

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:** Request IRG for consideration to accept the intralingual captions and the interlingual subtitles as evidence in future IRG WS  
**Source:** Eiso Chan (陈永聪, Culture and Art Publishing House)  
**Status:** Individual Contribution to IRG #61, Yale University  
**Action:** For consideration by IRG  
**Date:** 2023-10-02

In the newest version of IRG PnP, Section 2.2.1.d.(2).e) shows the following sentences.

<b>Usage (需求限制):</b> The use of characters must be for justifiable public interest. Examples of public use include evidence of: <i>governmental needs; scientific use; digitization projects for public use; and working systems of significance</i> as accepted by IRG. IRG has the right to reject characters that do not have sufficient evidence for IRG of justifiable public interest.
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Until now, we have not used the intralingual captions and the interlingual subtitles as evidence in IRG encoding works. However, I found some useful ideographs in our daily life and the academic research, which are used in the intralingual captions or the interlingual subtitles of the films/movies, the TV dramas, the documentaries, the video recordings of Chinese Xiqu Operas or Quyi Shows or traditional plays in East Asia or Southeast Asia, official music video of modern or pop songs and so on, but it is not easy to find them in any publications.

[W3C](#) has published some recommendations or candidate recommendations on the intralingual captions and the interlingual subtitles, like [TTML1](#) and [WebVTT](#), and they also kindly provide the clear [definitions](#) of the terms “caption” and “subtitle” as below. I added the adjectives before the terms based on the definitions.

<i>Captions</i> for the same language as the spoken audio.
------------------------------------------------------------

<i>Subtitles</i> for spoken audio translated into another language.
---------------------------------------------------------------------

As an example, U+2AA0A (𠄎) is a very common Hanzi in Chinese mainland, which it is not easy to find it in the official publications, but it is not hard to find it in the intralingual captions of some TV dramas.

The following picture shows U+2AA0A (𠄎) which is the screenshot cited from the Chinese TV drama [Zhuang Tai/Building the Stage](#) (《装台》) in 2020, which the director is Li Shaofei (李少飞), and starred by Zhang Jiayi (张嘉益), Yan Ni (闫妮) and so on. This TV drama is adapted from the novel with the same title, which is written by Chen Yan (陈彦) who is one famous writer and dramatist in Chinese mainland. This TV drama was previously broadcasted on [CCTV-1 \(中国中央电视台综合频道\)](#), and it is also transformed to a new play of Wuxi Xiju Opera (锡剧) with the same title by [the Wuxi Xiju Opera Troupe of Jiangsu Provincial Performing Arts Group \(江苏省演艺集团锡剧团\)](#) in 2022.



The above sentence is used for the theme song 《不愁》 sang by Sun Hao (孙浩). However, this character is still not easy to input via so many input methods now, so it was used as the homophonic replaced form as 忞 (U+6002) in the official music video shown as below.







When we use the intralingual captions and the interlingual subtitles as evidence, we still need to treat them as the common running texts and pay more attention to check if the sentences and the meanings are reasonable and right. When we cannot confirm if the character is right, it is better to check more sources or postpone for the future submission.

**(End of Document)**