INTERNATIONAL STANDARD

ISO 639-6

First edition
2009-12-01

Codes for the representation of names of languages —

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

Codes pour la représentation des noms de langues —

Partie 6: Code alpha-4 pour un traitement exhaustif des variantes linguistiques

Reference number
ISO 639-6:2009(E)

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

Contents

Page

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 639-6 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*

Introduction

Within language-dependent resources, where use, re-use and interchange are of significant importance to industry and academia alike, it is important to be able to fully document the data being captured within a resource. One important element of these resources is the language itself. It is important therefore to be able to identify the language of the resource as precisely and accurately as possible for the purposes of interoperability and quality of information content. The ISO 639 series of standards has been developed with these important goals in mind and this part of ISO 639 is concerned with capability for increased precision and accuracy. Specifically, this part of ISO 639 is concerned with the identification and documentation of language variants. Other parts of ISO 639 govern the identification of languages, language families and language groups and establish general principles for managing and developing codes.

The terms and definitions in this part of ISO 639 have, where appropriate, been harmonized with ISO 639-3.

This part of ISO 639 provides an alpha-4 language identifier and a unique language reference name forming the language code element for language variants. It also establishes a hierarchical framework that enables the relationship between language variants, language families, language groups and languages to be shown.

The alpha-4 language identifiers and language reference names, standardized in accordance with this part of ISO 639, are both complementary to, and compatible with, the alpha-2 and alpha-3 codes in other parts of the ISO 639 series of standards.

The ISO 639 series of standards provides a coherent means for the identification of languages from the highly generic to the highly specific. This part contributes to the objective of facilitating seamless transfer between these standards so that users are able to combine and use language identifiers for language variants, languages, macrolanguages, language groups and language families with minimal effort, whilst allowing a stand-alone system for future development and extensibility.

The alpha-4 language identifiers and the language reference names are provided for use in a broad range of applications, including terminology and lexicography, information and documentation (e.g. for education, archival and retrieval processes, information services, cultural heritage and “bridging the digital divide”), linguistics and information technology (search engines).

All language codes are to be regarded as open lists that can be extended and refined in accordance with the registration procedures within each standard. The registration and maintenance procedures for this part of ISO 639 have been adopted from ISO/IEC 11179-6.

Codes for the representation of names of languages —

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

1 Scope

This part of ISO 639 specifies a method for establishing four-letter language identifiers (alpha-4) and language reference names for language variants and a hierarchical framework for relating them to languages, language families and language groups. The alpha-4 language identifiers have been developed for use in a wide range of applications, especially in computer systems, where there is a potential need to cover the entire range of languages, language families and language groups as well as language variants within each identified language. Alpha-4 language identifiers can support the quantity of known language variants and accommodate any future expansion.

This part of ISO 639 provides a hierarchical framework, which facilitates backward compatibility with other ISO 639 codes, based on linguistic and/or geolinguistic relationships, within which a comprehensive enumeration of language variants is possible, including living, extinct, ancient and constructed languages, whether major or minor. As a result, this part of ISO 639 caters for a very large number of languages and their variants. This part of ISO 639 is not applicable to the registrations for languages designed exclusively for machine use, such as computer-programming languages and reconstructed languages.

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 639-4, *Codes for the representation of names of languages — General principles of coding of the representation of names of languages and related entities, and application guidelines*

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

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

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

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

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

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

3 Terms and definitions

For the purpose of this document, the terms and definitions in ISO 639-4 and the following apply.

3.1 Terms relating to alpha-4 language code

3.1.1
code
data transformed or represented in different forms according to a pre-established set of rules

[ISO 639-3:2007, definition 3.1]

3.1.2
code element
individual entry in a code table

[ISO 639-3:2007, definition 3.2]

3.1.3
language identifier
language symbol
unique string of three or four letters that represent a **language variant** (3.2.2)

NOTE In this part of ISO 639, three-letter language identifiers are those drawn from other parts of ISO 639. However, the four-letter language identifiers are of variant linguistic entities not found in any of the other parts of ISO 639.

3.1.4
reference name
name
appellation
linguistic expression used to designate an individual concept

NOTE 1 In this part of ISO 639, the reference name is used to designate a **language variant** (3.2.2).

NOTE 2 The reference name may be that by which the **language variant** (3.2.2) is known in any one of many languages.

NOTE 3 Adapted from ISO 639-3:2007, definition 3.4.

3.1.5
language code element
language (3.2.1) or **language variant** (3.2.2) entry in a code or code table consisting of categories of data that are transformed or represented in different forms according to rules

NOTE 1 The language code element for this part of ISO 639 consists of an alpha-4 **language identifier** (3.1.3) and a **reference name** (3.1.4) for **language variants** (3.2.2).

NOTE 2 The database for this part of ISO 639 includes the alpha-3 **language identifiers** (3.1.3) and their **reference names** (3.1.4); these are not part of this part of ISO 639 but provide information on the hierarchical links of **language variants** (3.2.2), **languages** (3.2.1), language families and language groups.

3.2 Terms relating to language and language variants

3.2.1
language
systematic use of sounds, characters, symbols or signs to express or communicate meaning or a message

3.2.2
language variant
uniquely identified use of **language** (3.2.1) based on **language variation** (3.2.3)

3.2.3
language variation
difference in the characteristics of **individual languages** (3.2.4)

3.2.4
individual language
language (3.2.1) that is distinguishable from other **languages** (3.2.1)

3.2.5
language documentation
information relating to the identification of a **language** (3.2.1)

4 Alpha-4 language identifier

4.1 Form of alpha-4 language identifier

A language identifier in accordance with this part of ISO 639 shall comprise a sequence of four letters from the following set of 26 letters of the Latin alphabet in lower case: a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z. Use shall not be made of diacritical marks or modified characters.

Alpha-4 language identifiers serve only as a device to uniquely represent language variants.

This part of ISO 639 makes provision for a reference name which is intended for use within a metadata registry. This reference name is selected from one of the names by which the language is or has been identified; this shall not be interpreted to imply that a name from any particular language is considered to be the preferred name.

As far as possible, an alpha-4 language identifier begins with the same letter as the corresponding reference name and includes one or more of the subsequent letters of that reference name; an effort being made to avoid pronounceable sequences. With thousands of languages and language variants and many similar names, alpha-4 language identifiers that resemble the reference name cannot be provided in every case. No alpha-4 language identifier shall have the same form as any other four-letter language reference name.

To maintain continuity and stability, alpha-4 language identifiers shall not change or be reused for a different purpose. In accordance with ISO/IEC 11179-6, when a language identifier is *retired* or *superseded*, the alpha-4 language identifier shall remain within the registry for this part of ISO 639 for backward compatibility.

When translating this International Standard to languages that are written using non-Latin scripts, alpha-4 language identifiers shall be formed using the Latin alphabet according to the principles of this part of ISO 639. The alpha-4 language identifier, and the language reference name, shall be treated as language-independent.

Knowledge of the world's languages and language variants at any given time is never complete or perfect. Additional language reference names and their alpha-4 language identifiers can be added when a language variant is found which differs from the language variants previously identified. It is essential that certain criteria be met before a new language code element can be accepted. These conditions conform to ISO/IEC 11179 (all parts) and are documented in 6.2. The denotation of an existing language code element can be revised. Existing reference names can also be revised. In addition, existing language code elements can be retired or become superseded when it is apparent that they no longer reflect language variant

distinctions in current use. When making changes, remaining language code elements shall not be adversely affected.

4.2 Syntax of the alpha-4 language identifier

Different types of language variants may be represented by an alpha-4 language identifier (e.g. living languages, ancient languages, artificially constructed languages) in a variety of communication modes including written, spoken or signed modes. The list of communication modes is not exhaustive and can be augmented over time. Table A.1 gives more detail.

Alpha-4 language code elements are linked to code elements provided by other standards in the ISO 639 series.

4.3 Use of language codes in the ISO 639 series of standards

The ISO 639 series of standards is intended to provide a compatible and interoperable set of language identifiers for language variants, languages and language groups and families. All parts of the series shall be considered when determining the most applicable language code, and the language identifiers may be applied, where applicable, in the order alpha-2, alpha-3, or alpha-4 moving from more generic to more specific forms of identification. The links provided between the alpha-4 and alpha-3 and between the alpha-3 and alpha-2 representations facilitate interoperability between systems, although each system can be used alone. No alpha-4 language identifier shall be assigned to language code elements that already have alpha-3 representations assigned to them in other parts of ISO 639. The Registration Authority of this part of ISO 639 is responsible, through representation on the ISO 639 Joint Advisory Committee (JAC), for determining the appropriate assignment of a language code element where there might be debate regarding the nature of the language or language variant being registered.

5 Language and language variants

5.1 Criteria for identifying language variants

No method of identification of a language is agreed by all or is appropriate for all purposes. As a result, there can be disagreement, even among speakers or linguistic experts, as to whether two language variants represent dialects of a single language or two distinct languages. For this part of ISO 639, judgements regarding whether two language variants are considered to be the same or different languages are based on a number of factors including linguistic similarity, intelligibility, a common literature and the views of speakers concerning the relationship between language and identity. They can also be affected by current views and local attitudes concerning the relationship between languages and ethnic identity and/or between languages and speakers' religious affiliations. Such judgements cannot always be absolute, given the impossibility of precisely defining degrees of linguistic similarity or inter-intelligibility and given the pressures of changing political situations.

The following basic criteria should be followed.

— Two related language variants are normally considered language variants of the same language if speakers of each language variant have inherent understanding of the other language variant (i.e. understanding based on knowledge of their own language variant without needing to learn the other language variant) at a functional level.

— Where spoken intelligibility between language variants is marginal, the existence of a common literature or of a common ethnolinguistic identity with a central language variant that both understand can be strong indicators that they should nevertheless be considered language variants of the same language.

— Where there is enough intelligibility between language variants to enable communication, the existence of well-established distinct ethnolinguistic identities can be a strong indicator that they should nevertheless be considered to be different languages. Some of the distinctions made on this basis might not be considered appropriate by some users or for certain applications.

These criteria shall be evaluated, where possible, according to open mediated discussion with identifiable experts; the results of discussion shall be documented.

5.2 Identifying languages

For the purposes of international reference, a category of identified languages (or language units), as recognized within the corpus of ISO 639 codes, shall be identified and represented as languages within this part of ISO 639.

Any unresolved cases of identification or re-identification will be referred to the Joint Advisory Committee of ISO 639.

5.3 Identifying spoken language variants

The application of alpha-4 language identifiers to spoken language variants within each language can be as detailed as practicable, involving not only language variants (known popularly as “dialects”) but also components (or “sub-dialects”) within each language variant.

The identification of language variants within a language is even more problematic than the identification of a “language” itself. For practical purposes, it is assumed that boundaries exist between neighbouring language variants or components of a spoken language, although there are most frequently gradual transitions among them (of pronunciation and/or vocabulary and/or morphology). The most tangible boundaries among the language variants and components of individual spoken languages are either geographic or ethnic, marked in particular by intervening highlands or open water or other areas of low or non-existent population or by areas where other languages predominate (although the relative dimensions of height, distance and/or population can vary greatly).

5.4 Identifying written language variants

The alpha-4 language identifiers of this part of ISO 639, where applied to written language variants, shall also denote the writing system, script and character set (where known) for use in identifying written and historical language variants and orthographies.

NOTE Whilst written language variants in scripts are already included within the code specified in this part of ISO 639, the framework also facilitates further extension to include transliteration and text to audio/audio to text representations.

5.5 Identifying transcription

The alpha-4 language identifiers of this part of ISO 639, where applied to code elements for transcription, denote a transcription of a spoken language or language variant. System designers may find it useful to further define these code elements by assigning an extension code from ISO 3166-1 or ISO 3166-2.

6 Structure

6.1 Model

6.1.1 General

The model for the code of this part of ISO 639 has been developed to be compatible with models being developed in ISO/TC 37 in general. ISO/TC 37 standards for computational use of terminology, specifically ISO 16642 and its combination with ISO 12620, emphasize the use of a metamodel in combination with metadata identifiers, referred to as data categories. This part of ISO 639 provides a specific model for language documentation and a list of metadata identifiers used within this model. These metadata identifiers are described and documented within ISO 639-4. Further discussion regarding this model can be found in References [15] and [16].

This model has also been developed in conformity with ISO/IEC 11179 (all parts) for the provision of a language metadata registry as follows:

— specified in accordance with ISO/IEC 11179-3;

— defined in accordance with ISO/IEC 11179-4;

— named in accordance with ISO/IEC 11179-5;

— registered in accordance with ISO/IEC 11179-6;

and is intended to be fully compatible with the metadata registry specified in accordance with ISO 12620.

The identifiers and associated data shall be managed within a metadata registry in conformity with ISO/IEC 11179-6.

6.1.2 Identification

Each language variant shall be provided with one language identifier. This language identifier is intended for use as the unique data identifier (DI), to be used in combination with a version identifier (VI) and registration authority identifier (RAI) for composing the international registration data identifier (IRDI).

6.1.3 Naming

The concepts of “language” and “language variant” are used in the model in a direct relationship with the metamodel. Each language code element shall have the language reference names for the language variant organized according to language sections. One or more names can be present in one or more languages that identify the language variant. Since the model is hierarchical, every language code element includes a link between the language variant and its parent. These links also provide a bridge between this part of ISO 639 and other codes in the ISO 639 series.

6.1.4 Representation

The alpha-4 language identifiers shall be regarded as representations in conformity with ISO/IEC 11179 (all parts).

6.1.5 Language documentation

Language documentation shall be maintained in the registry in conformity with ISO/IEC 11179 (all parts).

6.2 Data categories used in this model

Table 1 comprises descriptors used for documenting and identifying each language variant, some of which are deemed essential for purposes of registration.

A metadata registry for this part of ISO 639 shall be developed and maintained. The purpose of this metadata registry is to support the ability to identify whether similar suitable items of metadata already exist and, if not, to assist in the construction of a new description from similar existing language descriptions. The metadata registry will allow similarities between registered items to be identified as a function of the identifiers used for that item. Where similar or identical items are determined, appropriate decisions regarding the need for new items or the harmonization of existing items shall be undertaken, with the decisions being fully documented within the system and the status of the alpha-4 language identifiers being recorded accordingly.

These data categories are intended for capturing information that assists in the identification, the determination of the provenance and monitoring of the quality of the registered language variants. They shall be used to help avoid unnecessary variations when describing highly similar objects within the registry. The development of the metadata registry might result in the addition of further data categories, and the registration authority for this part of ISO 639 should be prepared to document such additions where it is

essential for interoperability within the ISO 639 series and for incorporation into ISO 12620 data category registry.

Table 1 shows the form to use to collect information from prospective applicants for new variants. Required information is denoted by the footnote reference letter “a”.

Table 1 — Descriptors

ISO 639-6 request for new variant	
1. Language	
reference name[a]	
parent[a]	
language family[a]	
mode of communication[a]	
status [standard or historic or extinct ...][a]	
level type [e.g. variety, component]	
superseded by	
2. Language naming	
language	
3. Name	
language name[a]	
other known names[a]	
4. Representation	
alpha-4 language identifier	
5. Geographical information	
toponym[a]	
UN region/ISO 3166 code element	
6. Linguistic information	
writing system	
script (including ISO 15924 code elements where available)	
bilingual	
language family	
7. Temporal information	
historical note	
modern events and changes	
date[a]	
8. Diachronic information	
historic class[a]	
historical classification	

Table 1 (*continued*)

ISO 639-6 request for new variant	
9. Registration information	
registration status	
information source[a]	
information last updated	
representation deprecated	
deprecated to	
maintenance notes	
10. Cultural/religious information	
community	
religious culture	
11. Social information	
population size	
social status	
legal status	
speaker identification	
migration	
census data	
age information	
health information	
supporting information[a] (to include closest linguistic affinity if applicable)	
[a] Required information.	

7 Extension coding and register

A key feature of the hierarchical system in this part of ISO 639 is found within its flexibility to be used as either an object-oriented stand-alone code or as part of an extension coding system.

This part of ISO 639 is mapped to other standards in the ISO 639 series in order to facilitate extension coding and backward compatibility.

When using data derived from this part of ISO 639 as part of an extension coding system, the syntax of the system, and thus the concept and model, shall be described in order that interoperability and compatibility with other systems can be maintained.

Coding systems from any number of ISO International Standards or other standards may be used as part of any extension coding system devised, but it is recommended, when using this part of ISO 639 as part of such a system, that the structure of the system be described in an open and unambiguous way in order to facilitate interoperability with as many applications as possible; application designers are encouraged to give consideration to backward compatibility and forward extensibility. Such extension coding systems can be designed for use in coding locales.

Other ISO standards can be combined with the hierarchical code of this part of ISO 639 to create extension coding systems. These may include:

— ISO 3166-1;

— ISO 3166-2;

— ISO 4217;

— ISO 8601;

— ISO 15924.

Care should be taken in creating extension coding systems to ensure that coding methods do not give rise to duplicate extension codes for the same linguistic entity.

The variants of this part of ISO 639 include sociolects and orthographic variants which are the most suitable aspects to be used as part of an extension coding system. To this end, this part of ISO 639 will provide a “flag” field within the published online data tables that will indicate those entities of this part of ISO 639 that are most suitable for use in an extension coding system that combines other standards of the ISO 639 series, ISO 15924 and ISO 3166-1 or UN.M49 region codes or, indeed, any other extension coding systems that may be designed in the future.

One such industry-standardized extension coding system already in use is the IETF (Internet Engineering Task Force) BCP 47. At the time of drafting this part of ISO 639, BCP 47 does not facilitate ISO 639-3, ISO 639-5 or this part of ISO 639 but provision for these standards is expected, although not guaranteed, in future updates of BCP 47.

The alpha-4 language identifiers of this part of ISO 639 can be further enhanced by attaching alpha-4 language identifiers that denote the linguistic register as part of an extension coding system.

The concept of linguistic register has been described as follows (taken from Reference [17]):

Linguistic varieties that are linked ... to occupations, professions or topics have been termed registers. The register of law, for example, is different from the register of medicine, which in turn is different from the language of engineering... and so on. Registers are usually characterized solely by vocabulary differences; either by the use of particular words, or by the use of words in a particular sense.

For example, alpha-4 language identifiers and language reference names could be created for the registers of “mechanical engineering”, “medicine” and “agriculture” as children within a language itself or the spoken or written variations within that language if there is a need to be more specific.

This part of ISO 639 offers a registration facility for extension codes to denote the linguistic register that can include variations across languages, professions and countries.

8 Administration of code assignments

The allocation of alpha-4 identifiers shall be administered by the Registration Authority (RA) of this part of ISO 639. The mandate and other responsibilities of this RA are outlined in Annex B.

Annex A
(informative)

Example data

Table A.1 provides an example of content forming the alpha-4 code of language reference names, their alpha-4 language identifier and immediate parent language identifier for this part of ISO 639. As no alpha-4 language identifier is assigned to a language variant that already has an alpha-3 representation, alpha-3 representations are included from the preceding standards where appropriate.

Table A.1 — Nilo-Saharan (part) Saharan (part — example data only)

Alpha-4 identifier	Alpha-4 parent identifier	Language reference name
scpz	nlsn	Saharan
wsnc	scpz	Western Saharan
kau	wsnc	Kanuri
bms	kau	Bilma Kanuri
bmss	bms	Bilma Kanuri spoken
fchi	bmss	Fachi
blma	bmss	Bilma
krt	kau	Tumari Kanuri
krtw	krt	Tumari Kanuri written
krta	krtw	Tumari Kanuri written Ajami script
krts	krt	Tumari Kanuri spoken
kby	kau	Manga Kanuri
kbyw	kby	Manga Kanuri written
kbya	kbyw	Manga Kanuri written Arab script
kbys	kby	Manga Kanuri spoken
mngn	kbys	Manga-N
mngs	kbys	Manga-S
knc	kau	Central Kanuri
kncw	knc	Central Kanuri written
knca	kncw	Central Kanuri written Arab script
kncs	knc	Central Kanuri spoken
dxgr	kncs	Dagara
njsk	kncs	Njesko
karb	kncs	Kabari
sgrt	kncs	Sugurti
lxre	kncs	Lare
ngzr	kncs	Ngazar
gvja	kncs	Guvja

Table A.1 (*continued*)

Alpha-4 identifier	Alpha-4 parent identifier	Language reference name
kgma	kncs	Kagama
fdwa	kncs	Fadawa
mdug	kncs	Maiduguri
mxao	kncs	Mao
kwyy	kau	Kwayyamo
kwys	kwys	Kwayyamo spoken
mvar	kau	Mavar
mvrs	mvar	Mavar spoken
kbl	kau	Kanembu
kblw	kbl	Kanembu written
kbll	kblw	Kanembu written Latin script
kbla	kblw	Kanembu written Arabic script
kbls	kbl	Kanembu spoken
kkwu	kbls	Karkawu
mndo	kbls	Mando
ngri	kbls	Nguri
hddd	kau	Haddad
hdds	hddd	Haddad spoken
txbu	scpz	Tubu cluster
tuq	txbu	Tedaga
tugw	tuq	Tedaga written
tuga	tdgw	Tedaga written Arab script
tugl	tdgw	Tedaga written Latin script
tugs	tuq	Tedaga spoken
gnda	tugs	Gunda
brwa	tugs	Brawia
chga	tugs	Chigaa
tmgr	tugs	Tomagra

Annex B
(normative)

Operation of the Registration Authority (ISO 639-6/RA) and the Registration Authorities Advisory Committee for ISO 639 (ISO 639/RA-JAC)

B.1 Mandates

B.1.1 ISO/IEC 11179 conformity

The alpha-4 language identifiers shall be considered as representations complying with ISO/IEC 11179 (all parts). The data associated with these representations and their language reference names shall be in conformity with ISO/IEC 11179 (all parts). The representations, identifiers and associated data shall be managed within a metadata registry in conformity with ISO/IEC 11179-6. This should ensure conformity with ISO 12620.

Any alpha-4 language identifier shall remain unchanged but may be deprecated or retired, even if a language variant represented by the alpha-4 language identifier is subsequently re-defined as a *language variant* or as a language group; the original alpha-4 language identifier shall remain in the registry for this part of ISO 639 for backward compatibility.

B.1.2 Registration Authority ISO 639-6/RA

The ISO 639-6/RA shall be responsible for the registration and maintenance of language reference names and alpha-4 language identifiers in accordance with the rules and processes specified in this annex.

A Registration Authority (RA) for this part of ISO 639 shall be established with the following mandate:

— to register language variants in accordance with this part of ISO 639;

— to act in conformity with ISO/IEC 11179-6 with respect to its registration procedures;

— to maintain a process that ensures the integrity of its registrations consistent with this annex;

— to establish and maintain a process that involves the Joint Advisory Committee (JAC) in any application for the registration of a major language or macrolanguage.

The Technical Management Board has appointed GeoLang Ltd[1)] to be the ISO 639-6 Registration Authority for the implementation of this part of ISO 639.

---

1) At the time of publication of this part of ISO 639, the ISO 639-6 Registration Authority is held by:

GeoLang Ltd
Corner House, Barn Street, Haverfordwest
Pembrokeshire SA61 1BW, Wales, United Kingdom.
The URL of the website is: http://www.geolang.com.

ISO maintains an on-line listing of maintenance agencies and registration authorities at http://www.iso.org/rama. Users are encouraged to consult this listing for the most up-to-date information concerning maintenance agencies and registration authorities.

B.1.3 Joint Advisory Committee ISO 639/RA-JAC

The Joint Advisory Committee (ISO 639/RA-JAC) shall advise and monitor the ISO 639-6/RA to ensure that the language code elements are developed in conformity with this part of ISO 639.

When the ISO 639-6/RA consults the ISO 639/RA-JAC about a proposed inclusion, deletion or change, the ISO 639/RA-JAC shall be obliged to respond within one month.

The ISO 639/RA-JAC shall resolve any questions that arise in relation to ambiguity in the relationships between denotations of language code elements from different parts of ISO 639.

B.1.4 Relationship between ISO 639-6/RA and ISO 639/RA-JAC

The decisions of the ISO 639-6/RA shall be restricted to the addition of new language code elements or changes to existing language code elements that are to be included in this part of ISO 639.

If the ISO 639-6/RA receives a registration that appears to designate a “major” language or a macrolanguage, the application shall be tabled for discussion with the ISO 639/RA-JAC, which will identify the appropriate form of registration.

The ISO 639-6/RA can consult the ISO 639/RA-JAC on any proposed additions or changes to the code specified by this part of ISO 639.

B.2 Responsibilities

B.2.1 Application for the registration of new language code elements and for the change of existing language code elements

The ISO 639-6/RA shall receive and review applications for the registration of new language code elements and for the change of existing ones. It shall act on such applications when the applicant:

a) has provided basic details of identification such as name, affiliation, and email address for correspondence purposes;

b) has provided suitable documentary evidence (at least the minimum documentary evidence needed as specified with reference to ISO/IEC 11179-6) supporting the requested change or inclusion.

When these criteria are met, the ISO 639-6/RA can consult relevant experts, and shall arrive at a decision in accordance with the rules given in B.1.4 and shall inform the applicant of the result as soon as is reasonably practicable but within a period of no more than four weeks from the time of application.

B.2.2 Maintenance of the registry

The ISO 639-6/RA shall:

a) maintain an accurate registry of information associated with registered language code elements in accordance with this part of ISO 639;

b) safeguard any confidential information;

c) process updates of registered language identifiers and alpha-4 language identifiers and distribute them on a regular basis to subscribers and other parties.

B.2.3 Reservation of alpha-4 language identifiers

When a request for inclusion of a new language code element has been rejected, ISO 639-6/RA might reserve the requested alpha-4 language identifier for the use of the applicant and other possible users. The ISO 639-6/RA shall keep a record of such reservations and shall inform the ISO 639-1/RA, the ISO 639-2/RA and the ISO 639-3/RA accordingly.

B.2.4 Creation of alpha-4 language identifiers

The creation of alpha-4 language identifiers is subject to the rules specified in ISO 639-1:2002, 4.1, and ISO 639-2:1998, 4.1.

B.2.5 Other general responsibilities

The ISO 639-6/RA shall:

a) handle all aspects of the registration process in accordance with good business practice;

b) indicate in its operations that it has been designated as ISO 639-6/RA by ISO Central Secretariat;

c) provide a six-monthly report to the ISO 639/RA-JAC of all change requests that have been processed;

d) provide an annual summary report on its activity to ISO Central Secretariat as well as to ISO/TC 37 and ISO/TC 46 Secretariats;

e) provide advice on the implementation and use of this part of ISO 639, as needed.

The web site for the ISO 639-6/RA is http://www.geolang.com/iso639-6/.

B.2.6 Status of registration

The ISO 639-6/RA shall provide a registration status for each language variant in accordance with the registration status list from ISO/IEC 11179-6:2005, Table 1, reproduced in Table B.1.

Table B.1 — ISO/IEC 11179-6 registration status list

Administered Item registration status category	Status criteria
Life cycle statuses	
Preferred standard	The Registration Authority confirms that the Administered Item is — preferred for use within the community that uses this metadata register.
Standard	The Registration Authority confirms that the Administered Item is — of sufficient quality, and — of broad interest for use in the community that uses this metadata register.
Qualified	The Registration Authority has confirmed that — the mandatory metadata attributes are complete, and — the mandatory metadata attributes conform to applicable quality requirements.
Recorded	The Registration Authority has confirmed that — all mandatory metadata attributes have been completed.
Candidate	The Administered Item has been proposed for progression through the registration levels.
Incomplete	The submitter[a] wishes to make the community that uses this metadata register aware of the existence of an Administered Item in their local domain.
Retired	The Registration Authority has approved the Administered Item as — no longer recommended for use in the community that uses this metadata register, and that — it should no longer be used.
Superseded	The Registration Authority determined that the Administered Item is — no longer recommended for use by the community that uses this metadata register, and that — a successor Administered Item is now preferred for use.
Documentation statuses	
Historical	The submitter[a] wishes to make the community that uses this metadata register aware of the existence of an Administered Item that was used in the past.
Application	The Registration Authority wishes to make the community that uses this metadata register aware of the existence of an Administered Item in their local domain that is in an application system and is not specified at the logical level. This item may be very well described.

[a] The term “submitter” in ISO/IEC 11179-6:2005 corresponds to “applicant” in this part of ISO 639.

Bibliography

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

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

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

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

[5] ISO 4217, *Codes for the representation of currencies and funds*

[6] ISO 8601, *Data elements and interchange formats — Information interchange — Representation of dates and times*

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

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

[9] ISO 16642, *Computer applications in terminology — Terminological markup framework*

[10] UN M49, Rev. 4, *Standard Country or Area Codes for Statistics Use*, 1999

[11] BCP 47 (IETF), *Tags for Identifying Languages*

[12] DALBY, D. *Linguasphere Register of the World's Languages and Speech Communities: Hebron (Wales)*. Linguasphere Press, 2000. 2 vol. ISBN 0-9532919-1-X, ISBN 0-9532919-2-8

[13] GRIMES, B.F. (ed.) *Ethnologue*, 14th ed. Dallas: SIL International, 2000. 2 vol. ISBN 1-55671-103-4, ISBN 1 55671 104 2

[14] PHILLIPS, A. and DAVIS, M. *Tags for identifying languages*. Internet Engineering Task Force, 2005. RFC 4646

[15] GILLAM, L. A metadata infrastructure using ISO standards. "*We Have to Talk about Metadata*" *Workshop at UK e-Science Programme All Hands Meeting 2007* (AHM 2007), Nottingham, 10-13 September 2007. Available at: http://www.allhands.org.uk/2007/proceedings/papers/791.pdf

[16] GILLAM, L., GARSIDE, D. and COX, C. Developments in Language Codes standards. In: Rehm, Witt and Lemnitzer (eds): *Datenstrukturen für linguistische Ressourcen und ihre Anwendungen/ Data Structures for Linguistic Resources and Applications*. Proc. of GLDV 2007, 11-13 April 2007, Tubingen, Germany: Gunter Narr Verlag. ISBN 978-3-8233-6314-9. Available at: http://www.cs.surrey.ac.uk/BIMA/people/l.gillam/download/publications/GLDV2007.final.LGDGCC.pdf

[17] TRUDGILL, P. *Sociolinguistics*. New York, Penguin, 1983, ISBN 0140224793

ICS 01.140.20

Price based on 16 pages