Universal Multiple-Octet Coded Character Set
International Organization for Standardization

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
Title: US/Unicode Activity Report for IRG #60
Source: Ken Lunde
Status: Member Body Contribution
Action: None
Date: 2023-02-07

The UTC (*Unicode Technical Committee*) and others in the US have been involved in the following standardization-related activities since the IRG #59 meeting took place:

1. The next version of the Unicode Standard will be Version 15.1, which is expected to be released on 2023-09-12. The only new characters that are expected to be included are five (5) new Ideographic Description Characters (IDCs) that can be used in Ideographic Description Sequences (IDSes). The Unicode Version 15.1 Alpha was issued as PRI (*Public Review Issue*) #473 on 2023-02-07, with a closing date of 2023-04-04.

2. The code charts for the CJK Unified Ideographs, CJK Unified Ideographs Extension A, and CJK Unified Ideographs Extension B blocks in Unicode Version 15.1 will *finally* include representative glyphs and source references for nearly 24,000 KP-source (aka *kIRG_KPSource* property) ideographs. Furthermore, the code charts for the CJK Unified Ideographs (aka URO) block will be enhanced to include a seventh column in order to accommodate KP-source ideographs in their own column.

3. The Unihan database that will be associated with Unicode Version 15.1 will include the following substantial changes (see PRI #467, *Proposed Update UAX #38*, for more details):

   - The provisional *kHKSCS*, *kIRGDaiKanwaZiten*, *kKPS0*, *kKPS1*, *kKSC0*, *kKSC1*, and *kRSKangXi* properties will be removed, because their property values are already covered by other properties, specifically *kIRG_HSource*, *kMorohashi*, *kIRG_KPSource*, *kIRG_KSource*, and *kRSUnicode*, respectively.
   - The provisional *kMorohashi* property, which currently covers less than 18,000 of the more than 51,000 entries in Japan's 大漢和辞典 ideograph dictionary, will be expanded to cover more than 49,000 of its entries.
   - The informative *kRSUnicode* property will be enhanced to specify non-Chinese simplified radicals, such as those used in Japan and Vietnam, through the use of two apostrophes ('').
   - The provisional *kJapanese*, *kMojiJoho*, *kSMSZD2003Index*, *kSMSZD2003Readings*, *kVietnameseNumeric*, and *kZhuangNumeric* properties will be added as new properties.

4. The US National Body submitted feedback and recommendations on the draft GB 18030-2022 amendment. See IRG N2591 (aka SC2 N4834, WG2 N5201, and L2/23-057). The only concern in this draft amendment is the encoding of nearly 900 ideographs in Plane 10, which is currently reserved in both ISO/IEC 10646 and the Unicode Standard.

5. UTN (*Unicode Technical Note*) #50, *KP-Source Property Value History*, was published as Version 1 on 2023-01-23. This new UTN documents changes that were made to the *kIRG_KPSource* (aka KP-source) property from Unicode Version 3.1.1 (2001) through Version 15.0 (2022), and will be kept up-to-date.

6. UTC experts reviewed and submitted comments against the UTC-assigned portion of IRG Working Set 2021 Version 4.0, specifically serial numbers 00001 through 01249, prior to the 2023-02-17 deadline, and will respond to comments against UTC-submitted ideographs prior to the 2023-03-16 deadline.

7. The UTC feels that it is prudent to reiterate what the UK stated in Section 2 of its Activity Report for IRG #59, IRG N2566, specifically that it is impossible to adequately review virtually all of the TCA-submitted ideographs due to the lack of sufficient evidence, and the review process of those ideographs is best described as frustrating. As a whole, the only aspect of these ideographs that can be meaningfully reviewed is their metadata, which has little to do with decisions related to unification. The UTC understands that these ideographs are used in personal names per IRG N2546 and IRG N2554R, but the disconnect between the the lack of actual/original evidence and the references to CNS 11643 and its representative glyphs is quite alarming, especially when considering the volume of such ideographs. The UTC therefore requests that the IRG Convenor clearly convey to IRG member bodies exactly what aspects of TCA-submitted ideographs in IRG Working Set 2021 are to be reviewed, because unification decisions and glyph-correctness cannot be adequately reviewed under such circumstances.

8. L2/22-279, *UCS Seal Script Source Mapping Data*, was posted to the UTC Document Register late last year, and serves as a new resource for Seal Script experts.

9. The Unicode CJK & Unihan Group, which reports to the UTC, continues to meet approximately two weeks prior to each quarterly UTC meeting. See L2/22-247 and L2/23-011 for the CJK & Unihan Group recommendations that were prepared for the UTC #173 and UTC #174 meetings, respectively. As a reminder, IRG experts are welcome to join the CJK & Unihan Group, which meets four times per year as a virtual Zoom meeting. The next CJK & Unihan Group meeting, in preparation for the UTC #175 meeting, is scheduled to take place on 2023-04-07 from 6 to 9PM PDT (GMT Minus 7).

That is all.