

PRACTICE OF THE REGISTRATION AUTHORITY FOR ISO/IEC 2375

2007-03-01

IPSJ/ITSCJ

**Information Processing Society of Japan
Information Technology Standards Commission of Japan**

Practice of the Registration Authority for ISO/IEC 2375

This document can be accessed at the following URLs:

<http://www.itscj.ipsj.or.jp/ISO-IR/practice/practice.html>

<http://www.itscj.ipsj.or.jp/ISO-IR/practice/practice.zip>

Contents

	Page
Foreword	iv
1. General	1
2. Basic principles.....	1
3. Terms	1
4. Application for registration	2
4.1 Components of an application.....	2
4.2 Cover page	3
4.3 Code table	4
4.4 List of character names	5
4.5 Mapping table	6
4.5.1 Header information	6
4.5.2 Mapping of individual characters	6
4.5.3 Supplementary information	7
5. Names of characters.....	7
6. Submission of an application.....	7
7. Access to the International Register.....	8
Annex A: Flow chart of registration process	9

Foreword

The International Register of Coded Character Sets to be used with Escape Sequences

The International Register contains coded character sets which have been registered in accordance with procedures specified in ISO/IEC 2375. Its purpose is to identify widely used coded character sets and associate with each a unique ISO/IEC 2022 escape sequence by means of which a character set can be designated conveniently.

The publication of this Register promotes compatibility in international information interchange and avoids duplication of effort in developing application-oriented coded character sets. Registration provides identification for a coded character set but implies nothing about its status; it may or may not be part of an international or a national standard, or of an application-oriented standard. However, when such a standard is issued subsequent to the registration of an escape sequence, it is appropriate to specify the escape sequence identifying the character set in the standard.

If it is desired to register a coded character set, submit an application as required in ISO/IEC 2375 and as specified in this document. Any coded character set can be a candidate for registration if it satisfies the formal requirements of ISO/IEC 2375. The characteristics of the coded character set will determine the type of escape sequence which can be allocated to it.

The Registration Authority

The registration procedure and the maintenance of the International Register are performed by an International Registration Authority. The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have designated IPSJ/ITSCJ, the Information Processing Society of JAPAN/Information Technology Standards Commission of Japan, as the Registration Authority for ISO/IEC 2375.

IPSJ/ITSCJ performs these duties as a free service to the international information interchange community. It advises applicants on the requirements to be met in applications, and circulates the applications in conformance with the registration procedure. It allocates the escape sequence and registers each coded character set with its specific escape sequence.

The Registration Authority is assisted in its work by an advisory committee of technical representatives from national standards bodies or liaison organizations. For further information about the composition and duties of the Registration Authority's Joint Advisory Committee, see clause 10 of ISO/IEC 2375.

1. General

This document presents the practice of the Registration Authority. It is intended as a guide for applicants. It describes the practical decisions that must be taken to provide a uniform presentation of the International Register.

2. Basic principles

The registration scheme of ISO/IEC 2375 has been designed to establish an International Register of Coded Character Sets, identified by specific ISO/IEC 2022 escape sequences. It is essential that the contents of the registered coded character sets be unambiguous and that differences between the registered coded character sets be easily understood. For these reasons the following principles have been established:

- uniqueness of designation (names),
- consistent style for the definition of control functions,
- coherent graphic layout for the tables,
- uniform typographical presentation of the registration documents.

While the Registration Authority implements these principles as far as feasible, sometimes special cases may require exceptional handling. Experience may also lead to some re-adjustment of the current practice.

The registration process of the Registration Authority is shown as a flow chart in Annex A.

3. Terms

3.1

Sponsoring Authority (SA)

A Sponsoring Authority is an organization that submits applications to the Registration Authority. For the purposes of ISO/IEC 2375, Sponsoring Authorities are limited to the following:

- any ISO or IEC technical committee or subcommittee
- any group within the ISO/IEC JTC 1/SC 2, appointed by the ISO/IEC JTC 1/SC 2 for purposes connected with code extension or the use of escape sequences
- any national body of ISO or IEC
- any organization having liaison status with ISO or IEC or with any of their technical committees or subcommittees

A Sponsoring Authority may, but need not, be the Owner of Origin and/or the Copyright Owner (defined below).

For further information regarding duties of the Sponsoring Authority (SA), see clause 9 of ISO/IEC 2375.

3.2

Owner of Origin

The Owner of Origin is the organization or individual responsible for the development of a coded character set.

Practice of the Registration Authority for ISO/IEC 2375

The Owner of Origin has ultimate authority over the content of its coded character sets.

A Sponsoring Authority may, but need not, be the Owner of Origin.

3.3

Copyright Owner

The Copyright Owner is the organization or individual holding the copyright for the publication that specifies a coded character set.

A Sponsoring Authority may, but need not, be the Copyright Owner.

3.4

The Registration Authority's Joint Advisory Committee (RA-JAC)

The Registration Authority's Joint Advisory Committee (RA-JAC) consists of a chair who is a technical representative of the Registration Authority and four other members who are technical representatives from national bodies of the ISO/IEC JTC 1/SC 2 or technical representatives from organizations with a liaison relationship to the ISO/IEC JTC 1/SC 2.

The RA-JAC advises the Registration Authority on technical matters. In particular, in the application process, the RA-JAC evaluates a mapping to ISO/IEC 10646 that is contained in an application in the following way:

- At the request of the Sponsoring Authority, the RA-JAC provides assistance in preparing a mapping to ISO/IEC 10646.
- The RA-JAC examines each application that contains a mapping to ISO/IEC 10646.
- The RA-JAC, in conjunction with the Sponsoring Authority, reviews comments on the mapping received from the members of the ISO/IEC JTC 1/SC 2.

4. Application for registration

4.1 Components of an application

An application for registration consists of the following four components:

- a cover page
- a code table
- a list or lists of character names in each of the languages used for character names in the coded character set submitted for registration
- a mapping table (optional)

Depending on the target to be registered, an application includes or excludes these components as follows.

Table 1. Components of an application

Type of coding system	Cover page	Code table	List of character names	Mapping table
An approved ISO or ISO/IEC coded character set standard	Required	Not applicable	Not applicable	Optional
Other than an approved ISO or ISO/IEC coded character set standard				
Graphic coded character set	Required	Required	Required	Optional
C0-control character set C1-control character set	Required	Required	Required	Optional
Single control function (ISO/IEC 2022, Fs escape sequence)	Required	Not applicable	Not applicable	Not applicable
Coding system not conformant with ISO/IEC 2022	Required	Required	Optional	Optional

Forms for these components are provided as Microsoft Word or ODF files in the attachments of this document.

4.2 Cover page

The form for a cover page is provided in the following attachment:

- Attachment A: Form for a cover page
["A-CoverPage.doc"](#)
["A-CoverPage.odt"](#)

The Sponsoring Authority shall specify all the following fields on the cover page form:

- **TYPE:** The type of coded character set registration.
 - graphic coded character set
 - 94-character graphic character set
 - 96-character graphic character set
 - multiple-octet graphic character set
Specify "(94ⁿ characters)" or "(96ⁿ characters)" (where n = 2, 3, ...).
 - control functions
 - C0-control character set
 - C1-control character set
 - single control function (ISO/IEC 2022, Fs escape sequence)
 - coding systems not conformant with ISO/IEC 2022
 - coding system that uses the standard return
 - coding system that does not use the standard return
- **NAME:** A short name for the coded character set.
 *This text shall be used as the title of each registration in the International Register.

Practice of the Registration Authority for ISO/IEC 2375

- **DESCRIPTION:** A short description. If any of the following conditions apply, their descriptions shall be included here in the application.
 - A mapping to ISO/IEC 10646 is included.
 - The coded character set is intended for use in combination with one or more other registered sets with the identification of the registered set or sets.
 - The coded character set is intentionally a subset or a superset of one or more other registered sets (If the coded character set is a part of one or more standards, reference(s) to the standard(s) shall be described either in the DESCRIPTION: field or in the ORIGIN: field.).
 - The registration is for a coding system not conformant with ISO/IEC 2022 and the ISO/IEC 2022 standard return (escape sequence ESC 2/5 4/0) applies.
 - Where a publicly available document describing the coding system not conformant with ISO/IEC 2022 can be obtained if the registration does not include the code table and list of character names.
 - The registration may be subject to future modification.
 - The registration is a revision to a previously registered standard. If so, then the identity of that coded character set shall be given.
- **SPONSOR:** The name of the Sponsoring Authority.
- **ORIGIN:** Source reference information of the coded character set.
- **OWNER OF ORIGIN:** The name of the Owner of Origin of the character or coded character set.
- **FIELD OF UTILIZATION:** A general indication of the intended field of application.

4.3 Code table

Forms for code tables are provided in the following attachments:

- Attachment B-1: Code table for a 94-character graphic character set
["B1-CodeTable94GCharSet.doc"](#)
["B1-CodeTable94GCharSet.odt"](#)
- Attachment B-2: Code table for a 96-character graphic character set
["B2-CodeTable96GCharSet.doc"](#)
["B2-CodeTable96GCharSet.odt"](#)
- Attachment B-3: Code table for a 2-octet (94²) graphic character set
["B3-CodeTable2Octet\(942\)GCharSet.doc"](#)
["B3-CodeTable2Octet\(942\)GCharSet.odt"](#)
- Attachment B-4: Code table for a C0 control character set
["B4-CodeTableC0CCharSet.doc"](#)
["B4-CodeTableC0CCharSet.odt"](#)
- Attachment B-5: Code table for a C1 control character set
["B5-CodeTableC1CCharSet.doc"](#)
["B5-CodeTableC1CCharSet.odt"](#)
- Attachment B-6: Example of a code table for an 8-bit graphic character set not conformant to ISO/IEC 2022

["B6-CodeTable8BitGCharSetNotConf2022.doc"](#)

["B6-CodeTable8BitGCharSetNotConf2022.odt"](#)

For graphic character sets, each code position of the code table shall display a graphic symbol of the character assigned to such a code position. If such a character has no graphic symbol, its acronym shall be displayed instead of a graphic symbol.

For control character sets, each code position of the code table shall display an acronym of the character assigned to such a code position.

If a code position of the code table is an "unused position", it shall be shaded.

4.4 List of character names

Forms for lists of character names are provided in the following attachments:

- Attachment C-1: List of character names for a 94-character graphic character set
["C1-ListCharName94GCharSet.doc"](#)
["C1-ListCharName94GCharSet.odt"](#)
- Attachment C-2: List of character names for a 96-character graphic character set
["C2-ListCharName96GCharSet.doc"](#)
["C2-ListCharName96GCharSet.odt"](#)
- Attachment C-3: List of character names for a 2-octet (94²) graphic character set
["C3-ListCharName2Octet\(942\)GCharSet.doc"](#)
["C3-ListCharName2Octet\(942\)GCharSet.odt"](#)
- Attachment C-4: List of character names for a C0/C1 control character set
["C4-ListCharNameC0C1CCharSet.doc"](#)
["C4-ListCharNameC0C1CCharSet.odt"](#)
- Attachment C-5: Example of a list of character names for an 8-bit graphic character set not conformant to ISO/IEC 2022
["C5-ListCharName8BitGCharSetNotConf2022.doc"](#)
["C5-ListCharName8BitGCharSetNotConf2022.odt"](#)

For all character sets with character names in more than one language, the application for registration shall include a list of character names in each language. Each list should use the form appropriate for the type of character set, with the title including the language name (ex. "List of character names in English" or "List of character names in French"). Each list should be provided in an individual file. The remaining rules in this clause apply to each language-specific list of character names.

For graphic character sets, the list of character names shall show all the code positions in the code table and indicate the name of the character allocated to each code position as the name appears in the coded character set to be registered. Combining characters shall be identified by adding the text "(combining character)" immediately following the character name.

For control character sets, the list of character names shall show all the control characters of the set by indicating the acronym, name and definition for each code position in the code table as the name appears in the coded character set to be registered.

An unused position shall be indicated by the text "(This position shall not be used)" instead of a character name. For a contiguous range of unused positions, the list may show the range of code positions as a single entry, where the code position shows the first code position in the range, the word

Practice of the Registration Authority for ISO/IEC 2375

“to”, and the last code position, and the text for the character name shall be “(These positions shall not be used)”.

To help in understanding a graphic character, a short note for a character may be included either in parentheses after a character’s name or in the “Note” column of the form.

4.5 Mapping table

A mapping table identifies ISO/IEC 10646 equivalents for the characters of the coded character set proposed for registration. This section covers the main features of a mapping table. The complete specifications are published in ISO/IEC 2375, clause A.2.

A mapping table shall be a text file, characters of which are limited to the repertoire for ISO/IEC 646 IRV and three control characters **HT (CHARACTER TABULATION)**, **CR (CARRIAGE RETURN)** and **LF (LINE FEED)**

A mapping table shall be submitted in machine readable form, either as an attachment to the application for registration or by reference to a downloadable file on the Internet. A printed copy that is identical to the content of the machine-readable form may also be included as part of the application, but only the machine readable copy is required.

A mapping table has three parts:

- header information to identify the mapping table (required)
- list showing each character and its ISO/IEC 10646 equivalent (required)
- supplementary information for clarification (optional)

4.5.1 Header information

A mapping table shall contain the following header information:

- the name of the coded character set mapped to ISO/IEC 10646
- a version number
- the date of creation
- general notes preceding the list of mapped characters
- the applicable edition of ISO/IEC 10646 plus any amendments and corrigenda on which the mapping is based
- a statement of the format of the table (character used for separation, encodings used for the characters, and the order of the characters)

4.5.2 Mapping of individual characters

The mapping for each character shall be in the following format:

- A mapping for a character is a single line of text.
- Each line of text in an individual character mapping shall contain the code position of a character of the registration coded character set on the left side of the line and the corresponding short identifier for the code position (UID) of ISO/IEC 10646 on the right side of the line. These two notations are separated by **HT**.

- The code position of a character of the registration coded character set shall be described with hexadecimal (base 16) notation (the digits “0” through “9”, and Latin letters “A” through “F” (or “a” through “f”)).
- When the character in the registration corresponds to a combining sequence of ISO/IEC 10646, its coded representation shall be described in the notation for UCS Sequence Identifier (USI) of ISO/IEC 10646, such as following format:
$$\langle UID_1, UID_2, \dots, UID_n \rangle$$
- If a mapping to ISO/IEC 10646 does not exist for a character in the registration, the word “none” shall be indicated on the right side of the line.
- If a corresponding ISO/IEC 10646 character is from private use area or planes, its code position shall be followed by the text “(Private use character)”.

The list of mappings shall be ordered by the code position of the character in the registration. The list of mappings shall be in the same order as the characters in the coded character set submitted for registration.

If the coded character set submitted for registration includes combining characters, the mapping table shall include a note saying whether combining character(s) precede or follow the base character on which the combining character(s) is/are to be positioned. This note shall follow the entire list of mappings.

4.5.3 Supplementary information

Supplementary information may be included following all other mapping table information. An example of supplementary information is a discussion of alternative choices for mapping of a character in the coded character set submitted for registration.

An example of a mapping table is provided in the following attachment:

- Attachment D: Example of a mapping table
[“D-ExampleOfAMappingTable.txt”](#)

5. Names of characters

The name of a character is strongly recommended to meet the requirements of clause 6.4 **Naming of characters** and Annex L (informative) **Character naming guidelines** of ISO/IEC 10646.

NOTE – The official French version of ISO/IEC 10646 presents different rules for the naming of characters in its clause 6.4 **Choix des noms des caractères** and Annexe L (informative) **Conseils pour le choix des noms de caractères**, and it shall also apply when French names are provided.

6. Submission of an application

An application shall be submitted by E-mail to the following destination address:

- IPSJ/ITSCJ
<mailto:iso-ir@itscj.ipsj.or.jp>

All the components of an application shall be electronic files; in particular the code table and the list of character names (if it contains graphic symbols) should be PDF files with embedded fonts.

7. Access to the International Register

The International Register can be accessed at the following URL:

- <http://www.itscj.ipsj.or.jp/ISO-IR/>

Annex A: Flow chart of registration process



