

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

**Doc Type:** Ideographic Research Group Document  
**Title:** Feedback on IRGN2554  
**Source:** Eiso Chan (陈永聪)  
**Status:** Individual Contribution to IRG #59, online meeting  
**Action:** For consideration by IRG and TCA  
**Date:** 2022-10-17

In my personal opinion, we can consider this issue more easily. The followings are my points.

(1) We have confirmed that the ideographs only used for the person names are necessary for encoding in UCS and Unicode. I think it's a very important consensus for our encoding works.

(2) We need to confirm if the TCA submitted evidence is authoritative. If the registration process stated in IRGN2554 is in line with reality and local laws, the evidence submitted by TCA should be acceptable for our encoding works. The characters of a person's name can't be decided by himself/herself directly. The name holder (or his/her parents) can advocate a name, but whether this name can be registered successfully requires the approval of local governments. When his/her name is registered, he/she needs to use it in the daily life. If the names are input by the person doesn't match the data of the household system, the working personnels of the banks, schools, hospitals, railway stations, airports and so on could reject to serve them. Therefore, a person's name can also be regarded as a legally protected contract concluded between an individual citizen and the local governments, at least in Taiwan Province. The situation is similar in mainland China.

(3) For the TCA submitted characters only used for the person names, some are unifiable with the encoded characters, but some are not. If the character could be unified with encoded character, we can unify it, and the submitter can't reject the unification decision for the person's name as the reason; if the character could not be unified or we don't have the sufficient additional evidence or explanation to unify, it is better to encode it separately like other common submitted characters. In this process, more new UCV rules may be pointed out based on the quantity statistics and rationale analyses, the submitter should not oppose for the same reason. Some experts concerned how to confirm if the character is the variant of any encoded character. The ID cards used in mainland China and Taiwan Province are different from the ones used in HKSAR and Macao SAR. There is not a stable and formal Romanization string for a person's name printed on the ID card and recorded in the household database in mainland China and Taiwan Province. If a person doesn't have to go abroad and only stay in a very narrow place all his/her life, and he/she has no glorious deeds or criminal facts that need to be broadcast on TV and radio, he/she even can't pass the college entrance examination who doesn't need to write a paper signed with his/her English name, and he/she is a person of few

words in school, company or on cultivated field, there is no chance to set a formal Romanization string for his/her name and the name will not be called in a public environment, but he/she still has the right to use the legal real name which he/she (or his/her parents) thinks suits him/her, and his/her name should be interchanged in all aspects of his/her daily life like an ordinary person. In our knowledge, the family members and the friends of a person must know how to read the person's name, but they also have rights to keep silent for us; another situation is they are willing to tell us all information we want to know, but the pronunciations they can provide in some Chinese dialects even ethnic languages and dialects will be very hard to translate/transform to the possible corresponding modern Mandarin pronunciation, or we can't handle the pronunciation in different dialects and languages for a phonology that has not been fully studied, that means the pronunciation provided by them are useless for us. It is too harsh for the submitters to provide the pronunciations of all the submitted characters. But, the reviewers could claim to postpone the character for further additional evidence or explanation when a character is regarded as incomprehensible based on the rationale of the abstract shape. Note that the Romanization string in the ID card in HKSAR and Macao SAR is necessary even if the spelling rules are not consistent which are mentioned by Prof. Kataoka and Prof. Cheng in their papers; the Kanji/Hanja forms are optional for the person's civic information in Japan and ROK (even DPRK?), that we can meet so many people's names can't be transformed to Hanzi easily, and there is not the legal real name systems in Japan and ROK only based on Kanji/Hanja like mainland China and Taiwan Province, so the experiences of Japan and ROK are not totally suitable for mainland China and Taiwan Province.

(4) In Chinese traditional culture of some areas, the parents would select the servile names as the legal real names of their children, because they think the servile names would help them raise the children well. (贱名好养活。) On the other hand, some parents would use some Hanzi with the unfriendly meanings as the children's names, because they hate the children. People's activity motives are complex, so we need not be surprised by a character only used for person name but with the pejorative meaning.

(5) As TCA shows, the ID card used in Taiwan Province recorded not only the information of the holder himself/herself. The information of his/her parents and grandparents should be treated as the digital heritage of one person. Since there is such a requirement in the local law, we should not refuse to accept the character to IRG WS because the person died.

(6) I can understand all the experts' rigorous and serious comments on these characters, but UCS and Unicode are just a maximum character range for modern digitization not a dictionary, which is better to make all the users convenient first. I have a private project named Hierarchical Character Subset of Practical Contemporary Chinese Hanzi (当代实用汉字分层字符子集) to collect all the eligible Hanzi used in China (mainland, HK SAR, Macao SAR and Taiwan Province included) and overseas Chinese regions (especially in Singapore, Malaysia, Thailand, North Myanmar, Central Asia and so on), and I set many levels. If a character is introduced with clear rationale, meanings, pronunciations and specific use cases, it will be included in more forward level; if not, it will be included in more posterior level or excluded.

Therefore, I think it's OK to accept the TCA submitted characters to WS2021 as the common characters for the further reviews.

**(End of Document)**