

Title: Concerns on the copyright of the IDS data referred by L2/21-118R.  
Source: suzuki toshiya (individual contribution)  
Date: 2021/09/17  
Reference: [L2/21-118R](#)

L2/21-118R proposed to include the IDS data to Unihan DB with new property, kIDS, and refers the data at the babelstone.co.uk. Before all, I express my thanks and appreciate the contribution for the “not copyrighted” IDS data, which might be hoped by many Unicode experts, and the long & ongoing efforts to improve and maintain the data. But please let me note about a concern on its copyright status.

The IDS data used by IRG was maintained by Taichi Kawabata until Ext F, which was licensed under the GNU GPL license, because the IDS for URO, CJK Unified Ideograph Extension A and B were taken from the CHISE project (currently led by Tomohiko Morioka from Kyoto University, Japan) who licensed these data under the GNU GPL. I heard that many experts had tried to have an agreement with the CHISE project to use the data in the software with different licenses, but could not reach an agreement. Maybe the note of “not copyrighted” is added to resolve this situation. However, I have to note 2 concerns on the “not copyrighted” status of the data referred by the L2/21-118R.

1) an issue of the license.

the header of the IDS.TXT data referred by L2/21-118R has a mention to the GitHub repository by Taichi Kawabata. The repository does not have IDS.TXT but has [LISP software](#) including IDS data covering from URO to Ext F. The program in the repository is licensed under the GPL version 3. I think the derivative works of the GPLv3 software cannot be public domain until the day of their copyright is expired.

2) an issue of the authorship.

It is strongly supposed that the program in the Kawabata’s repository has some relationship with the IDS data from the CHISE. Another repository <https://github.com/cjkvi/cjkvi-ids/> has the ids.txt file by Kawabata-san, it is stated as derived from the CHISE project’s data (\*). The git repository (not on GitHub) of the CHISE project has more authors; <https://gitlab.chise.org/CHISE/ids/-/blob/master/README.en>

I hope these concerns are resolved before the official inclusion of kIDS into Unihan DB (and redistribute through Unicode website), if the kIDS would be added to the characters in URO, Ext A and Ext B. Anyway, the data for Ext C and later would not have such copyright issues, it would have no copyright issues.

(\*) Maybe I’m overlooking some bits of the links to cjkvi-ids or CHISE in the repository [github.com/kawabata/ids](https://github.com/kawabata/ids). I’m asking to Kawabata-san now. Several times I’ve discussed with Kawabata-san about the license of IDS data, but he never mentioned as if he made his own data independently.

(end of document)