Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**

**Title:** **Review of Proposal to encode 114 Tangut components**

Source: **Andrew West**

Status: **Individual Contribution**

Action: **For consideration by JTC1/SC2/WG2 and UTC**

Date: **2024-01-29**

I carried out a detailed review of the *Proposal to encode 114 Tangut components* (WG2 N5218R2 = L2/23-247), and found two issues:

1. Tangut Component 783 (U+18D8E) 퐈 (퐈) is equivalent to and unifiable with the existing Tangut Component 324 (U+18943 퐈), and so should be removed.

2. Tangut Component 822 (U+18DB5) has the wrong glyph. The glyph shown in the proposal document is 恒 (4 strokes), which is identical to the existing Tangut Component 104 (U+18867 恒). The correct glyph should be 恒 (8 strokes), as shown in the evidence for RR356 on p. 32.

Therefore the repertoire of Tangut components in the new Tangut Components Supplement block should be modified from 116 characters in the range U+18D80..U+18DF3 (TANGUT COMPONENT-769 through TANGUT COMPONENT-884) to 115 characters in the range U+18D80..U+18DF2 (TANGUT COMPONENT-769 through TANGUT COMPONENT-883).

I can provide a revised font with the required glyph correction, and 18D8F..18DF3 moved to 18D8E..18DF2.