

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/IRG N2486

Date: 2021-05-07

Doc. Type:	Member body contribution
Title:	TCA's submission for the CJK extension IRG Working Set 2021
Source:	TCA
Status:	Input to IRG
Action:	For review by the IRG and WG2
Distribution:	IRG Member Bodies and Ideographic Experts
No. of pages:	4
Appendixes	4
Medium:	Electronic

1. Preface

TCA's submission for IRG Working Set 2021 consists of 1,000 ideographs. These ideographs are used for **personal name**, and came from Dept. of Household Registration, M.O.I. These ideographs are also collected into the Full Character Repertoire of CNS 11643. This submission is extremely important because Taiwan is planning a large New eID project.

For the personal name collection in the Full Character Repertoire of CNS 11643, there are about 10,000 ideographs are still outside of the UCS. Based on the following criteria, we selected 1,000 ideographs form that personal name collection as our submission:

- (1) higher usage frequency provide by Dept. of Household Registration,
- (2) may not be unified to existing UCS ideographs, and
- (3) prefer the length of IDS smaller than 5.

2. Evidence Explanation

TCA's Evidence is divided into two parts,

- (1) For the evidence 00:

The official document of Dept. of Household Registration, M.O.I., which proves that the 1,000 ideographs submitted by TCA are all from the household service database, and they are all highly used Characters for names.

- (2) For the evidence 01-04:

Dept. of Household Registration, M.O.I. extracted the usage data of the ideograph from the household service database, including CNS code, household registration internal code, ID,

and ideograph not yet included in the ISO/IEC 10646 appearing in the ID household name (or Spouse/ Father/ Mother/ Adoptive father/ Adoptive mother).

According to the Personal Identifiable Information Protection law, ID and name can only provide partial information. In addition, TCA are unable to provide individual scans of ID cards, in order to show the evidences provided by TCA are come from M.O.I., each page of the evidence is stamped with official seal. Hope IRG members can understand.

3.Documents

Submission Form

Appendix 1: IRGN2486_TCA_2021_sub_Attributes.xlsx

Appendix 2: IRGN2486_TCA_2021_sub_TrueType Font

Appendix 3: IRGN2486_TCA_2021_sub_Evidences

ISO/IEC JTC 1/SC 2/WG 2/IRG

**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITION OF CJK UNIFIED IDEOGRAPHS TO THE REPERTOIRE OF ISO/IEC 10646**

Submitters are reminded to:

1. Fill in all the sections below.
2. Read the Principles and Procedures Document (P & P) available at <https://appsrv.cse.cuhk.edu.hk/~irg/irg56/IRGN2424Confirmed.pdf> for guidelines and details before filling in this form.
3. Use the latest Form from <https://appsrv.cse.cuhk.edu.hk/~irg/irg56/IRGN2424SubmissionForm.xlsx>

See also <http://appsrv.cse.cuhk.edu.hk/~irg/irgwds.html> for the latest *Unifiable Component Variations*.

Administrative

1. IRG Project Code:	e.g. IRG Collection 2021
2. Title:	TCA's Submission for CJK-2021
3. Submitter's Region/Country Name:	TCA
4. Submitter Type (National Body/Individual Contribution):	Liaison member
5. Submission Date:	2021-05-07
6. Requested Ideograph Type (Unified or Compatibility Ideographs)	Unified Ideographs
If Compatibility, the submitter is strongly encouraged to instead register them as IVS in new or an existing IVD collection(See UTS #37) with the IRG's approval (Registration fee will not be charged if authorized by the IRG.).	
	N/A
7. Proposal Type (Normal Proposal or Urgently Needed)	Normal Proposal
8. Choose one of the following:	
This is a complete proposal.	Yes
(or) More information will be provided later.	

B. Technical – General

1. Number of ideographs in the proposal:	1,000
2. Glyph format of the proposed ideographs is in TrueType?	Yes
Are all the proposed glyphs put into BMP PUA area?	Yes
Are data for source references vs. character codes provided?	Yes
3. Source references:	
Do all the proposed ideographs have a unique, proper source reference (member body international consortium abbreviation followed by no more than 9 alphanumeric characters)?	Yes
4. Evidence:	
a. Do all the proposed ideographs have a separate evidence document which contains least one scanned image of printed materials (preferably dictionaries)?	Yes
b. Do all the printed materials used for evidence provide enough information to track the by a third party (ISBN numbers, etc.)?	Yes
5. Attribute Data Format: (Excel file or CSV text)	Excel

C. Technical - Checklist

Understanding of the Unification Principles

1. Has the submitter read ISO/IEC 10646 Annex S and does the submitter understand the unification principles? Yes
2. Has the submitter read the “Unifiable Component Variations” (contact the IRG technical editor through the IRG Convenor for the latest version) and does the submitter understand the unifiable variation examples? Yes
3. Has the submitter read the IRG PnP document and does the submitter understand the 5% Rule? Yes

Character-Glyph Duplication (<http://www.itscj.ipsj.or.jp/sc2/open/pow.htm> contains all the published ones and those under ballot)

4. Has the submitter checked that the proposed ideographs are **not unifiable** with any of the unified or compatibility ideographs of the latest version of ISO/IEC 10646? Yes

If the checking has been done against an earlier version of ISO/IEC 10646, please specify the version. (e.g. 10646:2012) 10646: 2017

5. Has the submitter checked that the proposed ideographs are **not unifiable** with any of the ideographs in the amendments, if any, of the latest version of ISO/IEC 10646? Yes

If yes, which amendment(s) has the submitter checked?

6. Has the submitter checked that the proposed ideographs are **not unifiable** with any of the ideographs in the proposed amendments, if any, of ISO/IEC 10646? Yes

If yes, which draft amendment(s) has the submitter checked?

7. Has the submitter checked that the proposed ideographs are **not unifiable** with any of the ideographs in the current working M-set and D-set of the IRG? (Contact IRG chief editor and technical editor through the IRG Convenor for the newest list) Yes

If yes, which document(s) has the submitter checked? Extension H

8. Has the submitter checked that the proposed ideographs are **not unifiable** with any of the over-unified or mis-unified ideographs in ISO/IEC 10646? (See Annex E of the IRG PnP document) Yes

9. Has the submitter checked whether the proposed ideographs have any **similar ideographs** in the current standardized or working sets mentioned above? Yes

10. Has the submitter checked whether the proposed ideographs have any **variant ideographs** in the current standardized or working sets mentioned above? Yes

Attribute Data

11. Do all the proposed ideographs have attribute data including the Kangxi radical code, stroke count, and first stroke(primary)? Yes

12. Do the proposed ideographs contain secondary radical code and their stroke count and first stroke are also provided? No

13. Are there any simplified ideographs (ideographs that are based on the policy described in 簡化字總表) among the proposed ideographs?

If yes, does the proposal include proper simplified/traditional indication flag for each proposed ideograph in the attribute data?

14. Do all the proposed ideographs have the document page number of evidence documents in the attribute data? Yes

15. Do all the proposed ideographs have the proper Ideographic Description Sequence (IDS) in the attribute data? Yes

If no, how many proposed ideographs do not have the IDS?

16. If the answer to question 9 or 10 is yes, do the attribute data include any information on similar/variant ideographs for the proposed ideographs?
17. Do all the proposed ideographs contain the total stroke count (kTotalStrokes)?

Yes

Yes