



ISO/IEC JTC 1/SC 2/WG 2 N 1839

ISO/IEC JTC 1/SC 2/WG 3 N 469

Date : 1998-09-28

**ISO/IEC JTC 1/SC 2/WG 3
7-bit and 8-bit codes and their extension
SECRETARIAT : ELOT**

DOC TYPE :	Other document
TITLE :	US Recommendations regarding procedures for character set registration
SOURCE :	NCITS/L2
PROJECT:	----
STATUS :	For the consideration of SC2
ACTION ID :	Forward to SC2
DUE DATE :	----
DISTRIBUTION :	Project editor P, O and L Members of ISO/IEC JTC 1/SC 2 WG Conveners, Secretariats WG 3 Members ISO/IEC JTC 1 Secretariat ISO/IEC ITTF
MEDIUM :	Open
NO OF PAGES :	4

Contact 1: Secretariat ISO/IEC JTC 1/SC 2/WG 3 ELOT Mrs K.Velli (acting)
Acharnon 313, 111 45 Kato Patissia, ATHENS – GREECE
Tel: +30 1 22 80 001 Fax : +30 1 22 86 219 E-mail : kkb@elot.gr

Contact 2 : Convenor ISO/IEC JTC 1/SC 2/WG 3 Mr E.Melagrakis
Acharnon 313, 111 45 Kato Patissia, ATHENS – GREECE
Tel: +30 1 22 80 001 Fax : +30 1 22 86 219 E-mail: eam@elot.gr

ISO
INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION

ISO/IEC JTC 1/SC 2/WG 2
Universal Multiple-Octet Coded Character Set
(U C S)

ISO/IEC JTC1/SC2/WG2 N 1839
L2/98-291

Title : US
Date : August 25, 1998 recommendations regarding procedures for character set registration
Source : NCITS/L2
Action : Forward to SC2

As NCITS/L2 considered the Applications for Registration in SC2 documents N 3113-15, N 3125-38, and N 3140-41, it became apparent that many of these applications had the same defects. The defects could be eliminated by defining the requirements for character set registration more precisely, and by strengthening the review process.

Recommendation 1 :

The application for registration must include an exact copy of the character set **as it was originally published**, which will be retained by the registration authority as its reference copy. A new version with redrawn glyphs, different names and/or missing text may not be substituted.

Rationale for Recommendation :

The registration agency must possess a **true and complete copy** of what is being registered, and maintain it for reference purposes. (The reference copy may be the actual publication, or a digitized version of it.)

The information available in a true and complete copy of the character set being registered is vital in determining ISO/IEC 10646 equivalents for its characters.

The redrawing of glyphs may introduce errors, and (worst case) create spurious characters.

Example: N 3126 includes a sideways Q identified as a signature mark (a feature of manuscript and early printed books). The image for this character is different from that in ISO 5426-2 : 1996, the standard being registered. The image in ISO 5426-2 : 1996 itself is different from that in the British Library character set from which the sideways Q came. (As the source, the British Library forms should be considered the authoritative image).

Renaming of characters with ISO/IEC 10646 names introduces two problems. Firstly, it introduces an implied mapping of characters in the set being registered to ISO/IEC 10646

characters. (Mapping to ISO/IEC 10646 characters should be an explicit process, as recommended below.)

Secondly, if characters are renamed, the name of a character **as published** is unavailable in the reference copy. In the case of ISO or national standards, the name **as published** is the name that was approved under ISO or NSB procedures. If a character name needs to be changed, this should be done by revising the standard, and not by authorized renaming of characters during registration.

Redrawn versions of four characters sets (N 3125, N 3127-29) caused unnecessary work for the SC2 Secretariat and national standards bodies. The character sets in question are already registered, and the official versions published by ISO include the assigned G0-G3 escape sequences. But because redrawn versions were submitted, and the clause dealing with escape sequences was omitted, the SC2 Secretariat had no way of knowing that the four character sets were already registered.

Recommendation 2 :

Mapping of characters in a particular character set to ISO/IEC 10646 characters should be an explicit process, and must be subject to review by qualified experts.

Rationale for Recommendation :

Our concern here is to establish correct and consistent correspondences between characters from source character sets (those being registered) and ISO/IEC 10646. If mapping is done by people who lack the appropriate expertise, the result can be mappings with erroneous and contentious content, as evidenced by many of the proposed registrations being reviewed.

Incorrect mappings can introduce havoc into the exchange of data.

Example: N 3134 gives ISO/IEC 10646 mappings for American punctuation marks which do not agree with the mappings authorized by WG2 and published in SC2/WG2 N 1616.

Incorrect mappings or insufficient knowledge about the characters being mapped can result in unnecessary proposals.

Example: SC2/WG2 N 1745 is a proposal for the addition of six characters that already exist in ISO/IEC 10646.

Many characters sets have character in common. When a character occurs in more than one character set, it is essential that its ISO/IEC 10646 mapping be the same.

Example: ISO 5426:1980 and ANSI/NISO Z39.47:1993 both contain the registered trademark sign ®, but the registration applications for these character sets (N 3125 and N 3138 respectively) have different mappings for this character, to U+2122 and to U+00AE (the latter is correct).

Recommendation 3:

To implement the preceding recommendations, NCITS/L2 recommends that SC2 re-activate/re-establish the REGISTRATION ADVISORY GROUP (RAG) charged with examining proposed registrations, reviewing and revising procedures for registration, advising submitters of registrations on improvements of submissions, etc. The RAG will operate under SC 2 charter, as an advisory group to the Registration Authority. The RAG will consist of experts nominated by SC2 member bodies and its liaisons.

Rationale for Recommendation.

There are formal registration procedures in place based on ISO 2375. However, these procedures were established before the publication of ISO/IEC 10646, so they do not address the issue of including mapping information in a registration.

ISO/IEC JTC 1/SC 2/WG 3 N 469

The revival of the RAG would ensure that the reference material obtained in the registration process serves the needs of the global computer industry as full as possible. The RAG will also ensure the quality and reliability of mappings established as part of the registration process.

NCITS/L2 and the Unicode urge SC2 to give this proposal immediate consideration, so that current flaws in the registration process may be remedied.