

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N2351

Date: 2019-05-10

Source:	China
Meeting:	IRG#52, Hong Kong SAR, China
Title:	Activity Report
Status:	Member's submission
Actions required:	FYI
Distribution:	IRG
Medium:	Electronic
Pages:	2

1. Hosting IRG#53 confirmed

China confirms to host IRG#53 in 2019, in Shenzhen(深圳), Guangdong Province. See IRGN2366 for logistics information.

2. Progress of the Project “Chinese Characters Repertoire” (中華字庫)

2.1 Future Submissions

The Project has completed the digitalization of the Twenty-Four Dynasty Histories (廿四史) during 2017 to 2019. There are still some characters used for places and person names in these books found not encoded, thus which will be submitted to IRG in late 2019.

There are more than 50,000 head characters collected also, from more than 100 different kind of ancient dictionaries. 20% of them, i.e. 10,000 or so are picked out and will be submitted in batchs.

The Chinese Character Repertoire and China National Committee for Terms in Sciences and Technologies are collaboratively working with the Project, the findings will be released in next meeting.

2.2 Fonts

There are 750 thousand of 1 million Hanzi fonts made by the early 2019.

3. Methods for Chinese font glyph evaluation based on graphics and deep learning

The Peking University is designing a series of quantitative and qualitative glyph evaluation methods for Hanzi fonts. The proposed algorithm, computer graphics-based methods and deep learning-based methods were comprehensively compared with other publicly known methods to prove its effectiveness. A user study was also carried out to measure its realism and style-similarity qualitatively. Furthermore, texts rendered in the styles generated according to the method were illustrated to indicate its feasibility in real applications. The expected advantages in both visual perceptions and quantitative evaluations were confirmed.