GHex, to look at its real composition.



```
echo -n "Useful to eliminate characters such as &ldquo;;&rdquo; which users  
can enter by mistake instead of a space.&quot; | od -A n -t x1  
55 73 65 66 75 6c 20 74 6f 20 65 6c 69 6d 69 6e  
61 74 65 20 63 68 61 72 61 63 74 65 72 73 20 73  
75 63 68 20 61 73 20 e2 80 9c 2c e2 80 9d 20 77  
68 69 63 68 20 75 73 65 72 73 20 63 61 6e 20 65  
6e 74 65 72 20 62 79 20 6d 69 73 74 61 6b 65 20  
69 6e 73 74 65 61 64 20 6f 66 20 61 20 73 70 61  
63 65 2e
```

So what's looking decent are actually Unicode characters



```
e2 80 9c 2c e2 80 9d
```

I have absolutely no idea how to escape that so PHP will find it in pattern matching... Devs, please do not use Unicode unless really really necessary...

Importance

3

Easy to solve?

10 easy

Priority

30

Demonstrate Bug (Tiki 19+)

Please demonstrate your bug on [show2.tiki.org](http://show2.tiki.org)

Version: trunk ▼

Ticket ID

7452

Created

Friday 31 July, 2020 17:08:49 GMT-0000

by hman

LastModif

Friday 31 July, 2020 17:08:49 GMT-0000

Comments



Same uncatchable Unicode in /lib/prefs/feature.php line 2927



```
&#039;feature_typo_quotes&#039; =&gt; [  
    &#039;name&#039; =&gt; tra(&#039;Smart &ldquo;curly&rdquo;  
quotes&#039;),  
    &#039;description&#039; =&gt; tra(&#039;Convert  
\&quot;straight&#039; quotes to &ldquo;curly&rdquo; ones, also convert &rsquo;  
apostrophe and &lsquo;single&rsquo; quotes&#039;),  
    &#039;type&#039; =&gt; &#039;flag&#039;, ,  
    &#039;tags&#039; =&gt; [&#039;basic&#039;],  
    &#039;default&#039; =&gt; &#039;y&#039;, ,  
    &#039;dependencies&#039; =&gt; [  
        &#039;feature_typo_enable&#039;, ,  
    ],
```

Cannot be translated with something like this in custom.php:



```
&quot;Convert \&quot;straight&#039; quotes to &ldquo;curly&rdquo; ones, also convert  
&rsquo; apostrophe and &lsquo;single&rsquo; quotes&quot; =&gt; &quot;Konvertiert  
\&quot;einfache&#039; Anf&uuml;hrungszeichen in typografische, sowie Apostrophe und  
einfache Anf&uuml;hrungszeichen.&quot;, ,
```

## Attachments

&nbsp;	filename	created	hits	comment	version	filetype	&nbsp;
--------	----------	---------	------	---------	---------	----------	--------

No attachments for this item

The original document is available at

<https://dev.tiki.org/item7452-Unicode-characters-that-make-life-of-translators-harder>