

**From:** Markus Scherer <markus.scherer@jtcsv.com>

**Subject:** UCD filename too long: DerivedNormalizationProperties.txt

One (only one) of the filenames in the Unicode Character Database is too long for MacOS (9.2) systems.

MacOS filenames can be at most 31 characters long.

DerivedNormalizationProperties.txt has a 34-character filename.

1234567890123456789012345678901234

1                      2                      3

This is an issue for

- Mac users downloading UCD files
- processing UCD files on MacOS with scripts & tools that rely on the UCD filenames
- libraries like ICU that contain UCD source files for data compilation

I propose that the UTC resolve to

- 1) set a policy that UCD filenames should be no longer than 31 characters
- 2) change this filename in future Unicode versions ( $\geq 3.2$ ) to something shorter

For 2) there are of course multiple reasonable ways to shorten the name.

- Just this one filename could be shortened to, e.g.,  
"DerivedNormalizationProps.txt" [my personal favorite] or  
"DerivedNormProperties.txt"                      or  
"DerivedNormProps.txt"  
1234567890123456789012345678901  
1                      2                      3

- The prefix on all "Derived" files could be shortened to "Der" or similar;

however, this would change several filenames, while only one is really problematic.

- etc.

If the UTC does not resolve to change this filename, then some users of the UCD will need to come up with ad-hoc, different, unstandardized, filenames. This could lead to confusion.

Thanks and best regards,  
markus