

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

Doc Type: Working Group Document
Title: Proposal on new CJK Unified Ideographs Extension for IRG
WS2021
Source: IRG
Status: Sub WG Contribution
Action: For consideration by JTC1/SC2/WG2
Date: 2024-06-04
Appendix: 4

This document proposes 4,300 CJK unified ideographs for encoding in UCS.

These characters are selected from the IRG Project Working Set 2021 (WS201), which was carried out in 2021, originally containing 4,952 characters. The characters were reviewed by IRG members and experts for no less than 7 rounds. Their attributes such as radicals and stroke counts, and review and discussion records could be found through IRG Online Review Tool (ORT, <https://hc.jsecs.org/irg/ws2021/app/>).

Source List:

- GDM Place name characters from the Public Order Administration, Ministry of Public Security, People's Republic of China
- GKJ Science and technology characters, such as names of geology, botanical and zoology terms
- GXM Human-name characters. Source from Public Order Administration, The Ministry of Public Security of the People's Republic of China
- GZ Ancient Zhuang Character Dictionary, (古壮字字典) 1989, ISBN 7-5363-0614-8
- T8 TCA-CNS 11643-1992 8th plane
- T9 TCA-CNS 11643-1992 9th plane
- T11 TCA-CNS 11643-1992 17th plane
- KC Korean History On-Line (한국 역사 정보 통합 시스템)
- SAT Taishō Shinshū Daizōkyō (大正新脩大藏經), 1924-1934
- UK The Unicode Technical Report #45, UK-source Ideographs, February 15, 2019
- UTC The Unicode Technical Report #45, U-source Ideographs, February 15, 2019
- V0 TCVN 5773:1993
- V1 TCVN 6056:1995
- V2 VHN 01:1998
- V3 VHN 02: 1998

V4 Kho Chữ Hán Nôm Mã Hoá (Hán Nôm Coded Character Repertoire), Hà Nội, 2007

VN Vietnamese Nôm characters not defined by V0 to V4 sources, major sources are: Nguyễn Quang Hồng. "Tự Điển Chữ Nôm Dẫn Giải". Hà Nội, 2014, and Anthony Trần Văn Kiệm "Giúp đọc Nôm và Hán Việt", Đà Nẵng, 2004

Note: Descriptions of V4 and VN are from ISO/IEC 10646:2020/Amd.1, others are from UCS2020.

Appendix 1: Summary Form

Appendix 2: Character List (IRG WS2021 v7.0 charts, PDF form)

Appendix 3: Character Attributes (XLS form)

Appendix 4: Fonts

End.