

Subject: Proposal to add kTongyongGuifanHanzibiao, kJoyoKanji, kJinmeiyōKanji, and kJinmeiyōKanjiVariant to the Unihan Database
Date: 2017-04-05
To: UTC
From: Jaemin Chung
Pages: 2

L2/17-088

I propose to add China's Tōngyòng Guīfàn Hànzìbiǎo (通用规范汉字表, general purpose normalized hanzi list, 2013), and Japan's Jōyō Kanji (常用漢字, everyday use kanji, 2010) and Jinmeiyō Kanji (人名用漢字, kanji for use in personal names) to the Unihan Database.

kTongyongGuifanHanzibiao

The values I propose to add to the kTongyongGuifanHanzibiao field are just natural numbers (1 to 8105). In the official Tōngyòng Guīfàn Hànzìbiǎo list (<http://www.gov.cn/gzdt/att/att/site1/20130819/tygfzhzb.pdf>), every character has a number assigned to it.

I used the following data to make the attached kTongyongGuifanHanzibiao text file.

- <http://blogs.adobe.com/CCJKType/files/2014/06/china-8105-06062014.txt>

kJoyoKanji

The values I propose to add to the kJoyoKanji field are natural numbers (1 to 2136, in the order appearing in the list) and the readings specified in the official Jōyō Kanji list (http://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kijun/naikaku/pdf/joyokanjihyo_20101130.pdf).

The number and the readings are separated by a semicolon (;). On'yomi (音読み; Sino-Japanese reading) is written in katakana, and kun'yomi (訓読み; native Japanese reading) is written in hiragana. Okurigana (送り仮名) in kun'yomi is separated by a hyphen-minus (-). Multiple readings of a single kanji are separated by a space character.

Kyūjitai (旧字体) characters are not included.

I used the following data to make the attached kJoyoKanji text file.

- <http://x0213.org/joyo-kanji-code/>
- <http://www.asahi-net.or.jp/~ax2s-kmtm/ref/jouyoukanji.html>
- <https://ja.wikipedia.org/wiki/%E5%B8%B8%E7%94%A8%E6%BC%A2%E5%AD%97%E4%B8%80%E8%A6%A7>

(Document continued on the next page)

kJinmeiyōKanji

The value I propose to add to the kJinmeiyōKanji field is simply “Yes” for every character appearing in the official Jinmeiyō Kanji list (<http://www.moj.go.jp/content/001131003.pdf>). Numbers and/or readings are not suitable because of the following reasons:

- The official Japanese Jinmeiyō Kanji list does not specify any readings to any kanji. (This is actually one of the reasons why an extremely unusual name such as Yuri Murayama (村山^{むらやま}彩希^{ゆいり}) exists – there is absolutely no way that the name 彩希 can be read Yuri (ゆいり) at all.)
- A new kanji is sometimes added in the middle of the list. If Jinmeiyō Kanji is maintained using a numbered list, new numbers have to be assigned whenever a new kanji is added to the list.
 - Jinmeiyō Kanji is revised much more frequently than Jōyō Kanji. When a new version of Jōyō Kanji is issued, it is usually not revised for multiple decades (1981, revised in 2010, then no change (until March 2017, as of this writing)). But Jinmeiyō Kanji is not that stable. After the 1981 version of Jōyō Kanji was issued, Jinmeiyō Kanji was revised in 1990, 1997, 2004, and 2009. After the 2010 version of Jōyō Kanji was issued, Jinmeiyō Kanji was revised in 2015.

I used the following data to make the attached kJinmeiyōKanji text file.

- <http://x0213.org/jinmeiyō-kanji-code/>
- <http://www.asahi-net.or.jp/~ax2s-kmtm/ref/jinmeikanji.html>
- <https://ja.wiktionary.org/wiki/Wiktionary:%E4%BA%BA%E5%90%8D%E7%94%A8%E6%BC%A2%E5%AD%97%E3%81%AE%E4%B8%80%E8%A6%A7>

kJinmeiyōKanjiVariant

The official Jinmeiyō Kanji list also contains some information about a variant of a kanji, with a hyphen or parentheses. The kJinmeiyōKanjiVariant field lists those variants.

I used the same data mentioned in the kJinmeiyōKanji section above to make the attached kJinmeiyōKanjiVariant text file.

(End of document)