

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest *Roadmaps*.

A. Administrative

| | |
|--|---|
| 1. Title: | BENGALI VOWEL LETTER AW (U+0985 U+09D7) USED IN KOKBOROK |
| 2. Requester's name: | <i>Sanghmitra Sahu (sanghmitra.sahu@gmail.com)</i> |
| 3. Requester type (Member body/Liaison/Individual contribution): | <i>Individual</i> |
| 4. Submission date: | <i>13 December, 2014</i> |
| 5. Requester's reference (if applicable): | <i>Email dated 11 Dec 2014 of Rick McGowan (rick@unicode.org) to barun_sahu@yahoo.com</i> |
| 6. Choose one of the following: | |
| This is a complete proposal: | <i>complete</i> |
| (or) More information will be provided later: | <i>-</i> |

B. Technical – General

| | |
|---|--|
| 1. Choose one of the following: | |
| a. This proposal is for a new script (set of characters): | <i>No</i> |
| Proposed name of script: | <i>-</i> |
| b. The proposal is for addition of character(s) to an existing block: | <i>Yes</i> |
| Name of the existing block: | <i>BMP, Bengali (U+0980 to U+09FF)</i> |
| 2. Number of characters in proposal: | <i>one</i> |
| 3. Proposed category (select one from below - see section 2.2 of P&P document): | |
| A-Contemporary <input checked="" type="checkbox"/> B.1-Specialized (small collection) <input type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/> | |
| C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/> | |
| F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/> | |
| 4. Is a repertoire including character names provided? | <i>Yes</i> |
| a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? | <i>Y</i> |
| b. Are the character shapes attached in a legible form suitable for review? | <i>Yes</i> |
| 5. Fonts related: | |
| a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? | <i>Samples in the enclosure</i> |
| b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): | <i>Samples in the enclosure</i> |
| 6. References: | |
| a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? | <i>yes</i> |
| b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? | <i>Yes</i> |
| 7. Special encoding issues: | |
| Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? | <i>Yes</i> |
| | <i>Enclosed</i> |

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

| | | |
|--|---|----------------------|
| 1. Has this proposal for addition of character(s) been submitted before? If YES explain | - | No |
| 2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents: | <i>www.tritripura.in , www.ttaadc.gov.in, www.kohm.twipra.com/</i> <i>Kokborok books, writings in Bengali script</i> | yes |
| 3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: | <i>http://en.wikipedia.org/wiki/Kokborok</i> | Tripura etc |
| 4. The context of use for the proposed characters (type of use; common or rare) Reference: | <i>See enclosure</i> | common |
| 5. Are the proposed characters in current use by the user community? If YES, where? Reference: | <i>Tripura (India) and nearby areas (including Bangladesh)</i> | yes |
| 6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference: | <i>Enclosed</i> | Yes Yes |
| 7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? | | Yes |
| 8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference: | <i>Enclosed</i> | Sometimes yes Yes |
| 9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference: | <i>Enclosed</i> | Yes Yes |
| 10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference: | - | No - |
| 11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference: | - - - | No - - |
| 12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) | | no - |
| 13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference: | - | No - |

ENCLOSURE

KOKBOROK LANGUAGE: Kokborok is a language spoken by a number of tribes in Tripura (a state in India) and neighbouring areas (Assam, Mizoram and parts of Bangladesh). The language is spoken by Tripura, Reang/Bru, Jamatia, Rupini etc). It is part of Bodo group of languages.

OFFICIAL STATUS OF KOKBOROK LANGUAGE: Kokborok is the second official language of Tripura since 1979. There currently is a debate over giving the language recognition as a National language of India.

SCRIPT OF KOKBOROK: At present Kokborok is officially written in a modified Bengali script. The modifications are on the sound “wa” as in “water”, use of apostrophe to denote a light “a” sound at the end of the word (sometimes, despite the implicit/explicit hasant U+09CD), and a special vowel and vowel sign (“w”). It is also written in Latin script by some groups.

THE PRESENT PROPOSAL: “Bengali Au Length Mark” (U+09D7, ৌ) occurs very much in Kokborok writing in Bengali. There are some Kokborok words, where the full vowel form of this vowel sign is used. For example in the word ঔ (‘awng’ meaning to be). At present, ঔ is being written as অ and ৌ in computers. But it is not a standard practice, and may not be universally supported. In fact, many word-processors do not produce the visual form correctly, and the “Use Sequence Checking” in Microsoft Word even prevents typing the two characters. For example, ঔ (U+0986) cannot be written as অ (U+0985) and ৌ (U+09BE).

So, it is proposed that a character vowel letter “Bengali Letter Aw” be encoded, corresponding to the vowel sign U+09D7.

SIMILARITY WITH DEVANAGARI SCRIPT: In Devanagari script, there is look-wise similar vowel letter and vowel sign. U+094A is the vowel sign, and U+0912 is the corresponding letter.

On the lines of Devanagari script, the Bengali character names may be changed to “Bengali Letter Short O” and “Bengali Vowel Sign Short O”. Also, the proposed letter may be U+0992 so that Bengali and Devanagari letters are relatively in the same location.

The features on sorting, searching and indexing etc for the proposed Bengali character may be same as those of the corresponding Devanagari character.

SAMPLE KOKBOROK TEXT:

লুকু নাইকলনাই তেই কামি হামকীরাই হাসতে তংনক (সিপাৰ্ড), ত্ৰিপুৱা, আগৰতলা

ফৌঙজাকনাইৱগনি ঞ্লাইথাই তেই ঞ্লাইথাইয়া:-

- ১। সিরিং খে তন্দি,
- ২। আয়াঙ-আয়াঙ তা বেরাইদি,
- ৩। ক্লাস ঔঙফুরক মোবাইল সুইস অফ, সিরিং এৰা কৌলইনাইখে তন্দি,
- ৪। ক্লাস বিসিংগ' মোবাইল' কক তা সাদি,
- ৫। ক্লাস বিসিংগ' ককৰগ তা সালাইদি,
- ৬। কৌচাঙ-স্ৰাইস্ৰাইখে তন্দি, সিত্ৰা-মানীইখর বিসিং দাদি,
- ৭। জের-মান্ খুকতাই তা থিবিদি, সুজাকনাই বেসিন্' সিমিসে খুকতাই থিবিদি,
- ৮। ধুমা তা নৌঙদি,
- ৯। সিপাৰ্ডনি মানীই-খুনাই তা পিয়াকদি এৰা তাজাৱা-মাজাৱা তা ঞ্লাইদি,
- ১০। কুবুই ক্লাস আংমানি জরানি বিসিং সরকারনি তাংৱাইদা তৌয়ীই সালাই মানাই,
- ১১। সামুঙ-আংনাই-আংয়া সেজনানি, ককহাম সানানি এৰা ককথুম ৱীনানি নাঙনাই,
- ১২। ককতিসামা হাকৰ' নিনি ককতিসামুঙ এৰা ককয়াফাৰমুঙ সুইয়ীই ৱীদি।

Translation of the above text in English:

State Institute of Public Administration and Rural Development (SIPARD), Tripura, Agartala

Dos and don'ts for trainees:-

- 1. Maintain silence,*
- 2. No loitering,*
- 3. Keep mobile in switched off, silent or vibrate mode in classroom,*
- 4. No talking on mobile in classroom,*
- 5. No crosstalk in classroom,*
- 6. Maintain cleanliness; and use trashcan,*
- 7. No spitting; spit only in washroom basin,*
- 8. No smoking,*
- 9. No damage to or fiddling with SIPARD property,*
- 10. Free to discuss government policy within the classroom during actual session,*
- 11. Feedback is mandatory,*
- 12. Submit your written complaints and suggestions in Complaint & Suggestion Box.*