
ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646 / UNICODE

A. Administrative

1. Title: **Proposal for Encoding Siloti Nagri Script in the BMP**
2. Requesters' names: Peter Constable (SIL International); James Lloyd-Williams and Sue Lloyd-Williams (Sylheti Translation And Research, London); Advocate Shamsul Islam Chowdhury (Chairman, Sylot Academy, Sylhet, Bangladesh); Professor Asaddar Ali (Vice Chairman, Sylot Academy, Sylhet, Bangladesh); Mohammed Sadique (First Secretary, Embassy of Bangladesh, Stockholm, Sweden); Matiar Rahman Chowdhury (Chairman, Sylot Academy (UK and Europe), London).
3. Requester type: Expert contribution
4. Submission date: 2002-11-01
5. Requester's reference (if applicable): WG2 Nxxxx (= US L2/02-xxx)
- 6a. Completion: This is a complete proposal
- 6b. More information to be provided: See referenced document in item 5

B. Technical—General**1a. New script? Name?**

Yes. Siloti Nagri.

1b. Addition of characters to existing block? Name?

No.

2. Number of characters

46

3. Proposed category

Category A

4. Proposed level of implementation and rationale

Level 3. Siloti Nagri script contains combining diacritics (see supporting documentation attached).

5a. Character names included in proposal?

Yes.

5b. Character names in accordance with guidelines?

Yes.

5c. Character shapes attached in a reviewable form?

Yes.

6a. Who will provide computerized font?

Sue Lloyd-Williams (Sylheti Translation And Research)

6b. Font currently available?

Yes.

6c. Font format?

TrueType

7a. Are references (to other character sets, dictionaries, descriptive texts, etc) provided?

Yes. See supporting document (reference provided above).

7b. Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?

Yes. See supporting document (reference provided above).

8. Does the proposal address other aspects of character data processing?

Yes. See supporting document (reference provided above).

C. Technical—Justification

1. Has this proposal for addition of character(s) been submitted before?

No.

2. Contact with the user community?

Yes.

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Sizeable communities in Sylhet region of Bangladesh, in Calcutta, in England and elsewhere.

3b. Reference

See supporting document (reference provided above).

4a. The context of use for the proposed characters (type of use; common or rare)?

Publishing in the script was known to have been done by several presses in Bangladesh and Calcutta as recently as the 1970s. More recently, no metal-type facilities have been available, but development of digital type has resulted in renewed use and growing interest.

4b.Reference

See supporting document (reference provided above).

5. Are the proposed characters in current use by the user community?

Yes.

5b. Where?

In London and Birmingham, and in Sylhet, possibly elsewhere.

5c. Reference

See supporting document (reference provided above).

6. After giving due considerations to the principles in N 1352 must the proposed characters be entirely in the BMP?

Yes.

6b. Rationale

Living script.

6c. Reference

See supporting document (reference provided above).

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes. One possible exception is SILOTI NAGRI SIGN FUL (U+xx71 in the chart): a similar character can be found in Bengali poetic texts, and thus it may make sense to unify this character across scripts. If unification were desired, then this character could be added to the General Punctuation block, as it is a punctuation character.

8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

9a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

Siloti Nagri has similarities with other North Indic scripts, but is distinct; it is no more like any other Indic scripts already in the UCS than are any of them to one another. Thus, for purposes of this proposal, these characters cannot be considered similar to existing characters.

10a. Does the proposal include use of combining characters and/or use of composite sequences (see clause 4.11 and 4.13 in ISO/IEC 10646-1)?

Yes.

10b. If YES, is a rationale for such use provided?

Yes.

10c. Reference

See supporting document (reference provided above).

10d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

Yes.

10e. Reference

See supporting document (reference provided above).

11. 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).

Yes. (See supporting document—reference provided above.)

D. Proposed Characters

Characters are presented so as to align with ISCII ordering of corresponding characters from related scripts. We do not necessarily prefer this layout over a more compact layout that keeps all characters in collation order, however. This has been done mainly to facilitate comparison with other blocks in the UCS.

(The code chart is presented on a new page.)

Code Chart:

	xx0	xx1	xx2	xx3	xx4	xx5	xx6	xx7
0			ଟ	ନ				ୈ
1			ଢ		୨			*
2	ଂ		ଟ	ନ				ଂ
3		ନ						ଃ
4			ନ					ଃଃ
5		ନା	ନା					ଃଃ
6	ନ	ନା	ନା					SN IVO
7	ହ	ନ	ନା		େ			
8		ବ	ନ	ନା				
9	ଢ			ନା				
A		ନ	ନା					
B		ନା	ନା		ୋ			
C		ନ	ବ			୨		
D		ନା	ନ		ଂ			
E			ନା	ା				
F	ପେ	ହ		ୀ				

Character Names:

Various signs

U+xx00 (This position shall not be used)
U+xx01 (This position shall not be used)
U+xx02 SYLOTI NAGRI SIGN ANUSVARA
U+xx03 (This position shall not be used)
U+xx04 (This position shall not be used)

Independent vowels

U+xx05 (This position shall not be used)
U+xx06 SYLOTI NAGRI LETTER A
U+xx07 SYLOTI NAGRI LETTER I
U+xx08 (This position shall not be used)
U+xx09 SYLOTI NAGRI LETTER U
U+xx0A (This position shall not be used)
U+xx0B (This position shall not be used)
U+xx0C (This position shall not be used)
U+xx0D (This position shall not be used)
U+xx0E (This position shall not be used)
U+xx0F SYLOTI NAGRI LETTER E
U+xx10 (This position shall not be used)
U+xx11 (This position shall not be used)
U+xx12 (This position shall not be used)
U+xx13 SYLOTI NAGRI LETTER O
U+xx14 (This position shall not be used)

Consonants

U+xx15 SYLOTI NAGRI LETTER KO
U+xx16 SYLOTI NAGRI LETTER KHO
U+xx17 SYLOTI NAGRI LETTER GO
U+xx18 SYLOTI NAGRI LETTER GHO
U+xx19 (This position shall not be used)
U+xx1A SYLOTI NAGRI LETTER CO
U+xx1B SYLOTI NAGRI LETTER CHO
U+xx1C SYLOTI NAGRI LETTER JO
U+xx1D SYLOTI NAGRI LETTERJHO
U+xx1E (This position shall not be used)
U+xx1F SYLOTI NAGRI LETTER TTO
U+xx20 SYLOTI NAGRI LETTER TTHO
U+xx21 SYLOTI NAGRI LETTER DDO
U+xx22 SYLOTI NAGRI LETTER DDHO
U+xx23 (This position shall not be used)
U+xx24 SYLOTI NAGRI LETTER TO
U+xx25 SYLOTI NAGRI LETTER THO
U+xx26 SYLOTI NAGRI LETTER DO
U+xx27 SYLOTI NAGRI LETTER DHO
U+xx28 SYLOTI NAGRI LETTER NO
U+xx29 (This position shall not be used)
U+xx2A SYLOTI NAGRI LETTER PO
U+xx2B SYLOTI NAGRI LETTER PHO
U+xx2C SYLOTI NAGRI LETTER BO

U+xx2D SYLOTI NAGRI LETTER BHO
U+xx2E SYLOTI NAGRI LETTER MO
U+xx2F (This position shall not be used)
U+xx30 SYLOTI NAGRI LETTER RO
U+xx31 (This position shall not be used)
U+xx32 SYLOTI NAGRI LETTER LO
U+xx33 (This position shall not be used)
U+xx34 (This position shall not be used)
U+xx35 (This position shall not be used)
U+xx36 (This position shall not be used)
U+xx37 (This position shall not be used)
U+xx38 SYLOTI NAGRI LETTER SO
U+xx39 SYLOTI NAGRI LETTER HO
U+xx3A (This position shall not be used)
U+xx3B (This position shall not be used)
U+xx3C (This position shall not be used)
U+xx3D (This position shall not be used)

Dependent vowel signs

U+xx3E SYLOTI NAGRI VOWEL SIGN A
U+xx3F SYLOTI NAGRI VOWEL SIGN I
U+xx40 (This position shall not be used)
U+xx41 SYLOTI NAGRI VOWEL SIGN U
U+xx42 (This position shall not be used)
U+xx43 (This position shall not be used)
U+xx44 (This position shall not be used)
U+xx45 (This position shall not be used)
U+xx46 (This position shall not be used)
U+xx47 SYLOTI NAGRI VOWEL SIGN E
U+xx48 (This position shall not be used)
U+xx49 (This position shall not be used)
U+xx4A (This position shall not be used)
U+xx4B SYLOTI NAGRI VOWEL SIGN OO
U+xx4C (This position shall not be used)

Various signs

U+xx4D SYLOTI NAGRI SIGN HASANTA
U+xx4E (This position shall not be used)
U+xx4F (This position shall not be used)
U+xx50 (This position shall not be used)
U+xx51 (This position shall not be used)
U+xx52 (This position shall not be used)
U+xx53 (This position shall not be used)
U+xx54 (This position shall not be used)
U+xx55 (This position shall not be used)
U+xx56 (This position shall not be used)
U+xx57 (This position shall not be used)

Additional consonants

U+xx58 (This position shall not be used)
U+xx59 (This position shall not be used)

U+xx5A (This position shall not be used)
 U+xx5B (This position shall not be used)
 U+xx5C SYLOTI NAGRI LETTER RRO
 U+xx5D (This position shall not be used)
 U+xx5E (This position shall not be used)
 U+xx5F (This position shall not be used)
 U+xx60 (This position shall not be used)
 U+xx61 (This position shall not be used)
 U+xx62 (This position shall not be used)
 U+xx63 (This position shall not be used)
 U+xx64 (This position shall not be used)
 U+xx65 (This position shall not be used)
 U+xx66 (This position shall not be used)
 U+xx67 (This position shall not be used)
 U+xx68 (This position shall not be used)
 U+xx69 (This position shall not be used)
 U+xx6A (This position shall not be used)
 U+xx6B (This position shall not be used)
 U+xx6C (This position shall not be used)
 U+xx6D (This position shall not be used)
 U+xx6E (This position shall not be used)

U+xx6F (This position shall not be used)

Siloti Nagri-specific additions

U+xx70 SILOTI NAGRI SIGN DVISVARA
 U+xx71 SILOTI NAGRI SIGN FUL
 U+xx72 SILOTI NAGRI POETRY MARK 1
 U+xx73 SILOTI NAGRI POETRY MARK 2
 U+xx74 SILOTI NAGRI POETRY MARK 3
 U+xx75 SILOTI NAGRI POETRY MARK 4
 U+xx76 SILOTI NAGRI INHERENT
 VOWEL O
 U+xx77 (This position shall not be used)
 U+xx78 (This position shall not be used)
 U+xx79 (This position shall not be used)
 U+xx7A (This position shall not be used)
 U+xx7B (This position shall not be used)
 U+xx7C (This position shall not be used)
 U+xx7D (This position shall not be used)
 U+xx7E (This position shall not be used)
 U+xx7F (This position shall not be used)

Unicode character properties

With the exception of the characters listed below, we propose that all characters have a General Category property of “Lo”, a Canonical Combining Class property of 0, a Bidirectional Class property of “L”, a Mirrored property of “N”, an East Asian Width property of “N”, and a Line Breaking property of “AL”. In the information provided below, only those values that differ from the defaults just mentioned are listed.

Characters	Properties
U+xx02, U+xx47, U+xx4D, U+xx70	General Category = “Mn”, Canonical Combining Class = 230, Bidi Class = NSM, Line Breaking = “Cm”
U+xx3E, U+xx3F, U+xx4B	General Category = “Mc”, Canonical Combining Class = 0, Line Breaking = “Cm”
U+xx41	General Category = “Mn”, Canonical Combining Class = 220, Bidi Class = NSM, Line Breaking = “Cm”
U+xx71	General Category = “Po”, Line Breaking = “QU”
U+xx72..U+xx75	General Category = “Po”, Line Breaking = “AL”

The canonical combining classes assigned to some of these do not match typical practice for Indic scripts. We propose these classes as we suspect they may potentially be more useful in the future, and less likely to lead to problems in use or implementation. If it is considered important to match comparable class assignments for other Indic scripts, then all Syloti Nagri combining marks should be assigned to class 0 with the exception of U+xx4D, which should be assigned to class 9.

Additional notes

Although Syloti is an Indic script, we are not proposing a viramic processing model for it. Detailed reasoning is provided in supporting documentation. Rather, we propose that conjunct forms are controlled by means of ZERO WIDTH JOINER: presence of ZWJ has the effect of requesting a conjunct form (displayed if supported by the font), while absence of ZWJ entails that a conjunct cannot appear in that position.

The character SILOTI NAGRI INHERENT VOWEL O is intended for use in specialised processing implementations and usage scenarios, and would not necessarily be utilised in most common implementations.