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

Doc Type:Working Group DocumentTitle:Proposal to encode one Tangut ideographSource:Andrew West and Viacheslav ZaytsevStatus:Individual ContributionAction:For consideration by JTC1/SC2/WG2 and UTCDate:2025-03-28

# 1. Introduction

We recently carried out a comprehensive analysis all 6,150 entries in the nine-volume dictionary of Tangut words, *Xīxiàwén Cídiǎn* (*shìsú wénxiàn bùfēn*) 西夏文词典(世俗文献部 分) (Běijīng: Zhōngguó shèhuì kēxué chūbǎnshè, 2021), compiled by Hán Xiǎománg 韓小 忙. This important new dictionary has already been the source for several submissions to WG2 and the UTC:

- *Proposal to encode 113 Tangut components* by West (WG2 N5218R3 = L2/23-150)
- *Glyph corrections for four Tangut ideographs* by West (WG2 N5237 = L2/23-148)
- *Proposal to encode two Tangut ideographs* by Chan, You, and Zaytsev (WG2 N5286 = L2/24-236)

The Tangut ideographs, Tangut components, and glyph changes proposed in these three documents have been incorporated into Unicode version 17.0.

In this and one other related document (WG2 N5305: *Glyph changes for 14 Tangut ideographs*) we aim to document and resolve any outstanding discrepancies between the Tangut character set in Hán Xiǎománg's dictionary (hereafter *XXWCD*) and the set of Tangut ideographs encoded in the ISO/IEC 10646 and Unicode standards. In this document we document cases where *XXWCD* has two or three entries for glyph variants corresponding to a single encoded Tangut ideograph, and propose to encode one Tangut ideograph which we do not think should be treated as a glyph variant.

#### 2. Discussion

There are 147 entries in *XXWCD* which are unencoded variant forms of another entry in the dictionary, and a further six entries for variant glyph forms which are coincidentally identical to the glyph form of a different unrelated Tangut ideograph (see entries for 17009, 17230,

1774A, 177DE, 17AF5, and 187CD in the Appendix).<sup>1</sup> In total, 295 entries in *XXWCD* correspond to 148 Unicode characters (143 Unicode characters correspond to two entries in *XXWCD*, and five Unicode characters correspond to three entries in *XXWCD*). For example, *XXWCD* has entries for  $\mathfrak{A}$  (v. 1 p. 39) and  $\mathfrak{A}$  (v.1 p. 40), which are glyph variants of the Tangut ideograph meaning 'lion' which is encoded as U+17E5B  $\mathfrak{A}$  (the Unicode glyph corresponds to the entry on v. 1 p. 40). In most cases the difference between the variants is minor, and either involves a difference in writing a single component (as in the above example) or a difference in positioning of components (e.g. entries for  $\mathfrak{A}$ ,  $\mathfrak{A}$  and  $\mathfrak{A}$  corresponding to U+177D4  $\mathfrak{A}$ ; and entries for  $\mathfrak{B}$  and  $\mathfrak{B}$  corresponding to U+182C6  $\mathfrak{E}$ ). One particularly common variation is that between Component 26 -, Component 28 -, and Component 75 -, which are often written in a very similar manner in primary Tangut sources, with the result that *XXWCD* has three entries ( $\mathfrak{A}$ ,  $\mathfrak{A}$ , and  $\mathfrak{A}$ ) corresponding to U+17219  $\mathfrak{A}$ .

For reference, all the *XXWCD* entries for variants corresponding to the same Unicode character are listed in the Appendix at the end of this document. Almost all of these variant forms can probably be considered to be unifiable glyph variants, and could potentially be represented using Ideographic Variation Sequences (IVS).<sup>2</sup>

> > Fig. 1: *XXWCD* vol. 5 p. 96

<sup>&</sup>lt;sup>1</sup> An example of a glyph variant for one character corresponding to the glyph form for another character is the entry for 黻 (v. 8 p. 469) which is a variant of U+1774A 萩 'roof beam' ( 萩 on v. 3 p. 142), but it is identical to the glyph for U+18736 談 'big' ( 談 on v. 8 p. 456).

<sup>&</sup>lt;sup>2</sup> Version 17.000 of the Tangut Yinchuan font maintained by West includes a provisional set of nearly 3,000 IVS for 2,319 Tangut ideographs, including all the glyph forms used in *XXWCD*. This set may be considered for registration with the Ideographic Variation Database (IVD) at a future date.

MSA(甲)	MS B (乙)	MS D (丁)	MS E (戊)	MS F (己)
漏	A.	霸	荔	大
爒	貓	蒳	颏	<i>嬌</i> ?

Table 1: Glyphs for *tsia<sup>2</sup>* in Manuscripts of *Joined Rimes of the Five Sounds* 

<b>Table 2: Proposed New</b>	<sup>,</sup> Tangut Ideograph
------------------------------	-------------------------------

Code Point	Glyph	IDS	Radical/ Stroke	Source Reference	<i>XXWCD</i> Vol./Page	<i>XXWCD</i> Glyph
*18D1F	麱	Ⅲ袤⊟৺Ⅲ乡肓	210.14	H2021-509604	5/96	茒

<sup>&</sup>lt;sup>3</sup> This variant is included in *XXWCD* as  $\ddot{m}$  on v. 2 p. 300, and as it is a simple repositioning of the top right component to cover the left component it can be considered to be a unifiable glyph variant of U+17A11.

# 3. Code Point Allocation and Character Properties

# 3.1. Code Point Allocation

To encode new (recently identified) Tangut ideographs, there are free code points only in the Tangut Supplement block (U+18D00–U+18D7F). The last occupied code point in this block is U+18D1E. This code point was occupied by the Tangut ideograph, which was proposed to be encoded in Proposal WG2 N5286 = L2/24-236 (see References). In this regard, we propose to encode the new character using the following free code point \*U+18D1F.

# 3.2. Unicode Character Properties

The entries for Tangut ideographs in the UnicodeData.txt file should be amended to:

```
18D00;<Tangut Ideograph Supplement, First>;Lo;0;L;;;;N;;;;
18D1F;<Tangut Ideograph Supplement, Last>;Lo;0;L;;;;N;;;;
```

#### 3.3. Tangut Sources

The following entries should be added to the TangutSources.txt file:

U+18D1F kTGT\_MergedSrc H2021-509604 U+18D1F kRSTUnicode 210.14

# 4. References

#### **4.1. Technical Documents**

N5218R3. Andrew West. Proposal to encode 113 Tangut components. JTC1/SC2/WG2 N5218R3. 2024-06-12.

URL: https://www.unicode.org/wg2/docs/n5218r3-Tangut-components.pdf

**N5218R**. Andrew West. *Proposal to encode 114 Tangut components*. JTC1/SC2/WG2 N5218R = L2/23-247. 2023-10-02.

URL: <u>https://www.unicode.org/wg2/docs/n5218r2-Tangut-components.pdf</u> URL: <u>https://www.unicode.org/L2/L2023/23247-wg2-n5218r2.pdf</u>

**N5218**. Andrew West. *Proposal to encode 114 Tangut components*. JTC1/SC2/WG2 N5218 = L2/23-150. 2023-05-31.

URL: <u>https://www.unicode.org/wg2/docs/n5218-Tangut-components.pdf</u> URL: <u>https://www.unicode.org/L2/L2023/23150-n5218-tangut-components.pdf</u>

N5237. Andrew West. *Glyph corrections for four Tangut ideographs*. JTC1/SC2/WG2 N5218 = L2/23-148. 2023-06-21.

URL: <u>https://www.unicode.org/wg2/docs/n5237-Tangut-glyph-corrections.pdf</u> URL: <u>https://www.unicode.org/L2/L2023/23148-n5237-tangut-glyph-corr.pdf</u>

N5286. Eiso Chan (陈永聪/巍融绿), Jerry You (游程宇/豫藏茂), Viacheslav Zaytsev (IOM RAS). Proposal to encode two Tangut ideographs. ISO/IEC JTC1/SC2/WG2 N5286 = L2/24-236. 2024-10-14.

URL: <u>https://www.unicode.org/wg2/docs/n5286\_TwoTangutIdeographs.pdf</u> URL: <u>https://www.unicode.org/L2/L2024/24236-n5286-two-tangut-ideographs.pdf</u>

# 4.2. Tangut Sources

Joined Rimes of the Five Sounds. ngwə ·ệi we mbju (<a href="mailto:kita:mail

# 4.3. Academic Literature

- ECHW 7. Shǐ Jīnbō 史金波, Wèi Tóngxián 魏同賢, and E. I. Kychanov (chief eds.). É cáng Hēishuǐchéng wénxiàn 俄藏黑水城文獻 = Памятники письменности из Хара-Хото хранящиеся в России = Heishuicheng Manuscripts Collected in Russia. Vol. 7. Shànghǎi 上海: Shànghǎi gǔjí chūbǎnshè 上海古籍出版社, 1997. ISBN 7-5325-2213-X.
- Xià-Hàn Zìdiǎn. Lǐ Fànwén 李範文 (comp.) and Jiǎ Chángyè 賈常業 (rev. and exp.). Xià-Hàn zìdiǎn 夏漢字典 [Tangut-Chinese Dictionary (Revised Ed.)]. Běijīng 北京: Zhōngguó shèhuì kēxué chūbǎnshè 中國社會科學出版社, 2008. ISBN 978-7-5004-2113-9.
- XXWCD. Hán Xiǎománg 韩小忙 (comp.). Xīxiàwén cídiǎn (shìsú wénxiàn bùfēn) 西夏文词典 (世俗文献部分):全9册 [Tangut Word Dictionary (Secular Literature Part).9 vols.]. Běijīng 北京: Zhōngguó shèhuì kēxué chūbǎnshè 中国社会科学出版社, 2021. (Xīxià wénzì yǔ wénxiàn yánjiū 西夏文字与文献研究 [Studies of Tangut characters and documents]). ISBN 978-7-5203-8011-9.

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17000	اللہ اللہ	1.0	1/10	灵	LR1.8
17009	λĹ	1.8	1/174	烮	LR1.9
17000	IH	2.12	1/99	厦	LR2.12
1708D	<u></u> 図	2.12	1/100	厦	LR2.12
—————————————————————————————————————	送記	2.14	1/143	版范	LR2.14
17080	IIXIL		1/133	뛦	LR2.13
17050	磙	9.10	1/174	碗	LR4.10
17070			2/593	癍	LR47.10
17105	5	11 11	1/192	鬏	LR5.11
1710E	承	11.11	1/200	鬏	LR5.11
17118	声	11 14	1/197	糵	LR5.14
	<b>文</b> 龍	11.14	1/197	鼐	LR5.14

# Appendix: XXWCD Entries for Variants of the Same Unicode Character

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17150	厚	17.0	1/331	僱	LR7.9
17150	<b>川</b> 代	17.9	1/331	帰	LR7.9
17155		17.10	1/335	顺	LR7.10
1/155	川図	17.10	1/335	顺	LR7.10
124	旧书	17 1 2	1/391	凲	LR7.13
171A0	川井井	17.13	1/391	凲	LR7.13
	膖	17.14	1/398	시紙	LR7.14
171AD			1/398	腦	LR7.14
171B0	旧泛	17.14	1/402	腻	LR7.14
17100	加交		1/403	膨	LR7.14
171RD	庫	1716	1/409	廬	LR7.16
1118D	川風	17.16	1/409	牘	LR7.16
171D8		26.13	1/430	薮	LR10.13
	荻		1/462	薮	LR12.13

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17155		20.0	1/438	雿	LR12.9
17125	丞上	28.9	2/235	瀻	LR36.10
17164	涩	20.10	1/440	쥹	LR12.10
ITIEA	效	28.10	1/424	쥺	LR10.10
171EB	苏	20 10	1/441	蘷	LR12.10
	X4	28.10	2/252	颏	LR36.11
	烫	28.11	1/448	즃	LR12.11
1/17/			2/274	莈	LR36.12
17150	小		1/449	쥹	LR12.11
171179	18년	20.11	1/428	햢	LR10.11
17150		20 11	1/450	鼒	LR12.11
1/1/6	为伟	28.11	2/277	蘒	LR36.12
171FF	うな	20 1 2	1/451	乘	LR12.12
	蘒	28.12	1/455	義	LR12.12

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17200	南	20.12	1/458	藏	LR12.13
17206	2)\$(1)\$	20.15	1/429	藏	LR10.13
17200	以	20.12	1/460	藏	LR12.13
17200	井仪	28.13	1/465	藏	LR12.13
금	调	28.13	1/460	黿	LR12.13
17200	们民		1/461	黿	LR12.13
17212		28.13	1/465	覈	LR12.13
17212			2/346	藏	LR36.14
		28.14	1/466	藏	LR12.14
17219	藏		1/431	藏	LR10.14
			2/357	藏	LR36.15
1721A		28.14	1/468	義	LR12.14
	蔌		2/358	蕔	LR36.15

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17210		20.14	1/468	霰	LR12.14
17218	承	28.14	2/359	蔽	LR36.15
			1/468	灵	LR12.14
1721C	灵	28.14	1/431	蘣	LR10.14
			2/359	蘣	LR36.15
	禝	28.15	1/472	禝	LR12.15
17211			1/432	鱫	LR10.15
17000	覈	28.16	1/474	覈	LR12.16
17223			2/388	藗	LR36.17
			1/474	靋	LR12.17
17224	巍	28.17	1/474	巍	LR12.17
			2/392	巍	LR36.18
17228	<u>-1</u> 	20 10	1/477	藏	LR12.19
	夏	28.19	1/477	题	LR12.19

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17220	中兴	21.0	1/498	輚	LR13.8
17230	衩	31.8	3/4835	狡	LR88.10
17241	乏反	22.0	1/516	額	LR14.9
17241	℃代	33.9	1/516	额	LR14.9
17289	印	267	1/566	升	LR16.7
	ク日	36.7	1/565	現	LR16.7
	뚏	頻 36.12	1/644	凝	LR16.12
17260			1/644	巍	LR16.12
17211	西	36.14	1/671	また	LR16.14
17311	ブダラ		1/671	题	LR16.14
17210	经久	26.14	1/675	籁	LR16.14
1/310	ŹĪX	36.14	1/675	癙	LR16.14
17000	环	41.0	2/12	孩	LR19.9
1733C	核	41.9	2/14	Ŧ貁	LR19.10

 $^5$  This variant of U+17230 is identical to U+17BE2 (炎 on v. 3 p. 482).

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17202	武治	47.10	2/79	ĨĨŔ	LR23.13
17382	1111	47.13	2/78	ĪĨĨ	LR23.12
17202	小茶	69.0	2/159	姘	LR32.9
17505	X1/1	00.9	2/163	姘	LR32.10
	¥2	75 10	2/224	藉	LR36.10
1/442	ধর	75.10	1/438	彦	LR12.9
1746B	嚨	龍 75.11	2/262	硫	LR36.11
17400			2/288	硫	LR36.12
17449	詳	75.13	2/294	旆	LR36.13
17489	们存		8/127	旆	LR303.13
17480	游	75 12	2/297	藐	LR36.13
17400	<u> </u> 刹市	75.13	2/297	藐	LR36.13
174DC	滋	75 1 4	2/334	襜	LR36.14
	循	75.14	2/337	蓧	LR36.14

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17451	浙	75 14	2/335	颜	LR36.14
17461	<u> </u>	/5.14	2/352	颜	LR36.15
17540	<i>71</i>	7( 0	2/403	偁	LR37.8
17540	IA	76.8	2/404	偁	LR37.8
	미	70.0	2/411	副	LR39.9
1754E	有乡	79.9	2/450	罰	LR39.9
	瀫	79.15	2/449	燾	LR39.15
17374			6/299	瓢	LR169.15
1762B	周	92.10	2/593	즀	LR47.10
17020	知此		8/445	亂	LR390.10
17643	邗名	99.10	3/11	雃	LR50.10
17643	ЛН	99.10	3/12	頎	LR50.10
176AD	世	106.11	3/64	饚	LR55.11
	T.S.		3/64	谐	LR55.11

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17(00	书	10( 11	3/67	游	LR55.11
17688	街多	106.11	3/212	世子	LR55.11
1771D	批	10( 14	3/124	萒	LR55.14
17710	<u> </u>	106.14	3/192	襚	LR55.17
17725	***	106 14	3/138	珳	LR55.14
1773E	袤处	106.14	3/138	拔	LR55.14
	蔱	106.14	3/142	譢	LR55.14
1774A			8/4696	莃灸	LR395.14
17745	书	106 14	3/143	薻	LR55.14
17746	兆	106.14	3/119	荍	LR55.13
17745	达	106 17	3/196	蠽	LR55.17
177AF	鍫	106.17	3/196	薮	LR55.17
17700	世	106.18	3/199	菜	LR55.18
1//88	義		3/199	業	LR55.18

Unicode 17.0 Alpha		XXWCD				
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S	
			3/208	讌	LR55.21	
177D4	ゴ	106.21	3/208	瀻	LR55.21	
			3/211	虃	LR55.21	
177DF	唇花	107.16	8/34	齀	LR284.16	
ITTDE	<b>TTP</b>	107.10	5 8/5777 <b>#</b>	LR435.16		
17754	祀	記	100 11	3/225	融	LR56.11
177EA		108.11	3/227	祇	LR56.12	
17050	反赛	112.15	3/289	國	LR58.15	
17850	KØ	112.15	3/289	顑	LR58.15	
17800	们天	138.0	3/350	很	LR68.9	
17070	11久	138.9	3/369	妪	LR74.9	
17001	们在	120.0	3/350	化	LR68.9	
1/091	11/2	130.7	3/369	桃	LR74.9	

 $<sup>^7</sup>$  This variant of U+177DE is identical to U+187CD  $\Lambda$  (靴 on v. 8 p. 576).

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17002	公正	120.0	3/350	循	LR68.9
17892	1117	138.9	3/369	全	LR74.9
17002	派	120.10	3/350	í	LR68.10
17893	1113	138.10	3/369	邂	LR74.10
17004	<b>刈</b> ヨ	120.10	3/351	循	LR68.10
17894	1177	7 138.10	3/369	4月	LR74.10
1700(	能	120.10	3/354	能	LR68.10
17896		旧匕 138.10	3/370	瘧	LR74.10
17007	御	120 11	3/355	彽	LR68.11
17897	IIIX	150.11	3/369     XII3       3/351     1福       3/351     4福       3/369     X福       3/354     1信       3/355     11       3/355     11       3/370     X13       3/355     11       3/356     11       3/370     X14       3/356     11       3/370     X14	LR74.11	
17000	训诂	120 11	3/356	鄁	LR68.11
1/077	IIFT	138.11	3/370	渊	LR74.11
17000	泅,	120.11	3/358	毹	LR68.11
1/876	1111 - 138.11 -	3/370	紪	LR74.11	

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17005	<i>6</i> 1441	120 12	3/358	徽	LR68.12
17895	11811	138.12	3/370	娰	LR74.12
17005	刘作	120.12	3/359	褦	LR68.12
1789F	TIXE	138.12	3/370	姚	LR74.12
17042	狄	120.12	3/360	彽	LR68.13
178A3	TKL	138.13	3/370	娫	LR74.13
1701(	665	141.10	4/123	颏	LR70.10
17916	881	141.10	4/217	纐	LR70.11
17002	从各	1 4 1 1 1	4/264	綷	LR70.11
17982	<u></u> ጸ۴ጉ	141.11	Vol./PageGlyph3/35813/35813/37023/35913/37023/36013/37023/37024/12324/21724/26424/26524/26524/26524/17624/36024/6152	熧	LR71.11
1700(	が三	1 4 1 1 1	4/265	縃	LR70.11
1/980	ХэП	141.11	4/176	紤	LR70.10
17004	历史	141.12	4/360	嫁	LR70.12
1/9DA	久永		4/615	篴	LR71.12

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
	刘省	141 12	4/379	猟	LR70.13
1/9/7	刘伟	141.15	4/378	猟	LR70.13
17050	从冬	141 12	4/386	颔	LR70.13
1796	XXI	141.15	4/447	銹	LR70.14
17400	が言	141 10	4/400	谻	LR70.13
TAUC	刘代	141.13	4/453	떘	LR70.14
			4/404	鵕	LR70.13
17A11	貓	141.13	2/300	蒳	LR36.13
			5/96	奫	LR116.14
17425	姤	141 12	4/428	纐	LR70.13
17A35	<u> </u>	141.13	4/362	纯	LR70.12
17426	加合	141 10	4/428	纐	LR70.13
1/A36	刘元	141.13	4/364	纐	LR70.12

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17400	必冬	14110	4/536	綐	LR70.16
17AC8	知代	141.16	4/536	縑	LR70.16
17455	んじ	142.0	4/26 <sup>8</sup>	紁	LR70.8
1745	X	142.8	4/589	災	LR71.8
17000	公商	142.15	4/616	颏	LR71.15
17000	<b>火</b>	142.15	4/616	颏	LR71.15
17045	奋	152.15	3/397	慮	LR79.15
17645		152.15	3/397	新     類       397     窗       397     窗       397     බ	LR79.15
17 <b>8</b> 4E	古红	台灯 1-5-10	3/410	ሸ	LR80.10
17046	IIXI	155.10	3/413	ሸ	LR80.11
17205	彩巾	170.7	3/474	靴	LR87.7
5007	116	170.7	3/482	犼	LR88.7
17000	初古	170.10	3/477	耞	LR87.10
17000	1111	170.10	3/482	釽	LR88.10

Unicode 17.0 Alpha		XXWCD				
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S	
17005	*	105 12	3/610	荻	LR99.12	
17065	災	185.12	3/609	荻	LR99.12	
17040	垂	105 16	6/410	蘒	LR177.16	
17049	徊	185.16	6/411	龙	LR177.16	
17040	我	105 16	3/648	蕔	LR99.16	
17CAB	兼後	185.16	3/647	蒹	LR99.16	
15005	Ħ	Ŧ	20(12	5/9	匙	LR112.13
17037	刑	206.13	8/559	型	LR428.13	
17042	赉	206.14	5/23	璦	LR112.14	
17042	131	200.14	5/48	蕟	LR112.15	
17041	芳女	216 12	6/230	鍫	LR163.13	
	田	216.13	8/551	盭	LR422.13	
170.00	土	21715	5/129	캛	LR119.15	
117068	刻臣	21/.15	5/129	虦	LR119.15	

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17550	万夜	240.12	1/40	孫	LR1.13
17658	以宜	240.13	1/39	硺	LR1.13
17565	比象	242.12	5/257	胤	LR132.13
17665	肝剤	243.13	5/257	將	LR132.13
17507	记	256.0	5/296	飛	LR137.9
17687	书匕	256.9	6/409	飛	LR176.9
17500	戚		5/296	戚	LR137.11
17688		X 256.11	6/409	飛	LR176.11
17502	名比	260.12	5/324	縬	LR139.12
17203	<del>11</del> /KII	200.12	Vol./PageGlyph1/40成領1/39成領5/257月稿5/257月稿5/296紀6/409紀5/296紙6/409紙5/324紙5/316紙5/334編5/397千5/397千	郙	LR139.11
17664	得	260.14	5/334	륢	LR139.14
	<b>††</b> \$\$	260.14	5/334	靜	LR139.14
17505	炎	264.16	5/397	截	LR142.16
1/505	111刻	204.10	5/397	癫	LR142.16

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
10025	ন	270.12	6/129	娍	LR149.13
18035	XIIX	278.13	6/129	娍	LR149.13
10001	ンヨ	200.10	6/207	巍	LR161.10
18081	Xtt	298.10	6/212	鋭	LR161.11
1011D	歌	210.17	6/299	颤	LR169.17
1011D	汤们	新们 310.17	2/452	翿	LR39.17
10122	马灯	212 11	6/308	靫	LR171.11
10123	ĦΧΙ	512.11	6/309	弒	LR171.12
18124	司书	314.9	6/313	誹	LR172.9
1012A	1111	514.5	6/314	耏	LR172.9
18144	洏	216 12	6/357	颏	LR173.12
1014A	XIIIC	316.12	6/357	瓢	LR173.12
1016E	茄		6/396	霢	LR173.15
1010	XHII	510.15	6/397	窥	LR173.15

Unicode 17.0 Alpha		XXWCD				
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S	
10144	石灰	220.10	6/434	秤	LR181.10	
181A4	If I	328.10	7/210	靜	LR238.11	
10207	E7	245 10	6/497	箙	LR195.10	
18207	113	345.10	1/169	皴	LR4.10	
10206	- TD N	401.0	7/161	葭	LR232.8	
18200	这	401.8	7/161	陵	LR231.8	
10207	山田	古言	401 14	7/161	顔	LR232.14
10207	汉世	401.14	7/161	旗	LR231.14	
18320	影	416 12	3/429	截	LR83.12	
18320		410.12	7/234	靗	LR239.12	
10245	汉米	420.11	7/254	随	LR242.11	
10345	KU KU	420.11	7/243	隊都	LR241.11	
10200	当沿	421 12	7/297	詂	LR248.14	
10200	1))	431.13	7/296	髚	LR248.13	

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
10204	兴水	426.10	7/346	斜	LR251.10
183BA	<b>浆</b> 名	436.10	7/346	荻	LR251.10
10422	光石	426.15	7/448	豘	LR251.15
18432	<b>※1</b> 甩	436.15	7/449	纐	LR251.15
10466	己冬	446 10	7/476	隧	LR256.10
18466	\$3	446.10	8/128	餞	LR304.11
10467	司中	446 12	7/476	郦	LR256.13
18467	IX11FF	440.13	8/131	刻	LR304.14
10460	記	446 12	7/476	屁	LR256.13
18408	IX RE	440.13	Vol./PageGlyph7/346第7/346第7/346第7/448第7/449第7/476夏8/128夏7/476夏8/131夏8/131夏7/476夏8/131夏7/476夏7/476夏7/476夏8/131夏7/489夏7/493夏7/493夏7/493夏7/493夏7/493夏7/493夏7/493夏7/493夏7/493夏7/493夏	蒑庑	LR304.14
10470	73	452.12	7/489	羂	LR261.12
104/0	<u> </u>	452.12	7/493	璭	LR262.12
10470	73	452.12	7/491	源	LR261.13
104/8	<u> </u>	《异·452.13 —	7/493	烮	LR262.13

Unicode 17.0 Alpha		XXWCD				
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S	
10475	74	452.11	7/492	頲	LR262.11	
1847F	烮	453.11	7/487	죩	LR261.11	
10520	冬瓜	F00 1F	8/77	쥃	LR293.15	
18539	\$XIX	500.15	8/77	氦級	LR293.15	
195 <i>4</i> P	至	512 11	8/111	剙	LR301.11	
10340	MH	512.11	8/121	瓶	LR301.11	
19565	衠	谷伯	E26 10	8/143	衠	LR309.10
1030r		520.10	8/143	傢	LR309.10	
1867D	否公	574 11	8/334	餃	LR340.11	
10070	IXX	574.11	Vol./PageGlyph7/492夏伊7/487夏伊8/77夏級8/77夏級8/111仟仟8/121仟仟8/143夏分8/143夏分8/334夏気2/593夏交3/213万秋3/213夏秋8/354夏秋8/354夏秋	扊	LR47.11	
10605	世纪	505 17	3/213	뙍	LR55.17	
10001	IITX	585.17	3/213	澈	LR55.17	
19640	<b>店</b> 次	600 12	8/354	駁	LR354.12	
TODAC	ĦΧ	群久 600.12	8/354	料	LR354.12	

Unicode 17.0 Alpha		XXWCD				
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S	
10(00	目見	(01.1.(	8/362	見 人 人 人	LR355.16	
18688	××.	601.16	8/362	月月 亥灸	LR355.16	
10750	半公	604 12	8/492	敽	LR399.13	
16/30	FIX.	084.15	8/487	散	LR399.13	
19750	半小	694 15	8/494	黻	LR399.15	
10/50	<u></u> ΓXL	684.15	8/487	靗	LR399.15	
19770	II.	記	606 12	8/517	酛	LR406.13
18779	8L	090.13	8/517	酛	LR407.13	
19774	副	606 15	8/517	鬷	LR406.15	
10/74	感义	090.13	8/517	鬷	LR407.15	
10702	一开山	709 17	1/202	藐	LR5.17	
10/ 73	ЯI К	708.17	7/16	就能	LR198.17	
10707	西	720.11	8/574	夏	LR432.11	
10/0/	义	/ 30.11	8/575	貢	LR433.11	

Unicode 17.0 Alpha		XXWCD			
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
187C8	酸	730.17	8/574	國美	LR432.17
			8/575	酸	LR433.17
187C9	醈	730.19	8/575	國紀	LR432.19
			8/575	戭	LR433.19
187CA	授	731.22	3/213	翪	LR55.22
			3/213	<b>拨</b>	LR55.22
187CD	が上	733.16	8/35 <sup>9</sup>	齀	LR284.16
			8/576	齀	LR435.16
187D6	酸多	739.15	2/595	酸	LR47.15
			2/595	嚴	LR47.15
187DD	新	744.16	8/593	截	LR440.16
			8/592	截	LR440.16

# **Proposal Summary Form**

SO/IEC JTC 1/SC 2/WG 2						
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 1064610						
Please fill all the sections A, B and C below.						
Please read Principles and Procedures Document (P & P) from _ <u>http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html</u> _ for						
guidelines and details before filling this form. Please ensure you are using the latest Form from <a href="http://www.dkuug.dk/.ITC1/SC2/WG2/docs/summaryform.html">http://www.dkuug.dk/.ITC1/SC2/WG2/docs/summaryform.html</a>						
See also .http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html_ for latest <i>Roadmaps</i> .						
A. Administrative						
1. Title: Proposal to encode one Tangut ideograph						
2. Requester's name: Andrew West and Viacheslav Zaytsev						
3. Requester type (Member body/Liaison/Individual contribution): Individual contribution						
4. Submission date: <u>2020-03-28</u>						
6. Choose one of the following:						
This is a complete proposal: YES						
(or) More information will be provided later:						
B. Technical – General						
1. Choose one of the following:						
a. This proposal is for a new script (set of characters): NO						
Proposed name of script:						
b. The proposal is for addition of character(s) to an existing block: YES						
Name of the existing block. Tangut Suppement						
3. Proposed category (select one from below - see section 2.2 of P&P document):						
C-Major extinct D-Attested extinct						
F-Archaic Hieroglyphic or Ideographic X G-Obscure or questionable usage symbols						
A is a repertoire including character names provided?						
a. If YES, are the names in accordance with the "character naming guidelines"						
in Annex L of P&P document? YES						
b. Are the character shapes attached in a legible form suitable for review? YES						
5. Fonts related:						
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the						
standard?						
h Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site,	etc.):					
Andrew West	510.7.					
6. References:						
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?						
b. Are published examples of use (such as samples from newspapers, magazines, or other sources)						
of proposed characters attached?						
7. Special encoding issues:						
Does the proposal address other aspects of character data processing (if applicable) such as input,						
	0					
8. Additional Information:						
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or S	cript					
that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.						
Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour						
information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default						
Collation benaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization						
see Unicode Character Database ( http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports						
for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.						

<sup>&</sup>lt;sup>10</sup>. Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

#### C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	NO				
If YES explain					
2. Has contact been made to members of the user community (for example: National Body,	VES				
user groups of the script of characters, other experts, etc.)?	TES				
If YES, with whom?					
If YES, available relevant documents:					
3. Information on the user community for the proposed characters (for example:	NO				
size, demographics, information technology use, or publishing use) is included?	NO				
4. The context of use for the proposed characters (type of use; common or rare) Reference:	rare				
5. Are the proposed characters in current use by the user community?	YES				
If YES, where? Reference:					
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely					
in the BMP?	NO				
If YES, is a rationale provided?					
If YES, reference:					
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	YES				
8. Can any of the proposed characters be considered a presentation form of an existing					
character or character sequence?	NO				
If YES, is a rationale for its inclusion provided?					
If YES, reference:					
9. Can any of the proposed characters be encoded using a composed character sequence of either					
existing characters or other proposed characters?	NO				
If YES, is a rationale for its inclusion provided?					
If YES, reference:					
10. Can any of the proposed character(s) be considered to be similar (in appearance or function)					
to, or could be confused with, an existing character?					
If YES, is a rationale for its inclusion provided?					
If YES, reference:					
11. Does the proposal include use of combining characters and/or use of composite sequences?	NO				
If YES, is a rationale for such use provided?					
If YES, reference:					
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?					
If YES, reference:					
12. Does the proposal contain characters with any special properties such as					
control function or similar semantics?	NO				
If YES, describe in detail (include attachment if necessary)					
13. Does the proposal contain any Ideographic compatibility characters?	NO				
If YES, are the equivalent corresponding unified ideographic characters identified?					
If YES, reference:					