INTERNATIONAL STANDARD

ISO 639-4

First edition
2010-07-15

Codes for the representation of names of languages —

Part 4: General principles of coding of the representation of names of languages and related entities, and application guidelines

Codes pour la représentation des noms de langue —

Partie 4: Principes généraux pour le codage de la représentation des noms de langue et d'entités connexes, et lignes directrices pour la mise en œuvre

Reference number
ISO 639-4:2010(E)

© ISO 2010

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

This document is a preview generated by EVS

COPYRIGHT PROTECTED DOCUMENT

© ISO 2010

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

 © ISO 2010 – All rights reserved

Contents

Page

Foreword iv
Introduction v
1 Scope 1
2 Normative references 1
3 Terms and definitions 2
4 Fundamental concepts of language coding 5
4.1 Language identifiers and languages 5
4.2 Individual languages 6
4.3 Macrolanguages 6
4.4 Linguistic norm 7
4.5 Dialects 7
4.6 Collective language code elements and language groups 7
4.7 Extinct, ancient and historic languages 8
4.8 Artificial languages 8
4.9 Writing systems and scripts 8
5 Relationship between the parts of ISO 639 8
5.1 Parts of ISO 639 8
5.2 ISO 639 as one code space 8
5.3 Principles 9
5.4 Common maintenance and language coding database 9
6 Implementation issues 9
6.1 Applications 9
6.2 Subsets of the code table 10
6.3 Language groups 10
7 Combining language identifiers with other standards and codes 11
7.1 Combining ISO 639 with ISO 3166 11
7.2 Combining ISO 639 with ISO 19111 and ISO 19112 11
7.3 Combining ISO 639 with ISO 15924 12
7.4 Other code combinations 12
7.5 Formats of combined identifiers 12
8 Language description format (LDF) 13
8.1 Compatibilities between the ISO 639 model and ISO 12620 13
8.1.1 General 13
8.1.2 Identification 15
8.1.3 Description of an ISO 639 language identifier 16
8.2 Extensions to ISO 12620 for ISO 639 LDF 20
8.2.1 Representation 20
8.2.2 Documentation 21
8.3 Language information 23
Annex A (informative) **Overall steering of ISO 639** 26
Bibliography 28

This document is a preview generated by EVS

© ISO 2010 – All rights reserved

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 639-4 was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 2, *Terminographical and lexicographical working methods*.

ISO 639 consists of the following parts, under the general title *Codes for the representation of names of languages*:

— *Part 1: Alpha-2 code*

— *Part 2: Alpha-3 code*

— *Part 3: Alpha-3 code for comprehensive coverage of languages*

— *Part 4: General principles of coding of the representation of names of languages and related entities, and application guidelines*

— *Part 5: Alpha-3 code for language families and groups*

— *Part 6: Alpha-4 code for comprehensive coverage of language variants*

This document is a preview generated by EVS

 © ISO 2010 – All rights reserved

Introduction

ISO 639 provides codes for the identification and specification of individual languages, language variants, and language groups. The identifiers may be used in a variety of applications, including specification of the language used in a text, the language of terms or words in a dictionary or terminological database, the language used in a spoken presentation, language proficiency, language capabilities of software, localization, etc. The various parts of ISO 639 are expected to be implemented in a number of environments.

Parts 1, 2, 3, and 5 of ISO 639 all contain some information about implementation issues. However, it is deemed necessary to expand those descriptions, and to have the implementation rules in a separate document. In future revisions of the other parts of ISO 639, it is expected that those standards will reference this part of ISO 639 rather than duplicating the information.

This document is a preview generated by EVS

© ISO 2010 – All rights reserved

Codes for the representation of names of languages —

Part 4:
General principles of coding of the representation of names of languages and related entities, and application guidelines

This document is a preview generated by EVS

1 Scope

This part of ISO 639 gives the general principles of language coding using the codes that are specified in the other parts of ISO 639 and their combination with other codes. Furthermore, this part of ISO 639 lays down guidelines for the use of any combination of the parts of ISO 639.

The terminology and general descriptions of this part of ISO 639 are intended to replace corresponding text of other parts of ISO 639 as relevant in future revisions.

Relevant metadata for the description of linguistic entities are also given, as a framework for databases of linguistic data to support the ISO 639 series of International Standards.

2 Normative references

The following referenced documents are indispensable for the application 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 639-1:2002, *Codes for the representation of names of languages — Part 1: Alpha-2 code*

ISO 639-2:1998, *Codes for the representation of names of languages — Part 2: Alpha-3 code*

ISO 639-3:2007, *Codes for the representation of names of languages — Part 3: Alpha-3 code for comprehensive coverage of languages*

ISO 639-5:2008, *Codes for the representation of names of languages — Part 5: Alpha-3 code for language families and groups*

ISO 3166-1:2006, *Codes for the representation of names of countries and their subdivisions — Part 1: Country codes*

ISO 3166-2:2007, *Codes for the representation of names of countries and their subdivisions — Part 2: Country subdivision code*

ISO 3166-3:1999, *Codes for the representation of names of countries and their subdivisions — Part 3: Code for formerly used names of countries*

ISO 8601:2004, *Data elements and interchange formats — Information interchange — Representation of dates and times*

ISO/IEC 11179-1:2004, *Information technology — Metadata registries (MDR) — Part 1: Framework*

© ISO 2010 – All rights reserved

ISO/IEC 11179-2:2005, *Information technology — Metadata registries (MDR) — Part 2: Classification*

ISO/IEC 11179-3:2003, *Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes*

ISO/IEC 11179-4:2004, *Information technology — Metadata registries (MDR) — Part 4: Formulation of data definitions*

ISO/IEC 11179-5:2005, *Information technology — Metadata registries (MDR) — Part 5: Naming and identification principles*

ISO/IEC 11179-6:2005, *Information technology — Metadata registries (MDR) — Part 6: Registration*

ISO 12620:2009, *Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources*

ISO 15924:2004, *Information and documentation — Codes for the representation of names of scripts*

ISO 19111:2007, *Geographic information — Spatial referencing by coordinates*

ISO 19112:2003, *Geographic information — Spatial referencing by geographic identifiers*

This document is a preview generated by EVS

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

In future versions of other parts of ISO 639, it is expected that some or all of the terms and definitions will be replaced by a reference to the terms and definitions in this part of ISO 639.

NOTE The definitions in this part of ISO 639 are intended for practical use within the context of the various parts of ISO 639 and their applications. For various linguistic purposes, there are needs for more detailed, and possibly deviating, definitions.

3.1
code
data transformed or represented in different forms according to a pre-established set of rules

NOTE The usage of the term “code” is not uniform in all standardized coding systems. According to the usage that is defined in this part of ISO 639, a “code” is to be understood as a **code table** (3.2) and the set of rules relating to the code table. Each individual row in a code table is a **code element** (3.4) (e.g. “de - German - allemand - Deutsch” in Part 1 of ISO 639), while the item “de” is the **language identifier** (3.5).

3.2
code table
table of **code elements** (3.4) as part of a **code** (3.1)

3.3
code space
totality of possible values for a set of identifiers within a **code** (3.1)

EXAMPLE All sequences of two letters (a–z) form the code space of the alpha-2 language code as specified in part 1 of ISO 639.

NOTE The alpha-3 language codes that are specified in parts 2, 3, and 5 of ISO 639 share the same code space, i.e. no language identifier assigned in one of the parts may be assigned to a different item in another part.

© ISO 2010 – All rights reserved