

INTERNATIONAL
STANDARD

ISO/IEC
30112

First edition
2020-09

**Information technology —
Specification methods for cultural
conventions**

*Technologies de l'information — Méthodes de spécification des
conventions culturelles*



Reference number
ISO/IEC 30112:2020(E)

© ISO/IEC 2020



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier; Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents	Page
Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
3.1 Bytes and characters	2
3.2 Cultural and other major concepts.....	2
3.3 FDCC-related categories	3
4 Notations	3
4.1 Notation for defining syntax	3
4.2 Portable character set.....	4
5 FDCC-set	6
5.1 General	6
5.2 FDCC-set description.....	7
5.2.1 General.....	7
5.2.2 Character representation.....	8
5.2.3 Continuation of lines.....	9
5.2.4 Names for copy keyword	9
5.2.5 Pre-category statements	9
5.3 LC_IDENTIFICATION.....	10
5.4 LC_CTYPE.....	12
5.4.1 General	12
5.4.2 Character classification keywords	13
5.4.3 Character string transliteration	17
5.4.4 "i18n" LC_CTYPE category.....	17
5.5 LC_COLLATE	42
5.5.1 General	42
5.5.2 Collation statements	44
5.5.3 "copy" keyword.....	46
5.5.4 "coll_weight_max" keyword.....	46
5.5.5 "section-symbol" keyword.....	47
5.5.6 "collating-element" keyword	47
5.5.7 "collating-symbol" keyword	47
5.5.8 "symbol-equivalence" keyword	48
5.5.9 "order_start" keyword	48
5.5.10 "order_end" keyword.....	49
5.5.11 "reorder-after" keyword	49
5.5.12 "reorder-end" keyword.....	50
5.5.13 "section" keyword	50
5.5.14 "reorder-section-after" keyword	51
5.6 LC_MONETARY	53
5.7 LC_NUMERIC.....	57
5.8 LC_TIME.....	58
5.8.1 General	58

5.8.2	Date field descriptors	62
5.8.3	Modified field descriptors	63
5.8.4	"i18n" LC_TIME category	64
5.9	LC_MESSAGES	65
5.10	LC_XLITERATE	65
5.10.1	General	65
5.10.2	Transliteration statements	66
5.10.3	"include" keyword	67
5.10.4	Example of use of transliteration	67
5.11	LC_NAME	68
5.12	LC_ADDRESS	69
5.13	LC_TELEPHONE	72
5.14	LC_PAPER	73
5.15	LC_MEASUREMENT	73
5.16	LC_KEYBOARD	74
6	CHARMAP	74
6.1	General	74
6.2	Character Set Description Text	74
7	Repertoiremap	79
8	Functionality	117
8.1	General	117
8.2	The "strpcoll" function	117
8.3	The "setmedia" function	118
8.4	String, encoding, repertoire and locale data types	118
8.4.1	General	118
8.4.2	String data type	118
8.4.3	Encoding data type	118
8.4.4	Repertoire data type	121
8.4.5	Locale data type	121
8.4.6	Character handling	123
8.4.7	String comparison	124
8.4.8	Message formatting	125
8.4.9	Conversion between string and other data types	127
8.4.10	Utilities	131
9	Messages format	133
Annex A	(informative) Differences from ISO/IEC/IEEE 9945	134
Annex B	(informative) Rationale	136
Annex C	(informative) BNF grammar	149
Annex D	(informative) Relation to taxonomy	155
Annex E	(informative) Implementation in glibc	158
Annex F	(informative) Relation between categories and keywords, and APIs	159
Annex G	(informative) Bindings guidelines	160

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see <http://patents.iec.ch>).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 35, *User interfaces*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document defines general mechanisms to specify cultural conventions. It also defines formats for a number of specific cultural conventions in the areas of character classification and conversion, sorting, number formatting, monetary formatting, date formatting, message display, addressing of persons, postal address formatting, and telephone number handling.

The benefits from this document are:

- | | |
|------------------------------|--|
| Rigid specification | Using this document, a user can rigidly specify a number of the cultural conventions that apply to their information technology environment. |
| Cultural adaptability | If an application has been designed and built in a culturally neutral manner, the application can use the specifications as data to its application programming interfaces (APIs), and thus the same application can accommodate different users in a culturally acceptable way to each of the users, without change of the binary application. |
| Productivity | This document specifies cultural conventions and how to specify data for them. With that data, an application developer is released from getting the different information to support all the cultural environments for the expected customers of the product. The application developer is assured of culturally correct behaviour as specified by the customer, and more markets can potentially be reached as customers can provide the data themselves for markets that were not targeted. |
| Uniform behaviour | When a number of applications share one cultural specification, which may be supplied from the user or provided by the application or operating system, their behaviour for cultural adaptation becomes uniform. |

The specification formats are independent of platforms and specific encoding and they are designed to be usable from a wide range of programming languages.

A number of cultural conventions, such as spelling, hyphenation rules and terminology, are not specifiable with this document, but the document provides mechanisms to define new categories and also new keywords within existing categories. An internationalized application can take advantage of information provided with the FDCC-set (such as the language) to provide further internationalized services to the user.

This document defines a format compatible with the one used in ISO/IEC 14651.

This document is upward compatible with elements of ISO/IEC/IEEE 9945, especially those on POSIX locales and charmaps – a locale or charmap conformant to POSIX specifications will also be conformant to specifications in this document, while the reverse condition will not hold. Some of the descriptions are intended to be coded in text files to be used via APIs developed for a number of systems which comply with ISO/IEC/IEEE 9945.

This document has enhanced functionality in a number of areas such as ISO/IEC 10646 support, more classification of characters, transliteration, dual (multi) currency support, enhanced date and time formatting, personal name writing, postal address formatting, telephone number handling, keyboard handling, and management of categories. There is enhanced support for character sets including ISO/IEC 2022 handling and an enhanced method to separate the specification of cultural conventions from an actual encoding via a description of the character repertoire employed. A standard set of values for all the categories has been defined covering the repertoire of ISO/IEC 10646.

This document has been developed to align with ISO/IEC/IEEE 9945. The major extensions from ISO/IEC/IEEE 9945 are listed in Annex A.

A rationale for elements of this document is found in Annex B.

A BNF specification of the syntax for formats in this document is given in Annex C.

The relation to the taxonomy of ISO/IEC TR 24785 is listed in Annex D.

A listing of the implementation of the specifications of this document in the GNU libc compiler product is given in Annex E.

The relation between formats and APIs of this document is listed in Annex F.

A guideline for a method to bind APIs of other programming languages to APIs defined in this document is specified in Annex G.

Information technology — Specification methods for cultural conventions

1 Scope

This document specifies description formats and functionality for the specification of cultural conventions, description formats for character sets, and description formats for binding character names to ISO/IEC 10646, as well as a set of default values for some of these items.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 639 (all parts), *Codes for the representation of names of languages*

ISO/IEC 2022, *Information technology — Character code structure and extension techniques*

ISO 3166 (all parts), *Codes for the representation of names of countries and their subdivisions*

ISO 4217, *Codes for the representation of currencies*

ISO 8601, *Date and time — Representations for information interchange*

ISO/IEC 9899, *Information technology — Programming languages — C*

ISO/IEC/IEEE 9945, *Information technology — Portable Operating System Interface (POSIX) Base Specifications, Issue 7*

ISO/IEC 10646, *Information technology — Universal Coded Character Set (UCS)*

ISO/IEC 14651, *Information technology — International string ordering and comparison — Method for comparing character strings and description of the common template tailorable ordering*

ISO/IEC 15897:2011, *Information technology — User interfaces — Procedures for the registration of cultural elements*

ISO 15924, *Information and documentation — Codes for the representation of names of scripts*