



a) Existing Standards

ISCII (Indian Standard Code for Information Interchange)–IS13194:1991

With the advent of computerization considerable work has been undertaken to facilitate the use of Indian languages on computers. These activities were generally limited to specific languages and were independent exercises of various organizations Thus making data-interchange impossible. In such a scenario, it was important to have a common standard for coding Indian scripts.

ISCII is the character code for Indian languages which suggests the encoding for all Indian languages that originate from Brahmi script. In 1991, The Bureau of Indian Standards adopted the Indian Standard Code for Information Interchange. ISCII was evolved by a standardization committee under the DIT(erstwhile Department of Electronics) during 1986-88. The ISCII document is available as **IS13194:1991 from the BIS offices and from <http://www.bis.org.in/>**.

Unicode Version 3.0

The Unicode Consortium is responsible for defining the behavior and relationships between characters. Unicode has incorporated a number of scripts prevalent in the world. Indian scripts also form a part of Unicode standard now. The Unicode standard encodes Indian language characters in the same relative positions A0-F4 in ISCII-88 standard. This parallel code layout emphasizes the structural similarities of the Brahmi scripts and follows the intention of the standard to enable one to one mappings between analogous coding positions in different scripts in the family. The details of the Unicode standard are available at <http://www.unicode.org>.

Proposed changes in Unicode with respect to Indic Scripts

Some modifications are necessary to incorporate in the Unicode Standard for adequate representation of Indian Scripts. Department of Information Technology, Ministry of

Communications & IT, is the voting member of the Unicode Consortium. Department of Information Technology has finalized the proposed changes in the existing Unicode Standards in consultation with respective State Government, Indian IT Industry & Linguists. These have been published in TDIL quarterly Vishwabharat@tdil issue 4 (Devanagari based languages Sanskrit, Hindi, Marathi, Nepali, Konkani & Sindhi), issue 5 (Gujarati, Malayalam, Telugu, Gurmukhi, Oriya), issue 6 (Bangla, Manipuri & Assamese), and issue 7 (Tamil, Kannada, Urdu, Sindhi & Kashmiri).

b) Draft Standards

i) **INSROT** (Indian Script to Roman Transliteration) : INSROT is the standard that defines transliteration from Devanagari into Roman. In order to facilitate websearch, it is case-independent, published in VishwaBharat@tdil, July 2002, issue 6th.

ii) **INSFOC** (Indian Script Font Code) : INSFOC is essential for exchange of content. Indian scripts are not linearly juxtaposed as Roman. Shape of vowel / consonant changes in a syllable or conjunct. Hence vowel-consonant combinations are fairly large in number. So far industry consensus has been evolved for common INSFOC standard for Devanagari, Gujarati, Gurmukhi & Malayalam, published in VishwaBharat@tdil, July 2002, issue 6th.

iii) **Vedic Code Set** : Vedic Code Set proposes inclusion of extended Devanagari code set with reserved locations for speech sound variations. Pure consonants are represented. Syllable / conjuncts are generated based on the script transformation rules. There is no separate code representation for vowel-modifier or matra. It was published in VishwaBharat@tdil, Oct 2002, issue 7th.

The summary of the proposed modification in UNICODE standard for Indian scripts and digital copy of all the Newsletters covering details of all the Standards is available at <http://tdil.mit.gov.in/standard.htm>.