

Re: Shift_JIS conflict with CSS?

Source: <http://sci.tech-archive.net/Archive/sci.lang.japan/2004-09/0157.html>

From: Shez (*UseReplyAddress_at_nospam.uk.invalid*)

Date: 09/01/04

Date: Wed, 1 Sep 2004 20:38:18 +0100

In the faraway land of sci.lang.japan, dc <d3ntaku@gmail.com> said:

>> *Anyhow, if anyone has tips on using Open Office (western version) for
>> mixed language documents and webpages that include Japanese, I'd be
>> interested to hear them. My needs are basically for English stuff with
>
>I find OpenOff/write basically creates awful HTML, the preview is
>nothing like a browser, even for basic word-processing, and it rips
>out style sheet references and reformats everything.*

That's odd, I found the preview looked almost exactly the same as in Firefox and IE. However the fact that it replaces external stylesheets with internal ones is indeed a pain, as I have to reinstate the "link" metatag each time.

My previous experience with WYSIWYG programmes was that they wrote dreadfully convoluted HTML, however with OO the only cluttering it did was because it was trying to make my HTML compliant by sticking <p>...</p> round any loose bits of text and repeating my font tags within every paragraph. I usually stick in one global font tag that spans from <body> to </body>, but despite its elegance this is unfortunately considered bad HTML. (One of my niggles with the evolution of HTML is that elegant and intuitive constructions are being replaced by obscure long winded ones, e.g.

```
background=pic.jpg  
by  
STYLE="background: url(pic.jpg) repeat scroll"
```

>Nvu does a good job for S-Jis pages:
><http://www.nvu.com/>

I don't use Shift_JIS any more as I found it caused my other special characters such as opening & closing quotes, em-dashes and so on being rendered as mojibake. UTF-8 fixes all this, but until OO the only editor I had which understood UTF-8 was JWPce.

>but for UTF8 my text editor comes up with gibberish (*editplus*).

sci.lang.japan: Re: Shift_JIS conflict with CSS?

Ordinary ASCII type editors show UTF-8 as triplets or pairs of bytes, however this isn't a problem as long as you didn't want to insert any new unicode or proof-read what's there: I find I can cut and paste the stuff and resave the file without any corruption of the unicode. So for general tweaking of webpages I often just use Notepad or somesuch without any proble