

TO: UTC
FROM: Deborah Anderson, Head of Delegation, USNB
RE: WG2 Report
DATE: 30 October 2009

The WG2 meeting took place from 26 to 30 October and was hosted by the Information Processing Society of Japan/Information Technology Standards Commission of Japan in Tokyo. The facilities were ideal and the needs of the participants were promptly attended to by the Japanese National Body hosts, and in particular Toshiko Kimura.

1. Disposition of Comments on the Amendments

The meeting's main focus was to work through the ballot comments on three amendments and have them advance, if possible.

- FPDAM 7 presented only a couple of significant hurdles: Ireland wanted two graphic-like characters re-inserted into Batak and Japan requested a name change for the Historic Kana block. Both were resolved (the two Batak characters were added into AMD 8 instead of AMD 7) and a suitable name was found for the new block (Kana Supplement), so the amendment progressed to FDAM.
- PDAM 8 presented more obstacles, due to the proposed LATIN LETTER MIDDLE DOT and Emoji. Both were discussed in ad hocs (see below), and although the US's vote remained "no" because of the failure to reach consensus on the disposition for the MIDDLE DOT, the amendment progressed to FPDAM.
- The CD was perhaps the most difficult to wade through, due to lengthy comments regarding fonts and sources from Japan and Korea. The Japanese issues were resolved, but Korea's vote remained "no." Nevertheless, the CD passed, and the US comments were on the whole accommodated.

2. Scripts and Character Proposals

All of the script contributions that had been approved by the UTC and L2 were also approved by the WG2. New scripts were earmarked for "the bucket" (targeted to a future amendment that will not be started until the subsequent WG2 meeting), and contributions that requested only a few characters were approved for Amendment 8.

The Government of India representative, Swaran Lata, and invited expert Peri Bhaskararao presented in person the GoI proposals for Kashmiri, which reflected an urgent request from the user community. The Devanagari additions were accepted, and the GoI agreed to the two additions requested by SEI in the

Arabic block. Requests for more characters in the Arabic block for Kashmiri are expected in the future from India.

Time was short at the meeting, so several script proposals (Pahawh Hmong, Manichaean, Warang Citi, Hentaigana, etc.) were postponed until the next meeting, including the contentious Hungarian script contributions. A Miao expert was in attendance (Erich Fickle) and an ad hoc meeting was able to resolve major architectural issues, with input from USNB representatives. A new proposal reflected the agreed upon changes is expected at the February UTC/L2. No experts were present for other historic ideographic proposals (Jurchen, Nüshu) or for a proposal to encode Obsolete Simplified Chinese characters, so discussion on these proposals was postponed until the next meeting.

a. Emoji

A successful emoji ad hoc meeting was chaired by Peter Constable, who went through various outstanding issues contained in the ballot comments. Members present included national body representatives, attendees from Google, and Katsuhiro Ogata, an independent journalist who authored N3711 (=L2/09-370) on the emoticons. This document was presented to the ad hoc, and served as the basis of another informal ad hoc between Ogata-san, Michael Everson, and Japanese national body representatives on names and glyphs for the emoticons. Further discussion between these parties is expected in the next few months. (Note that characters were added beyond those contained in PDAM 8, in response to Irish and German national body ballot comments, and it is possible a few more will be requested in the future from Ireland.)

The outcome of the ad hoc was:

- Decisions were made on characters as requested in the Irish and German ballot comments, and are documented in the Consent Docket.
- A source file containing the mappings between the emoji characters in the UCS and the characters in the Japanese telecommunication companies.
- After a discussion of the different options on the regional indicators (the USNB request for 10 and the Irish/German proposal for 256 two-letter codes), the US proposed an alternative solution to encode 26 symbols for letters of the Latin alphabet, which would be used in sequences to represent the 10 regional indicators in the emoji set. The US solution was approved.
- No action was taken regarding the remaining 65 characters (corporate logos, etc.) in the Japanese carrier sets, as no urgent need for their use in interchange has been demonstrated.

b. LATIN LETTER MIDDLE DOT

The LATIN LETTER MIDDLE DOT issue remains unresolved, with no movement from either the US or Ireland and UK. The character remains in Amendment 8 and the issue will again be taken up at the next meeting.

3. OWG-Sort

The OWG-Sort PDAM 2 ballot will start with Unicode 5.2 and will update to Unicode 6.0 before it is finished. The Korean Annex text has been drafted by Ken Whistler, with comments to come from Prof. Kim.

4. Future Meetings

The next meeting will be held in San Jose from 19-23 April 2010, with the fall 2010 meeting to be held in Busan, Korea, from

Oct. 4-8 2010, and in Helsinki, Finland, from April 11-15. The U.S. is the tentative host for Fall 2011.