



Economic and Social Council

Distr.
LIMITED

E/CONF.91/CRP.11
12 January 1998

ENGLISH ONLY

SEVENTH UNITED NATIONS CONFERENCE
ON THE STANDARDIZATION OF
GEOGRAPHICAL NAMES
New York, 13-22 January 1998
Item 6(c) of the provisional agenda

TOPONYMIC DATA FILES: TOPONYMIC DATA TRANSFER
STANDARDS AND FORMATS

Report of the Working Group on Toponymic Data Exchange
Formats and Standards

Paper submitted by the United Nations Group of Experts
on Geographical Names**

* E/CONF.91/1

** Prepared by the Working Group

**Seventh United Nations Conference on the
Standardization of Geographical Names
New York, 13 - 22 January 1998**

Item 6.c.
of the Provisional Agenda

TOPONYMIC DATA FILES:

TOPONYMIC DATA TRANSFER STANDARDS AND FORMATS

REPORT OF THE UN GROUP OF EXPERTS ON GEOGRAPHICAL NAMES
WORKING GROUP
ON
TOPONYMIC DATA EXCHANGE FORMATS AND STANDARDS
TO THE
SEVENTH UNITED NATIONS CONFERENCE
ON THE STANDARDIZATION OF GEOGRAPHICAL NAMES*

Summary

At the 18th session of the United Nations Group of Experts on Geographical Names (UNGEGN), in Geneva, in 1996, a working group on "Toponymic Data Exchange Formats and Standards" was formed to investigate and recommend on the requirements, standards, and formats which are available for the encoding, processing, international exchange and promotion of nationally standardized geographical names for international use. This report presents the findings of the working group for consideration by UNGEGN and the Seventh UN Conference on the Standardization of Geographical Names.

* Submitted by the UNGEGN Working Group on Toponymic Data Exchange Formats and Standards.

Report of the UN Group of Experts on Geographical Names

Working Group

on

“Toponymic Data Exchange Formats and Standards”

to the

**Seventh United Nations Conference
on the Standardization of Geographical Names.**

New York.

January, 1998.

Report Contents

Working Group Report

Annex A Master List of roman characters required for geographical names processing, standardisation, promotion and exchange.

Part 1 Arranged alphabetically and by diacritical marks

Part 2 Arranged by international standard encoding address : ISO/Unicode.

Annex B World Survey of character encoding requirements (Roman and non Roman for geographical names processing, standardisation, promotion and exchange.

Part 1 List of Countries, Languages, Writing Systems, Romanisation Systems and International Standard Table References. (Note. The list follows the UN Terminology Bulletin order).

Part 2 Tables of Characters and ISO/Unicode encoding for each language, writing system and romanisation system. (Note. The tables are ordered by ISO language abbreviation),

Annex C Draft Toponymic Data Exchange Standard.

Annex D *Summary* description of the **ISO/IEC 10646 (Unicode)** standard.

Report of the UNGEGN Working Group on “Toponymic Data Exchange Formats and Standards”

Introduction

1. At the 18th session of the United Nations Group of Experts on Geographical Names (UNGEGN), in Geneva, in 1996, a working group on “Toponymic Data Exchange Formats and Standards” was formed to investigate and recommend on the requirements, standards and formats which are available for the encoding, processing, international exchange and promotion of nationally standardised geographical names for international use.

2. This report presents the findings of the WG for consideration by UNGEGN and the Seventh UN Conference on the Standardization of Geographical Names.

Need for the Working Group

3. Experts at the 18th UNGEGN recognised that many countries had made significant progress with the recording of geographical names files and data bases in support of map and gazetteer production and to meet other administrative, toponymic, national and international purposes. However, this was being achieved in individual and different ways and in the absence of internationally agreed formats and standards for the coding, processing and exchange of names data.

4. It was felt that the introduction of an internationally agreed digital format for gazetteers and the use of international standards for the encoding of character sets could assist international standardisation of geographical names.

5. Accordingly, the WG was set up to investigate and report on the potential of the existing and developing standards to allow countries to record, preserve and promote digitally their individual toponymic heritage whilst encouraging, through international exchange and publication, the international use of nationally approved and standardised names.

6. At UNGEGN it was recognised that the establishment of toponymic websites on the Internet by several countries gives considerable impetus to the promotion, standardisation and international use of nationally approved geographical names, But, for that potential to be realised by more countries, there is a need for achievable, low cost hardware and software solutions for names data processing and a need for easy to implement international **standards** and formats for toponymic **data**.

Working Group Membership and Meetings

7. **As** established at the 18th UNGEGN the working group members **were**:

Mr R Marsden Head of Geography and Geodesy, Military Survey, London, UK; and member of the UK Permanent Committee on Geographical Names (PCGN) (Convenor of the WG).

Mr R Flynn Executive Secretary for Foreign Names, United States Board on Geographic Names (BGN); and Geographer, National Imagery and Mapping Agency, Washington, USA.

Mr P Päll Head of Department of **Grammar**, Institute of the Estonian Language, Ministry of Education, Tallinn, Estonia.

8. The Secretaries to the WG were Mr D Whittington, Military Survey, UK and Mrs C Burgess, Research Assistant, PCGN, UK. In addition, the WG wishes to acknowledge the considerable assistance of the staff of the UK PCGN and the US BGN in completing the detailed research and analysis of requirements which is presented in this report.

9. To accomplish the work, the WG has met three times in the PCGN offices at the Royal Geographical Society, London and once at the Institute of the Estonian Language, Tallinn. However, significantly, much of the work has been exchanged between London, Tallinn and Washington using e-mail via the Internet. This report will highlight those exchanges to emphasise the extent to which international exchange of toponymic data is now readily achievable with low hardware and software costs.

Text Encoding Standards

10. The WG has examined the suitability of the existing international text encoding standards to meet the requirements for geographical names recording, processing and exchange. The following standards have been considered:

a) ISO/IEC 8859: 1987 International Organisation for Standardisation. Information processing - 8 bit single byte coded graphic character sets. [Geneva], 1987.

b) ISO/IEC 10646-1:1993 International Organisation for Standardisation. Information Technology - Universal Multiple-Octet Coded Character Sets (UCS) [Geneva] 1993. [i.e. 16 bit codes sometimes referred to as wide **ASCII**]

c) The Unicode Standard Version 2.0, The Unicode Consortium, Addison-Wesley Developers Press, Reading, Massachusetts. July 1996.

11. The Unicode Consortium comprises the principal hardware and software corporations acting together to promote the standard and to introduce applications software which is consistent with the standard. The Unicode coding standard is now identical to the ISO/IEC 10646 standard and the two are maintained and revised in parallel. The analysis in this report therefore treats Unicode and ISO/IEC 10646 as a single 16 bit standard.

12. The Unicode/ISO 10646 standard is a fixed-width, uniform encoding scheme for written characters and text. The repertoire of this international character code for information processing includes characters for the major scripts of the world. The character encoding treats alphabetic characters, ideographic characters, and symbols identically, which means that they can be used in any mixture. The Unicode Standard is modelled on the **ASCII** character set, but uses a 16 bit encoding to support **full** multilingual text. No escape sequence or control code is required to specify any character in any language. There are 65,000 sixteen bit codes available of which 39,000 have been allocated to date. The Unicode/ISO 10646 standard builds upon the ISO 8859-1 standard as can be seen in Table 1.

ISO 8859-1 Text	ISO 10646/Unicode Text
0100 0001	0000 0000 0100 0001
0101 0011	0000 0000 0101 0011
0100 0011	0000 0000 0100 0011
0100 1001	0000 0000 0100 1001

Table 1

13. A technical **summary** of the design principles and the advantages of the ISO 10646/Unicode standard is given at Annex D. The standard clearly has considerable potential for the processing of standardised and nationally approved geographical names both in roman and in most non-roman scripts.

Survey of Text Encoding Requirements for Geographical Names.

14. One of the main **tasks** of the WG has been to assess the completeness and the utility of the above standards for the exchange of geographical names. This has been achieved by conducting a world wide survey, on a country and language basis, of the main writing and romanisation systems required for the recording, processing, standardisation, exchange and promotion of nationally approved geographical names for international use.

15. The survey has identified for each country, language, writing system and romanisation system the availability of 8 bit, ISO 8859 and 16 bit, ISO 10646/Unicode codes for each character. The survey has concentrated on the encoding requirements for roman and romanised characters since these are the principal vehicle for international **names** exchange and standardisation. However, the survey **has** also included a reference to the 16 bit, ISO 10646/Unicode **codes** for the non roman scripts where available since these **are** essential for national names recording and standardisation purposes and may be required for international use in the future.

16. The survey **has** focused on the writing and romanisation systems which **are** in current use by names standardisation authorities. These **are** the systems which have the most potential for **future** geographical **names** exchange and **international** standardisation. **Although** the standards can, generally, **also** handle **other** romanisation systems and **ancient** or historic writing systems and **can** be **used** for bibliographic **purposes**, those systems **are** of more limited use for international geographical **names** standardisation.

Results of the Survey of Requirements

17. The results of the survey of requirements are tabulated at **Annexes A and B**.

- a) **Annex A Master List** Provides two master lists of characters which summarise on a world-wide basis the results of the survey. The part 1 master list is sorted alphabetically and by diacritical mark. The part 2 master list provides the same data sorted by ISO/Unicode code order and is provided for ease of reference to those standards.
- b) **Annex B Part 1 List of Countries** Provides for each country the main languages, writing systems and romanisation systems which are required for the processing of most of the current geographical names which occur in the country concerned. The lists are not fully comprehensive. Additional languages and writing and romanisation systems may be required for some countries. The list has taken account of the UN Toponymic Guidelines for each **country**, where these are available. The final column in Annex B Part 1 identifies the ISO/Unicode code **address** for the non-roman writing systems and provides references to the tables of characters at Annex B **Part 2**.
- c) **Annex B Part 2 Tables of Characters** Provides the tables of characters for basic roman and each of the extended **roman** alphabet character sets required for the processing of **geographical names**. These tables are ordered by language codes (ISO 639). Each table provides the character, the character name (following the Unicode standard), the ISO 8859 code and the ISO 10646/Unicode code for each character. From these tables the completeness of the 8 bit and 16 bit standards to meet the requirements for encoding **digital** geographical names for any purpose may be readily assessed.

Analysis of the Requirements

18. From Annex A it can be seen that, for the languages, writing systems and romanisation systems included in the survey, the ISO 10646/Unicode encoding standard has the potential to meet almost all of the requirements for national and international geographical names processing and exchange. Of almost **500** roman characters identified, the ISO 10646/Unicode standard **accounts** for all but 7 characters. These exceptions are given in Table 2 and their existence may require additions to the ISO/Unicode standard **after** further consideration by the names, standardisation and other authorities of the countries concerned. Before the potential of the relatively new 16bit encoding **standard can** be fully realised, improved applications software for text processing and transfer will be required.

	Character	Language/s	Countries involved
1.	Āī	Devanagari	India
2.	āī	Devanagari	"
3.	Ī	Turkmen	Turkmenistan
4.	G‘	Uzbek	Uzbekistan
5.	g‘	Uzbek	"
6.	O‘	Uzbek	"
7.	o‘	Uzbek	"

Table 2

19. From Annex B, it *can* be seen from the number of gaps in the **standard** (indicated with a dash) that the 8 bit, ISO 8859 character encoding requirements for geographical names processing are less well met. Although the ISO 8859 standard has been established for some time and now has 14 parts there are many roman characters required for geographical **names** which are not covered by the **standard**. However, for ISO 8859 the applications software for processing and transfer is more widely available; particularly for ISO 8859 Latin 1 characters which can be easily accomplished with word processing, data base and production software and readily exchanged via the Internet.

Implementation of Encoding Standards

20. The WG has completed some analysis of the practical aspects of the implementation of the ISO standards for names processing. The analysis is flagged in the notes to the tables at Annex B **Part 2**. The following aspects **are** noted:

- a) **Note 1** There are cases where romanisation systems have been approved by the United Nations but have not been implemented by the **country** concerned. *eg. romanisation system for Devanagari, approved by the 2nd UN Conference on the Standardization of Geographical Names, 1972(Conf. Res: II/11).*
- b) **Note 2** There are cases where non-Roman writing systems have not yet been catered for in the ISO10646/Unicode standard. *eg. Khmer, Divehi, Mongolian. It would be the responsibility of the national standards authorities of the countries concerned to approach ISO/Unicode with the possibility of extending the standard to include these scripts and also individual roman-script characters which are not yet available (see c) below).*
- c) **Note 3** There are instances where Roman characters are not available in: (a) ISO 8859 ,(b) ISO 10646. *eg. Ŏ ŏ used in the McCune-Reischauer romanisation system for Korean is not possible in ISO 8859. All such cases are indicated in Annex B with a dash. Table 2 at paragraph 18 provides a full list of characters which are not available in ISO 10646/Unicode.*
- d) **Note 4** There are instances where characters could not be encoded as a single unified character but only in conjunction with a combining diacritical mark. *eg. in romanised Hebrew, Ĥ can only be encoded in ISO 10646/Unicode using the letter H (code 0048) plus the combining diacritic "line below" (0331), giving rise to the encoding 0048 + 0331 as shown in Annex B, part 2.*
- e) **Note 5** There are cases where the **word** processing software and fonts **used** by the WG could not display or print the character. Formatting **commands** (e.g. for underline) or, as a last resort, manuscript additions had to be made to the hardcopy. *Using the Windows 95 in-built font sets, the following characters could only be managed using formatting commands in Microsoft Word for Windows 95 version 7.0: Ĳ ĳ Ľ ľ Ń ń Ŝ ŝ Ũ ů Ż ż and the following with manuscript additions: Ā ā Ą ą Ć ć Ę ę Ħ ħ Ĭ ĭ Ū ū Ŭ ŭ Ű ű*

In WordPerfect 6.1 the overstrike facility or manuscript additions had to be used to produce approximately 50 characters

- f) **Note 6** Multiple coding. There are several cases identified by the survey where the character appears more than once in the ISO 10646/Unicode standard with different code references. *eg. the letter "a" found in Azerbaijani Roman is given as both 01DD and 0259 in ISO 10646/Unicode. If both users have different subsets of the ISO 10646/Unicode sets, the characters may be treated differently by different software packages*
- g) **Note 7** There are cases where codes are available in ISO 8859 but are difficult to implement as they occur in different parts of the standard. *eg. Tajik*

Transfer of Toponymic Data

21. The WG has considered the practical **aspects** which are involved for countries to implement the international encoding standards for digital **geographical names** recording, standardisation and promotion through international publication and exchange. The WG has been particularly mindful of the **need**, for all **countries**, to achieve low cost hardware and software solutions. The WG has demonstrated, through the e-mail exchanges between London, Tallinn and Washington of most of the required characters at Annex **A** and **B**, the feasibility of utilising the encoding standards. This has been done with relatively low cost, commercial word processing and font sets and PC hardware. The Working Group used the following resources* :

- a) Operating Systems: Microsoft Windows 95; Microsoft Windows NT v. 3.51
- b) Wordprocessing packages: Microsoft Word v. 7.0 for Windows 95, WordPerfect v. 6.1
- c) Font sets: Windows 95 fonts including MS LineDraw, Times New Roman CE, Times New Roman Special G1 & G2 (©Monotype Corporation), Armenian (©WYSIWYG Corporation); GDE Systems GeoName fonts (©Bitstream Inc) including GeoAME_CentSchbk GeoAPC_CentSchbk, GeoEEU_CentSchbk, GeoVTM_CentSchbk; WordPerfect 6.1 special fonts, UHoài 1.1 font for Vietnamese (©Cuong Bui - TriChlor Groupe Distribute available from the Internet)
- d) Database software: bespoke software (text-type **data** files managed with programmes using TurboPascal)
- e) Spreadsheet software: Microsoft Excel v. 7.0 for Windows 95
- f) E-mail software: Eudora Light; Microsoft Mail, Netscape Mail

22. The principal difficulty experienced by the Working Group in exchanging information was the incompatibility of encoded text (beyond ISO 8859 Latin 1) between WordPerfect and the Microsoft products, due primarily to the proprietary encoding **standards** used by the two **software** developers. When the WordPerfect v. 6.1 files were converted into Word v. 7.0, many of the characters beyond ISO 8859 **Latin-1** were lost or misrepresented. Characters created by using formatting **commands** were, however, preserved. The incompatibility could have been avoided for characters supported by the full suite of ISO 8859 sets had a set of common fonts been available at all **three** Working Group sites. Unicode compatibility could have been achieved if the latest **versions** of Microsoft Internet Explorer or Netscape Communicator were used at the **three** locations.

* The identification of the **software** listed here is **not** an endorsement by the Working Group or its **members** of the products or the manufacturers.

23. Analysis of Annex B to this Report shows that international exchange of digital toponymic information is largely achievable if the exchanging parties coordinate the content and format of the data sets. In the context of international standards for text encoding, an operational implementation of ISO 8859 in an eight-bit environment would satisfy data exchange requirements for 60 of the 114 language-specific character tables in Annex B. The implementation of ISO 8859 is dependent only on the availability of common fonts. Implementation of a sixteen-bit environment, including the ISO 10646/Unicode standard, would support the exchange of data covering 113 of the 114 character tables. This would require specific sixteen-bit-enabled software (to include operating systems, applications, and fonts) for the data exchange to succeed.

Toponymic Data Exchange Formats

24. In addition to the review of the toponymic data text encoding requirements and standards the second task of the WG has been to identify the toponymic data and formats which might be exchanged in digital form to promote international names standardisation. The WG has considered the format and content requirements for the toponymic data itself as well as for the data about the data (the metadata) which is essential for the successful transfer, understanding and utilisation of the names data.

25. The approach taken by the WG was to review the fields, content and format from a range of gazetteer specifications, some of which are in international use and to compare these with the UN resolution I/4 (E) recommended gazetteer fields. The gazetteer specifications and data sets examined included:

- a) Australian Geographic Place Names Gazetteer, Australian Surveying & Land Information Group (Internet server). <http://kaos.erin.gov.au/database/MAN200R.html>
- b) Canada's Official Geographical Names Web Site, Canadian Permanent Committee on Geographical Names, Natural Resources Canada (Internet server). <http://geonames.nrcan.gc.ca/>
- c) Luxembourg Gazetteer of Villages and Towns, Service Central de la Statistique et des Etudes Économiques (Internet server). <http://www.igd-leo.lu/igd-leo/onomastics/villages/viliages.html>
- d) NATO Standardization Agreement 2213 "Gazetteers," 3rd edition, 1980.
- e) Pan American Institute of Geography and History Gazetteer Specification (draft), no date.
- f) Répertoire Toponymique du Québec, Commission de toponymie, Québec, 1987.
- g) U.K. Ordnance Survey Gazetteer of Great Britain, 1987.
- h) U.S. Defense Mapping Agency Product Specification for Gazetteers, 3rd edition, 1983.
- i) U.S. Defense Mapping Agency Military Specification - Digital Gazetteer (draft), 1994.
- j) U.S. Board on Geographic Names Geographic Names Information System, U.S. Geological Survey (Internet server). <http://nsdi.usgs.gov/nsdi/products/gnis.html>
- k) U.S. Board on Geographic Names GEOnet Names Server, U.S. National Imagery and Mapping Agency (Internet server). <http://www.nima.mil/gns/html/index.html>

26. The WG recommendations on a proposed toponymic data exchange standard are at Annex C. The WG recommendations follow UN Resolution I/4 (E) for gazetteer content. The recommended minimum content for toponymic data exchange and for digital gazetteers is therefore as follows:

- a Name (including variants).
- b Feature Type.
- c Latitude and Longitude.
- d Country and administrative division code (in accordance with ISO 3166)
- e Map series, sheet number, edition, where appropriate.

Annex C provides a proposed specification for these fields together with guidelines for population and some example names records.

27. From the discussions noted above on the importance of metadata for the efficient transfer, understanding and use of nationally standardised toponymic data internationally, it was concluded by the WG that any proposed standard for digital toponymic data must include some provision for metadata. The difficulty is that there are no international metadata standards for any form of geographic data. There are some draft metadata **standards** but these are immature at present. The WG therefore recommends, **pending** the development of metadata **standards** for geographic data, that any digital names data exchanged on magnetic media or **made** available via the Internet should **contain** some accompanying **data** which describes the **data** sufficiently to assist the user in reading and using the data particularly in respect of the text encoding standards and software used to create the data. The WG recommendations on the content of the metadata are given at Annex C.

Summary

28. The WG has completed an initial survey of the requirements for the text encoding of nationally approved and standardised geographical names and a review of the suitability of the existing international standards to meet the encoding requirements for national and international names standardisation and international exchange. The WG has proposed a first international toponymic data exchange and digital gazetteer format for names data and suggestions for the essential descriptive metadata which must accompany the names data to assist users.

Conclusions

29. The WG has reached the following conclusions:

- a) The promotion of nationally standardised and approved geographical names for international standardisation and use requires international agreement on "Toponymic Data Exchange Formats and Standards".
- b) Digital recording, processing, publication and exchange of nationally standardised and approved geographical names written in roman, extended roman and non-roman writing systems is technically achievable now, by **most countries**, at relatively low costs for hardware and **software**.
- c) The publication of nationally approved and **standardised geographical names** on the Internet, or on magnetic media, **has** considerable potential to meet UN goals and conference resolutions for the international **standardisation** of names and significant potential to **allow countries** to preserve and promote their individual toponymic heritage.

- d) The current international text encoding standards (ISO/Unicode) meet most of the requirements for geographical names recording, processing, exchange and publication,
 - (i) 8 bit and/or 16 bit text encoding standards can be implemented now for most languages)
 - (ii) The remaining languages, with the exception of those shown in Table 2, can be supported only with a 16-bit text encoding standard
- e) The implementation of a practical programme of digital toponymic data exchange is handicapped by the incompatibility of some current commercial software with international text encoding standards.
- f) The Working Group has proposed an exchange format for **digital** names **data**, gazetteers and accompanying explanatory data which is based upon previous UN resolutions for hardcopy gazetteers (Annex C). **This digital** format could be used to promote international standardisation and use of nationally approved geographical names far more effectively than has been possible through hardcopy means.

Recommendations

30. The WG makes the following recommendations:

- a) Countries should aim to promote their toponymic heritage and the international standardisation and use of their nationally approved and standardised names by all available means including the Internet and digital gazetteer publication and exchange.
- b) Countries should review the text encoding requirements for geographical names which are included in this report to identify any additional requirements. These might be notified to UNGEGN through “Toponymic Guidelines”. Additional requirements may require extensions to the text encoding standards which should be taken up by each country’s standardisation authorities with ISO/Unicode.
- c) Countries should promote the international standardisation of geographical names through the digital exchange and publication of names and associated data using the formats and standards which have been proposed by the Working Group. Those dealing with the automation of geographical names are encouraged to use software which is compliant with the international text encoding standards identified in this report.
- d) The requirements for “Toponymic Data Exchange Formats and Standards” should be kept under review by the Working Group on Toponymic Data Files and Gazetteers.
- e) The Working Group on Toponymic Data Files and Gazetteers should encourage the implementation of the gazetteer and encoding standards by **countries**, through the UNGEGN Divisions; and with the Unicode consortium as a user group and as an example of application of the **standards**.

The WG invites comments, suggestions **and** proposals which should be addressed to:

Roger Marsden
c/o PCGN
Royal Geographical Society
1 Kensington Gore
London
s w 7 2AR
United Kingdom

Fax: **(44)**171 591 3127

E-mail may be directed to the following addresses:

peeter@eki.ee
flynnre@nima.mil
cburgess@pcgn.u-net.com

Master List of Characters

Key to Columns below:

Char. (Column 1) - characters are in alphabetical order with upper case characters before lower case characters. The order of diacritics is taken from the ISO/Unicode Table 8 Row 03 (Combining Diacritical Marks). Ligatures are given in the order of the first letter of the ligature, after combinations of the same letter with diacritical marks. Other variant characters are given next to their closest counterpart in the Basic Roman character table.

Order of combining diacritical marks (only those occurring in the tables are indicated):

0300 - grave accent (Varia)	030C - caron	032b - circumflex below
0301 - acute accent (Oxia)	0310 - candrabindu	032E - breve below
0302 - circumflex	0312 - turned comma above	0331 - line (macron) below
0303 - tilde	0315 - comma above right	0335 - short stroke overlay
0304 - macron	031B - horn	0337 - short solidus overlay
0306 - breve (Vrachy)	0323 - dot below	0338 - long solidus overlay
0307 - dot above	0324 - diaeresis below	0340 - grave tone mark (Vietnamese)
0308 - diaeresis (Dialytika)	0326 - comma below	0341 - acute tone mark (Vietnamese)
030A - ring above	0327 - cedilla	
030B - double acute accent	0328 - ogonek	

Character Name (Column 2) - in general the ISO/Unicode terms are used with the exception of "Latin" being replaced by "Roman". Names given in brackets follow the same pattern, but describe characters which cannot be found in a unified form in this standard.

ISO/Unic. (Column 3) - four digit code of the ISO 10646/Unicode is given unless the character is not Unified. In the latter case two codes with + sign are given, consisting usually of the code of the basic character and that of the combining diacritical mark.

References to Tables of Characters (Column 4) - gives references to tables in Annex B Part 2 where the characters occur.

Note: to acquire the ISO 8859 codes for characters refer to individual tables in Annex B Part 2

Part 1. Sorted alphabetically

<u>Char.</u>	<u>Character Name</u>	<u>ISO/Unic.</u>	<u>References to Tables of Characters</u>
&	Ampersand	0038	3
	Inverted apostrophe	02BB (2018)	am_R2, ar_R1, fa-R, hawa, he_R, ps_R, sm, to
	Reversed apostrophe	02BD (201B)	ar_R2
	Acute accent	00B4	et
	Quotation mark	0034	ru_R1, ru_R2
A	Roman capital letter A with grave	00C0	ca, gd, fr, it, mt, pt, rm, sámí2, tl, vi, zh_R1
a	Roman small letter a with grave	00E0	ca, gd, fr, it, mt, pt, rm, sámí2, tl, vi, zh_R1
Á	Roman capital letter A with acute	00C1	ar_R1, cs, cy, el-R, es, fa-R, fo, ga, gd, gl, gn, hu, is, km_R, mg, nl, pt, sámí1, sk, tl, vi, zh_R1
á	Roman small letter a with acute	00E1	ar_R1, cs, cy, el-R, es, fa-R, fo, ga, gd, gl, gn, hu, is, km_R, mg, nl, pt, sámí1, sk, tl, vi, zh_R1

Â	Roman capital letter A with circumflex	00C2	ar_R2, ar_R4, br, cham, cy, fr, fy, km_R, nl, pt, ro, tl, tr, vi, wo
â	Roman small letter a with circumflex	00E2	ar_R2, ar_R4, br, cham, cy, fr, fy, km_R, nl, pt, ro, tl, tr, vi, wo
Ă	Roman capital letter A with circumflex and grave	1EA6	vi
ă	Roman small letter a with circumflex and grave	1EA7	vi
Ȃ	Roman capital letter A with circumflex and acute	1EA4	vi
ȃ	Roman small letter a with circumflex and acute	1EA5	vi
Ã	Roman capital letter A with circumflex and tilde	1EAA	vi
ã	Roman small letter a with circumflex and tilde	1EAB	vi
Ą	Roman capital letter A with circumflex and hook above	1EA8	vi
ą	Roman small letter a with circumflex and hook above	1EA9	vi
Ȧ	Roman capital letter A with circumflex and dot below	1EAC	vi
ȧ	Roman small letter a with circumflex and dot below	1EAD	vi
ȧ	(Roman small letter a with circumflex and ogonek)	00E2+0328	3
Ä	Roman capital letter A with tilde	00C3	pt, vi
ä	Roman small letter a with tilde	00E3	gn, pt, vi
Å	Roman capital letter A with macron	0100	2, am_R2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hawa, hi_R, kn_R, lívo, lv, mars, mi, ml_R, or-R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
å	Roman small letter a with macron	0101	2, 3, am_R2, ar_R1, ar_R3, as_R, bn_R, fa-R, gu_R, hawa, hi_R, kn_R, lívo, lv, mars, ml_R, mi, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
Ȧ̆	(Roman capital letter A with macron and breve)	0100+0306	hi-R
ȧ̆	(Roman small letter a with macron and breve)	0101+0306	hi-R
Ȧ̇	Roman capital letter A with breve	0102	km_R, ro, sámi2, vi, wo
ȧ̇	Roman small letter a with breve	0103	km_R, ro, sámi2, vi, wo
Ȧ̈	Roman capital letter A with breve and grave	1EBO	vi
ȧ̈	Roman small letter a with breve and grave	1EB1	vi
Ȧ̉	Roman capital letter A with breve and acute	1EAE	vi
ȧ̉	Roman small letter a with breve and acute	1EAF	vi
Ȧ̊	Roman capital letter A with breve and tilde	1EB4	vi

ă	Roman small letter a with breve and tilde	1EB5	vi
Ă	Roman capital letter A with breve and hook above	1EB2	vi
ȃ	Roman <i>small</i> letter a with breve and hook above	1EB3	vi
Ȑ	Roman capital letter A with breve and dot below	1EB6	vi
ȑ	Roman small letter a with breve and dot below	1EB7	vi
Ȓ	(Roman capital letter A with dot above)	0041+0307	ulit
ȓ	(Roman small letter a with dot above)	0061+0307	ulit
Ȕ	Roman capital letter A with dot above and macron	01E0	sámi2
ȕ	Roman small letter a with dot above and macron	01E1	sámi2
Ä	Roman capital letter A with diaeresis	00C4	az, cy, de, et, fi, kk_R, lb, livo, sk, sv, tk, tk_R, yape
ä	Roman small letter a with diaeresis	00E4	az, cy, de, et, fi, kk_R, lb, livo, sk, sv, tk, tk_R, yape
Ȧ	Roman capital letter A with diaeresis and macron	01DE	livo, sámi2
ȧ	Roman small letter a with diaeresis and macron	01DF	livo, sámi2
Ȧ	Roman capital letter A with hook above	1EA2	vi
ȧ	Roman small letter a with hook above	1EA3	vi
Ȧ	Roman capital letter A with ring above	00C5	da, no, sv
ȧ	Roman small letter a with M g above	00E5	da, no, sv
Ȧ	Roman capital letter A with caron	01CD	bg_R1, zh_R1
ȧ	Roman small letter a with caron	01CE	bg_R1, zh_R1
Ȑ	Roman capital letter A with dot below	1EA0	vi
ȑ	Roman small letter a with dot below	1EA1	vi
Ȓ	Roman capital letter A with ogonek	0104	lt, pl
ȓ	Roman small letter a with ogonek	0105	lt, pl
Ȕ	Roman capital letter schwa	018F	az
ȕ	Roman <i>small</i> letter schwa	0259	az
		(OIDD)	
Æ	Roman capital letter AE (ash)	00C6	da, fo, is, no
æ	Roman small letter ae (ash)	00E6	da, fo, is, no
Ȧ	Roman capital letter AE (ash) with macron	01E2	sámi2
ȧ	Roman small letter ae (ash) with macron)	01E3	sámi2
Ȧ	Roman capital letters AI with breve		hi-R
ȧ	Roman <i>small</i> letters ai with breve		hi-R
Ȧ	Roman <i>capital</i> letter B with hook	0181	ha
ȧ	Roman <i>small</i> letter b with hook	0253	ha
Ȧ	Roman <i>capital</i> letter C with acute	0106	bosn, hr, mk_R, pl, sorbl, sorb2, sr_R
ć	Roman <i>small</i> letter c with acute	0107	bosn, hr, mk_R, pl, sorbl, sorb2, sr_R
Ȧ	Roman <i>capital</i> letter C with dot above	010A	mt

c	Roman small letter c with dot above	010B	mt
Č	Roman capital letter C with caron	010C	bg_R1, be-R, bosn, cs, hr, lt, lv, mk_R, ru_R1, ru_R3, sámi1, sorb1, sorb2, sk, sl, sr_R
č	Roman small letter c with caron	010D	bg_R1, be-R, bosn, cs, hr, lt, lv, mk_R, ru_R1, ru_R3, sámi1, sorb1, sorb2, sk, sl, sr_R
Ç	Roman capital letter C with cedilla	00C7	ar_R4, az, ca, eu, fr, sq, pt, tk, tr
ç	Roman small letter c with cedilla	WE7	ar_R4, az, ca, eu, fr, sq, pt, tk, tr
Ď	Roman capital letter D with caron	010E	cs, sk
ď	Roman small letter d with caron	010F	cs, sk
D	Roman capital letter D with dot below	1E0C	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or-R, pa_R, ta_R, te_R, ur_R
d	Roman small letter d with dot below	1E0D	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or-R, pa-R, ta_R, te_R, ur_R
Ð	Roman capital letter D with cedilla	1E10	ar_R1, lívo
ð	Roman small letter d with cedilla	1E11	ar_R1, lívo
Ɖ	Roman capital letter D with circumflex below	1E12	vend
ɖ	Roman small letter d with circumflex below	1E13	vend
Ḑ	Roman capital letter D with line below	1E0E	ar_R3, ps_R
ḑ	Roman small letter d with line below	1E0F	ar_R3, ps_R
Đ	Roman capital letter D with stroke	0110	bosn, hr, mk_R, sámi1, sr_R, vi
đ	Roman small letter d with stroke	0111	bosn, hr, mk_R, sámi1, sr_R, vi
Ḍ	Roman capital letter D with hook	018A	ha
ḏ	Roman small letter d with hook	0257	ha
Ð	Roman capital letter ETH	00D0	fo, is
ð	Roman small letter eth	00F0	fo, is
Ɖ	Roman capital letter EZH	01B7	sámi2
ɖ	Roman small letter ezh	0292	sámi2
Ḑ	Roman capital letter EZH with caron	01EE	sámi2
ḑ	Roman small letter ezh with caron	01EF	sámi2
E	Roman capital letter E with grave	00C8	af, ar_R4, be_R, ca, fr, gd, it, lo_R, mt, my_R, nl, pt, rm, ru_R1, sámi2, tl, vi, wo, zh_R1
è	Roman small letter e with grave	00E8	af, ar_R4, be_R, ca, fr, gd, it, lo_R, mt, my_R, nl, pt, rm, ru_R1, sámi2, tl, vi, wo, zh_R1
É	Roman capital letter E with acute	00C9	af, ar_R4, bi, ca, cs, cy, da, el_R, es, fr, ga, gd, gl, gn, he_R, hu, id, is, it, km_R, lb, lo_R, nl, no, pt, rm, sk, sv, tl, vi, wo, zh_R1
é	Roman small letter e with acute	00E9	af, ar_R4, bi, ca, cs, cy, da, el_R, es, fr, ga, gd, gl, gn, he_R, hu, id, is, it, km_R, lb, lo_R, nl, no, pt, rm, sk, sv, tl, wo, vi, zh_R1
Ê	Roman capital letter E with circumflex	00CA	af, ar_R4, br, cy, fr, fy, km_R, nl, ps_R, pt, tl, tn, vi, zh_R2, zh_R1
ê	Roman small letter e with circumflex	00EA	af, ar_R4, br, cy, fr, fy, km_R, nl, ps_R, pt, tl, tn, vi, zh_R2, zh_R1
È	Roman capital letter E with	1EC0	vi

è	circumflex and grave Roman small letter e with circumflex and grave	1EC1	vi
É	Roman capital letter E with circumflex and acute	1EBE	vi
é	Roman small letter e with circumflex and acute	1EBF	vi
Ê	Roman capital letter E with circumflex and tilde	1EC4	vi
ẽ	Roman small letter e with circumflex and tilde	1EC5	vi
Ë	Roman capital letter E with circumflex and hook above	1EC2	vi
ë	Roman small letter e with circumflex and hook above	1EC3	vi
Ĕ	Roman capital letter E with circumflex and dot below	1EC6	vi
ĕ	Roman small letter e with circumflex and dot below	1EC7	vi
Ë	Roman capital letter E with tilde	1EBC	vi
ẽ	Roman small letter e with tilde	1EBD	gn, vi
Ê	Roman capital letter E with macron	0112	am_R2, livo, lv, mi, zh_R1
ē	Roman small letter e with macron	0113	3, am_R2, livo, lv, mi, zh_R1
Ē	Roman capital letter E with breve	0114	gu_R, hi_R, km_R, kn_R, ml_R, sámí2, ta_R, te_R
ē	Roman small letter e with breve	0115	gu_R, hi_R, km_R, kn_R, ml_R, sámí2, ta_R, te_R
Ė	Roman capital letter E with dot above	0116	kk_R, lt, tg_R, ulit
ė	Roman small letter e with dot above	0117	kk_R, lt, tg_R, ulit
Ë	Roman capital letter E with diaeresis	00C3	af, be_R, cy, fr, lb, nl, ru_R1, ru_R2, sq, yape
ë	Roman small letter e with diaeresis	00EB	af, be-R, cy, fr, lb, nl, ru_R1, ru_R2, sq, yape
Ë	Roman capital letter E with hook above	1EBA	vi
ë	Roman small letter e with hook above	1EBB	vi
Ě	Roman capital letter E with caron	011A	cs, sorbl, sorb2, zh_R1
ě	Roman small letter e with caron	011B	cs, sorbl, sorb2, zh_R1
Ė	Roman capital letter E with dot below	1EB8	vi, yo
ė	Roman small letter e with dot below	1EB9	vi, yo
Ę	Roman capital letter E with ogonek	0118	lt, pl
ę	Roman small letter e with ogonek	0119	lt, pl
Ǝ	(Roman capital letter E with line below)	0045+0331	am_R1
ɛ	(Roman small letter e with line below)	0065+0331	am_R1
Ǝ	(Roman capital letter F with line below)	0046+0331	el-R
ƒ	(Roman small letter f with line below)	0066+0331	el_R
Ĝ	(Roman capital letter G with tilde)	0047+0303	tl
ĝ	(Roman small letter g with tilde)	0067+0303	gn, tl
Ĝ	Roman capital letter G with breve	011E	az, tr
ĝ	Roman small letter g with breve	011F	az, tr
Ġ	Roman capital letter G with dot above	0120	mt

ġ	Roman small letter g with dot above	0121	mt
Ĝ	Roman capital letter G with caron	01E6	sámi2
ǧ	Roman small letter g with caron	01E7	sámi2
ǧ	Roman small letter g with turned comma above	0123	lv
Ġ	(Roman capital letter G with dot below)	0047+0323	pa_R, ur_R
ğ	(Roman small letter g with dot below)	0067+0323	pa_R, ur_R
Ĝ	Roman capital letter G with cedilla	0122	lv
Ġ	(Roman capital letter G with line below)	0047+0331	3
ġ	(Roman small letter g with line below)	0067+0331	3
Ĝ	Roman capital letter G with stroke	01E4	sámi2
ǧ	Roman small letter g with stroke	01E5	sámi2
G´	(Roman capital letter G with inverted apostrophe)		uz
g´	(Roman small letter g with inverted apostrophe)		uz
Ĥ	Roman capital letter H with dot below	1E24	ar_R2, ar_R4, as_R, bn_R, hi_R, kn_R, ml_R, or-R, te_R
ĥ	Roman small letter h with dot below	1E25	ar_R2, ar_R4, as_R, bn_R, hi-R, kn_R, ml_R, or-R, te_R
Ĥ	(Roman capital letter H with diaeresis below)	0048+0324	wo
ĥ	(Roman small letter h with diaeresis below)	0068+0324	wo
Ħ	Roman capital letter H with cedilla	1E28	ar_R1, fa_R, ps_R
ĥ	Roman small letter h with cedilla	1E29	ar_R1, fa_R, ps_R
Ĥ	(Roman capital letter H with line below)	0048+0331	ar_R3, be_R, he_R
ĥ	Roman small letter h with line below	1E96	ar_R3, be_R, he-R
Ĥ	Roman capital letter H with stroke	0126	mt
ĥ	Roman small letter h with stroke	0127	mt
İ	Roman capital letter I with grave	00CC	gd, it, mt, rm, tl, vi, zh_R1
ı	Roman small letter i with grave	00EC	gd, it, mt, rm, tl, vi, zh_R1
Í	Roman capital letter I with acute	00CD	ca, cs, el-R, es, fo, ga, gl, gn, hu, is, it, nl, pt, sk, tg_R, tl, vi, zh_R1
ı	Roman small letter i with acute	WED	ca, cs, el-R, es, fo, ga, gl, gn, hu, is, it, nl, pt, sk, tg_R, tl, vi, zh_R1
Î	Roman capital letter I with circumflex	00CE	af, ar_R2, ar_R4, cy, fy, fr, mt, rm, ro, tl, tr
ı	Roman small letter i with circumflex	WEE	af, ar_R2, ar_R4, cy, fy, fr, mt, rm, ro, tl, tr
İ	Roman capital letter I with tilde	0128	vi
ı	Roman small letter i with tilde	0129	gn, vi
Ī	Roman capital letter I with macron	012A	2, am_R2, ar_R1, ar_R3, as_R, bn_R, fa-R, gu_R, hi-R, kk_R, kn_R, livo, lv, mi, ml_R, or-R, pa-R, ps_R, ta_R, te_R, ur_R, zh_R1
ı	Roman small letter i with macron	012B	2, am_R2, ar_R1, ar_R3, as_R, bn_R, fa-R, gu_R, hi_R, kk_R, kn_R, livo, lv, mi, ml_R, or_R, pa_R, ps_R, ta_R,

			te_R, ur_R, zh_R1
İ	Roman capital letter I with breve	012C	km_R, mn_R
ï	Roman small letter i with breve	012D	km_R, mn_R
ƒ	Roman capital letter I with dot above	0130	az, tr
Ī	Roman capital letter I with diaeresis	00CF	af, ar_R4, bi, ca, cy, el_R, fr, mn_R, nl, uk_R
ï	Roman small letter i with diaeresis	00EF	af, ar_R4, bi, ca, cy, el_R, fr, mn_R, nl, uk_R
Ī	(Roman capital letter I with diaeresis and acute)	00CF+0301	el_R
ï	(Roman small letter i with diaeresis and acute)	00EF+0301	el-R
İ̇	Roman capital letter I with hook above	1EC8	vi
ï̇	Roman small letter i with hook above	1EC9	vi
ƒ̃	Roman capital letter I with caron	01CF	zh_R1
Ī	Roman small letter i with w o n	01D0	zh_R1
İ̇	Roman capital letter I with dot below	1ECA	ibo, vi
ï̇	Roman small letter i with dot below	1ECB	ibo, vi
ł̣	Roman capital letter I with ogonek	012E	lt
ł̣	Roman small letter i with ogonek	012F	3, lt
İ̅	(Roman capital letter I with line below)	0049+0331	am_R1, el-R
ï̅	(Roman small letter i with line below)	0069+0331	am_R1, el-R
İ̅	(Roman capital letter I with line below and acute)	00CD+0331	el-R
ï̅	(Roman small letter i with line below and acute)	00ED+0331	el-R
ı	Roman small letter i with no dot	0131	az, tr
Ŭ	Roman capital ligature IJ	0132	nl
ij	Roman small ligature ij	0133	nl
Ĵ	(Roman capital letter J with line below)	004A+0331	as_R, bn_R
ĵ	(Roman small letter j with line below)	006A+0331	as_R, bn_R
Ķ	Roman capital letter K with caron	01E8	sámi2
ķ	Roman small letter k with caron	01E9	sámi2
Ķ̇	Roman capital letter K with dot below	1E32	pa_R, ur_R
ķ̇	Roman small letter k with dot below	1E33	pa_R, ur_R
Ķ̃	Roman capital letter K with cedilla	0136	lv
ķ̃	Roman small letter k with cedilla	0137	lv
Ķ̅	Roman capital letter K with line below	1E34	3
ķ̅	Roman small letter k with line below	1E35	3
Ķ̆	Roman capital letter K with hook	0198	ha
Ķ̆	Roman small letter k with hook	0199	ha
Ł̣	Roman capital letter L with middle dot	013F	ca
ł̣	Roman small letter l with middle dot	0140	ca
Ł̇	Roman capital letter L with acute	0139	sk
ł̇	Roman small letter l with acute	013A	sk
Ł̃	Roman capital letter L with caron	013D	sk
ł̃	Roman small letter l with caron	013E	sk
Ł̅	Roman capital letter L with dot below	1E36	gu_R, kn_R, ml_R, mr_R, or_R, ta_R, te_R
ł̅	Roman small letter l with dot below	1E37	gu_R, kn_R, ml_R, mr_R, or_R, ta_R, te_R

Ľ	(Roman capital letter L with diaeresis below)	004C+0324	as_R
ł	(Roman small letter l with diaeresis below)	006C+0324	as_R
Ļ	Roman capital letter L with cedilla	013B	livo, lv, mars
ļ	Roman small letter l with cedilla	013C	livo, lv, mars
Ľ	Roman capital letter L with circumflex below	1E3C	vend
ł	Roman small letter l with circumflex below	1E3D	vend
Ľ	(Roman capital letter L with breve below)	004C+032E	ta_R
ł	(Roman small letter l with breve below)	006C+032E	ta_R
Ľ	(Roman capital letter L with–line below)	1E3A	ml_R
ł	(Roman small letter l with–line below)	1E3B	ml_R
Ł	Roman capital letter L with stroke	0141	3, pl, sorbl, sorb2
ł	Roman small letter l with stroke	0142	3, pl, sorbl, sorb2
LJ	Roman capital letter LJ	01C7	sr_R
Lj	Roman capital letter L with small letter j	01C8	sr_R
lj	Roman small letter lj	01C9	sr_R
Ĺ	(Roman capital letter M with circumflex)	004D+0302	lb
ĺ	(Roman small letter m with circumflex)	006D+0302	lb
Ļ	(Roman capital letter M with macron)	004D+0304	bi
ļ	(Roman small letter m with macron)	006D+0304	bi
Ļ	Roman capital letter M with dot above	1E40	as_R, hi_R, kn_R, ml_R, or_R, te_R
ļ	Roman small letter m with dot above	1E41	as_R, hi_R, kn_R, ml_R, or_R, te_R
Ĺ	(Roman capital letter M with candrabindu)	004D+0310	or_R, te_R
ĺ	(Roman small letter m with candrabindu)	006D+0310	or_R, te_R
Ļ	(Roman capital letter M with cedilla)	004D+0327	mars
ļ	(Roman small letter m with cedilla)	006D+0327	mars
Ñ	Roman capital letter N with acute	0143	ne_R, pl, sorbl, sorb2
ñ	Roman small letter n with acute	0144	ne_R, pl, sorbl, sorb2
Ñ	(Roman capital letter N with circumflex)	004E+0302	lb
ñ	(Roman small letter n with circumflex)	006E+0302	lb
Ñ	Roman capital letter N with tilde	00D 1	as_R, bn_R, br, cham, es, eu, gl, gn, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, tk, tl
ñ	Roman small letter n with tilde	00F 1	as_R, bn_R, br, cham, es, eu, gl, gn, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, tk, tl
Ñ	(Roman capital letter N with macron)	004E+0304	mars
ñ	(Roman small letter n with macron)	006E+0304	mars
Ñ	Roman capital letter N with dot above	1E44	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, or_R, pa_R, ta_R, te_R, vend
ñ	Roman small letter n with dot above	1E45	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, or_R, pa_R, ta_R, te_R, vend
Ñ	(Roman capital letter N with	004E+0308	mg

	diaeresis)		
ñ	(Roman small letter n with diaeresis)	006E+0308	mg
Ñ	Roman capital letter N with caron	0147	cs, sk
ň	Roman small letter n with caron	0148	cs, sk
Ṅ	Roman capital letter N with dot below	1E46	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or-R, pa-R, ta_R, te_R
ṅ	Roman small letter n with dot below	1E47	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or-R, pa-R, ta_R, te_R
Ṇ	Roman capital letter N with cedilla	0145	livo, lv, mars
ṇ	Roman small letter n with cedilla	0146	livo, lv, mars
N̂	Roman capital Letter N with circumflex below	1E4A	vend
n̂	Roman small letter n with circumflex below	1E4B	vend
N̅	Roman capital letter N with line below	1E48	el_R, ps_R, ta_R
n̅	Roman small letter n with line below	1E49	el_R, ps_R, ta_R
Ŋ	Roman capital letter ENG (Sami)	014A	sámi1
ŋ	Roman small letter eng (Sami)	014B	sámi1
NJ	Roman capital letter NJ	01CA	sr_R
Nj	Roman capital letter N with small letter j	01CB	sr_R
nj	Roman small letter nj	01CC	sr_R
Ō	Roman capital letter O with grave	00D2	ca, gd, it, mt, no, pt, rm, sámí2, tl, vi, zh_R1
ò	Roman small letter o with grave	00F2	ca, gd, it, mt, no, pt, rm, sámí2, tl, vi, zh_R1
Ó	Roman capital letter O with acute	00D3	ca, cs, fo, el_R, es, ga, gd, gl, gn, hu, is, it, km_R, nl, pl, pt, sorb2, sk, tl, vi, zh_R1
ó	Roman small letter o with acute	00F3	ca, cs, fo, el_R, es, ga, gd, gl, gn, hu, is, it, km_R, nl, pl, pt, sorb2, sk, tl, vi, zh_R1
Ô	Roman capital letter O with circumflex	00D4	af, ar_R2, ar_R4, br, cy, fr, fy, ja_R2, km_R, lo_R, mg, my_R, nl, no, pt, sk, tn, vi
ô	Roman small letter o with circumflex	00F4	af, ar_R2, ar_R4, br, cy, fr, fy, ja_R2, km_R, lo_R, mg, my_R, nl, no, pt, sk, tn, vi
Û	Roman capital letter O with circumflex and grave	1ED2	vi
õ	Roman small letter o with circumflex and grave	1ED3	vi
Ố	Roman capital letter O with circumflex and acute	1ED0	vi
ó̂	Roman small letter o with circumflex and acute	1ED1	vi
Õ	Roman capital letter O with circumflex and tilde	1ED6	vi
õ̃	Roman small letter o with circumflex and tilde	1ED7	vi
Ö	Roman capital letter O with circumflex and hook above	1ED4	vi

ó	Roman small letter o with circumflex and hook above	1ED5	vi
Ô	Roman capital letter O with circumflex and dot below	1ED8	vi
ô	Roman small letter o with circumflex and dot below	1ED9	vi
Õ	Roman capital letter O with tilde	00D5	et, o, pt, vi
õ	Roman small letter o with tilde	00F5	et, gn, livo, pt, vi
Ō	(Roman capital letter O with tilde and macron)	00D5+0304	livo
ō	(Roman small letter o with tilde and macron)	00F5+0304	livo
Ȫ	Roman capital letter O with macron	014C	hawa, ja_R1, livo, mars, mi, zh_R1
ȫ	Roman small letter o with macron	014D	3, hawa, ja_R1, livo, mars, mi, zh_R1
ȴ	Roman small letter o with macron and acute	1E53	3
ȵ	(Roman small letter o with macron, acute and ogonek)	1E53+0328	3
Ȭ	Roman capital letter O with breve	014E	gu_R, hi_R, km_R, kn_R, ko_R1, ko_R2, ml_R, sámí2, tá_R, te_R
ȭ	Roman small letter o with breve	014F	gu_R, hi_R, km_R, kn_R, ko_R1, ko_R2, ml_R, sámí2, tá_R, te_R
Ȯ	(Roman capital letter O with dot above)	004F+0307	livo, ulit
ȯ	(Roman small letter o with dot above)	006F+0307	livo, ulit
Ȱ	(Roman capital letter O with dot above and macron)	004F+0307+0304	livo
ȱ	(Roman small letter o with dot above and macron)	006F+0307+0304	livo
Ö	Roman capital letter O with diaeresis	00D6	az, cy, de, et, fi, hu, is, kk_R, ky_R, lb, mn_R, nl, sv, tk, tk_R, tr, yape
ö	Roman small letter o with diaeresis	00F6	az, cy, de, et, fi, hu, is, kk_R, ky_R, lb, mn_R, nl, sv, tk, tk_R, tr, yape
Ȫ	Roman capital letter O with hook above	1ECE	vi
ȫ	Roman small letter o with hook above	1ECF	vi
Ȭ	Roman capital letter O with double acute	0150	hu
ȭ	Roman small letter o with double acute	0151	hu
Ȯ	Roman capital letter O with caron	01D1	zh_R1
ȯ	Roman small letter o with caron	01D2	zh_R1
Ȱ	Roman capital letter O with horn	01A0	vi
ȱ	Roman small letter o with horn	01A1	vi
Ȳ	Roman capital letter O with horn and grave	1EDC	vi
ȳ	Roman small letter o with horn and grave	1EDD	vi
ȴ	Roman capital letter O with horn and acute	1EDA	vi
ȵ	Roman small letter o with horn and acute	1EDB	vi

Ō	Roman capital letter O with horn and tilde	1EE0	vi
ō	Roman small letter o with horn and tilde	1EE1	vi
Ȫ	Roman capital letter O with horn and hook above	1EDE	vi
ȫ	Roman small letter o with horn and hook above	1EDF	vi
Ȭ	Roman capital letter O with horn and dot below	1EE2	vi
ȭ	Roman small letter o with horn and dot below	1EE3	vi
Ȯ	Roman capital letter O with dot below	1ECC	ibo, vi, yo
ȯ	Roman <i>small</i> letter o with dot below	1ECD	ibo, vi, yo
Ȱ	(Roman capital letter O with cedilla)	004F+0327	mars
ȱ	(Roman <i>small</i> letter o with cedilla)	006F+0327	mars
Ȳ	Roman capital letter O with ogonek	01EA	sámi2
ȳ	Roman <i>small</i> letter o with ogonek	01EB	sámi2
ȴ	Roman capital letter O with ogonek and macron	01EC	sámi2
ȵ	Roman small letter o with ogonek and macron	01ED	sámi2
ȶ	(Roman capital letter O with line below)	004F+0331	el-R
ȷ	(Roman small letter o with line below)	006F+0331	el-R
ȸ	(Roman capital letter O with line below and acute)	00D3+0331	el-R
ȹ	(Roman <i>small</i> letter o with line below and acute)	00F3+0331	el-R
Ⱥ	Roman capital letter O with stroke	00D8	da, fo, no
Ȼ	Roman small letter o with stroke	00F8	da, fo, no
ȼ	(Roman capital letter O with inverted apostrophe)		uz
Ƚ	(Roman small letter o with inverted apostrophe)		uz
Œ	Roman capital ligature OE	0152	fr, km_R
œ	Roman small ligature oe	0153	fr, km_R
Œ̆	(Roman capital ligature OE with breve)	0152+0306	km_R
œ̆	(Roman small ligature oe with breve)	0153+0306	km_R
ȼ̄	(Roman capital letter P with macron)	0050+0304	bi
Ƚ̄	(Roman small letter p with macron)	0070+0304	bi
Ŕ	Roman capital letter R with acute	0154	sorb1, sk
ŕ	Roman small letter r with acute	0155	sorb1, sk
Ŗ	(Roman capital letter R with dot below and macron)	1E5C	te_R
ŗ	(Roman small letter R with dot below and macron)	1E5D	te_R
Ř	Roman capital letter R with dot above	1E58	as_R, bn_R, hi_R, kn_R, or-R, pa-R, ur_R
ř	Roman <i>small</i> letter r with dot above	1E59	as_R, bn_R, hi_R, kn_R, or-R, pa-R, ur_R
Ř̆	Roman capital letter R with caron	0158	cs, sorb2
ř̆	Roman small letter r with caron	0159	a, sorb2

R	Roman capital letter R with dot below	1E5A	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, te_R
ṙ	Roman small letter r with dot below	1E5B	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, te_R
Ṛ	Roman capital letter R with cedilla	0156	livo
ṛ	Roman small letter r with cedilla	0157	livo
Ṙ	Roman capital letter R with line below	1E5E	ml_R, ps_R, ta_R
ṙ	Roman small letter r with line below	1E5F	ml_R, ps_R, ta_R
Ŝ	Roman capital letter S with acute	015A	ne_R, pl, sorbl
ŝ	Roman small letter s with acute	015B	ne_R, pl, sorbl
Š	(Roman capital letter S with macron)	0053+0304	fa_R, ps_R
š	(Roman small letter s with macron)	0073+0304	fa_R, ps_R
Ṣ̌	Roman capital letter S with caron	0160	bg_R1, be-R, bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámil, sorbl, sorb2, sk, sl, sr_R
ṣ̌	Roman small letter s with caron	0161	bg_R1, be_R, bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámil, sorbl, sorb2, sk, sl, sr_R
Ṛ̣	Roman capital letter S with dot below	1E62	ar_R2, ar_R4, as_R, bn_R, gu_R, hi-R, kn_R, ml_R, ne_R, or-R, ta_R, te_R, yo
ṛ̇	Roman small letter s with dot below	1E63	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or-R, ta_R, te_R, yo
Ṛ̣,	Roman capital letter S with comma below	0218	ro
ṛ̇,	Roman small letter s with comma below	0219	ro
Ṛ̣̣	Roman capital letter S with cedilla	015E	ar_R1, az, fa_R, ps_R, ro, tr
ṛ̣̇	Roman small letter s with cedilla	015F	ar_R1, az, fa-R, ps_R, ro, tr
Ṛ̣̣̣	(Roman capital letter S with line below)	0053+0331	ar_R3
ṛ̣̣̇	(Roman small letter S with line below)	0073+0331	ar_R3
\$	Dollar sign	0024	tk
¢	Cent sign	00A2	tk
Ṛ̣̣̣̣	(Roman capital letters SH with line below)	0053+0332 +0048+0332	ps_R
Ṛ̣̣̣̣	(Roman capital letter S and small letter h with line below)	0053+0332 +0068+0332	ps_R
ṛ̣̣̣̇	(Roman small letters sh with line below)	0073+0332 +0068+0332	ps_R
ß	Roman small letter sharp s (German)	00DF	de
Ṛ̣̣̣̣̣	(Roman capital letter T with diaeresis and dot below)	1E6C+0308	ml_R
ṛ̣̣̣̣̇	(Roman small letter t with diaeresis and dot below)	1E6D+0308 (1E97+0323)	ml_R
Ṛ̣̣̣̣̣̣	Roman capital letter T with caron	0164	cs, sk
ṛ̣̣̣̣̣̇	Roman small letter t with caron	0165	cs, sk
Ṛ̣̣̣̣̣̣̣	Roman capital letter T with dot below	1E6C	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, ur_R
ṛ̣̣̣̣̣̣̇	Roman small letter t with dot below	1E6D	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, ur_R
Ṛ̣̣̣̣̣̣̣̣	(Roman capital letter T with	0054+0324	wo

ẗ	diaeresis below) (Roman small letter t with diaeresis below)	0074+0324	wo
T̸	Roman capital letter T with comma below	021A	ro
ı	Roman small letter t with comma below	021B	ro
T̸	Roman capital letter T with cedilla	0162	ar_R1, livo ro
t̸	Roman small letter t with cedilla	0163	ar_R1, livo, ro
T̂	Roman capital letter T with circumflex below	1E70	vend
t̂	Roman small letter t with circumflex below	1E71	vend
T̄	Roman capital letter T with line below	1E6E	ar_R3, ps_R
t̄	Roman small letter t with line below	1E6F	ar_R3, ps_R
T̶	Roman capital letter T with stroke	0166	sámi1
t̶	Roman small letter t with stroke	0167	sámi1
Þ	Roman capital letter Thorn (Icelandic)	OOFE	is
þ	Roman small letter thorn (Icelandic)	OOFE	is
Û	Roman capital letter U with grave	00D9	br, cy, fr, gd, it, mt, rm, tl, vi, zh_R1
u̸	Roman small letter u with grave	00F9	br, cy, fr, gd, it, mt, rm, tl, vi, zh_R1
U̇	Roman capital letter U with acute	OODA	ca, cs, el_R, es, fo, fy, ga, gl, gn, hu, is, it, nl, pt, sk, tl, vi, zh_R1
ú	Roman small letter u with acute	OOFA	ca, cs, el_R, es, fo, fy, ga, gl, gn, hu, is, it, nl, pt, sk, tl, vi, zh_R1
Û	Roman capital letter U with circumflex	00DB	af, ar_R2, ar_R4, cy, fr, fy, ja_R2, nl, tl, tr
û	Roman small letter u with circumflex	00FB	af, ar_R2, ar_R4, cy, fr, fy, ja_R2, nl, tl, tr
Û	Roman capital letter U with tilde	0168	vi
ũ	Roman small letter u with tilde	0169	gn, vi
Ū	Roman capital letter U with macron	016A	2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hi_R, ja_R1, kk_R, kn_R, livo, lt, lv, mars, mi, ml_R, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
ū	Roman small letter u with macron	0168	2, 3, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hi_R, ja_R1, kk_R, kn_R, livo, lt, lv, mars, mi, ml_R, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
Ū	Roman capital letter U with breve	016C	bg_R2, km_R, ko_R1, ko_R2, tg_R, uz_R, zh_R2
ũ	Roman small letter u with breve	0160	bg_R2, km_R, ko_R1, ko_R2, tg_R, uz_R, zh_R2
Û	Roman capital letter U with diaeresis	00DC	az, bi, br, ca, cham, cy, de, et, es, eu, fr, gl, hu, kk_R, ky_R, lb, mn_R, pt, tk, tk_R, tr, zh_R2, zh_R1
ü	Roman small letter u with diaeresis	00FC	az, bi, br, ca, cham, cy, de, et, es, eu, fr, gl, hu, kk_R, ky_R, lb, mn_R, pt, tk, tk_R, tr, zh_R2, zh_R1
Û	Roman capital letter U with diaeresis and grave	01DB	zh_R1
ü	Roman small letter u with diaeresis and grave	01DC	zh_R1

Ú	Roman capital letter U with diaeresis and acute	01D7	zh_R1
ú	Roman small letter u with diaeresis and acute	01D8	zh_R1
Û	Roman capital letter U with diaeresis and macron	01D5	zh_R1
ü	Roman small letter u with diaeresis and macron	01D6	zh_R1
Ů	Roman capital letter U with diaeresis and caron	01D3	zh_R1
ů	Roman small letter u with diaeresis and caron	01D4	zh_R1
Ű	Roman capital letter U with hook above	1EE6	vi
ű	Roman small letter u with hook above	1EE7	vi
Û	Roman capital letter U with ring above	016E	cs
ú	Roman small letter u with ring above	016F	cs
Ů	Roman capital letter U with double acute	0170	hu
ű	Roman small letter u with double acute	0171	hu
Ů	Roman capital letter U with caron	01D3	zh_R1
ů	Roman small letter u with caron	01D4	zh_R1
U'	Roman capital letter U with horn	01AF	vi
u'	Roman small letter u with horn	01B0	vi
Ù	Roman capital letter U with horn and grave	1EEA	vi
ù	Roman small letter u with horn and grave	1EEB	vi
Ú	Roman capital letter U with horn and acute	1EE8	vi
ú	Roman small letter u with horn and acute	1EE9	vi
Û	Roman capital letter U with horn and tilde	1EEE	vi
ü	Roman small letter u with horn and tilde	1EEF	vi
Û	Roman capital letter U with horn and hook above	1EEC	vi
ű	Roman small letter u with horn and hook above	1EED	vi
U'	Roman capital letter U with horn and dot below	1EF0	vi
u'	Roman small letter U with horn and dot below	1EF1	vi
U	Roman capital letter U with dot below	1EE4	ibo, vi
u	Roman small letter u with dot below	1EE5	ibo, vi
Ů	Roman capital letter U with ogonek	0172	lt
ů	Roman small letter u with ogonek	0173	lt
U	(Roman capital letter U with line below)	0055+0331	be-R
u	(Roman small letter u with line below)	0075+0331	be-R
V	(Roman capital letter V with line below)	0056+0331	el-R

ŷ	(Roman small letter v with line below)	0076+0331	el-R
Ŵ	Roman capital letter W with circumflex	0174	chic, cy
ŵ	Roman small letter w with circumflex	0175	chic, cy
Ẁ	Roman capital letter W with diaeresis	1E84	cy
ẁ	Roman small letter w with diaeresis	1E85	cy
Ẃ	Roman capital letter W with dot below	1E88	ur_R
ẃ	Roman small letter w with dot below	1E89	ur_R
Ẅ	(Roman capital letter X with line below)	0058+0331	3
ẅ	(Roman small letter x with line below)	0078+0331	3
Ỳ	Roman capital letter Y with grave	1EF2	vi
ỳ	Roman small letter y with grave	1EM	vi
Ỳ	Roman capital letter Y with acute	00DD	cs, el-R, fo, is, sk , vi
ỳ	Roman small letter y with acute	00FD	cs, el-R, fo, is, sk , vi
Ỳ	Roman capital letter Y with circumflex	0176	cy
ỳ	Roman s mall letter y with circumflex	0177	cy
Ȳ	Roman capital letter Y with tilde	1EF8	vi
ȳ	Roman small letter y with tilde	1EF9	gn, vi
Ÿ	Roman capital letter Y with diaeresis	0178	cy, fr
ÿ	Roman small letter y with diaeresis	00FF	cy, fr, tk
Ỳ	Roman capital letter Y with hook above	1EF6	vi
ỳ	Roman small letter y with hook above	1EF7	vi
Ỳ	Roman capital letter Y with dot below	1EF4	or-R, ur_R, vi
ỳ	Roman small letter y with dot below	1EF5	or-R, ur_R, vi
¥	Yen sign	00A5	tk
Ž	Roman capital letter Z with acute	0179	pl, sorbl, sorb2
ž	Roman small letter z with acute	017A	pl, sorbl, sorb2
Ț	(Roman capital letter Z with macron)	005A+0304	fa-R, ps_R
ț	(Roman small letter z with macron)	007A+0304	fa-R, ps_R
Ț	Roman capital letter Z with dot above	017B	mt, pl
ț	Roman small letter z with dot above	017C	mt, pl
Ž	Roman capital letter Z with caron	017D	bg_R1, be-R , bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámil, sorbl, sorb2, sk , sl, sr_R
ž	Roman small letter z with caron	017E	bg_R1, be_R , bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámil, sorbl , sorb2, sk , sl, sr_R
Z	Roman capital letter Z with dot below	1E92	ar_R2, ar_R4, ta_R
z	Roman small letter z with dot below	1E93	ar_R2, ar_R4, ta_R
Ƶ	(Roman capital letter Z with cedilla)	005A+0327	ar_R1, fa-R, ps_R
ƶ	(Roman small letter z with cedilla)	007A+0327	ar_R1, fa-R, ps_R
Ẑ	Roman capital letter Z with line below	1E94	fa-R, he-R, ps_R
ẑ	Roman small letter z with line below	1E95	fa-R, he-R, ps_R
Ẓ	(Roman capital letters ZH with line below)	005A+0332 +0048+0332	ps_R
ẓ	(Roman capital letter Z and small letter h with line below)	005A+0332 +0068+0032	ps_R
Ẕ	(Roman small letters zh with line below)	007A+0332 +0068+0332	ps_R
£	Pound sign	00A3	tk
ƶ	(Turkmen small letter zh)		tk

Number of characters: 484

Part 2. Sorted by ISO/Unicode

<u>ISO/Unic.</u>	<u>Char.</u>	<u>Character Name</u>	<u>References to Tables of Characters</u>
0024	\$	Dollar sign	tk
0034		Quotation mark	ru_R1, ru_R2
0038	&	Ampersand	3
0041+0307	A	(Roman capital letter A with dot above)	ulit
0045+0331	E	(Roman capital letter E with line below)	am_R1
0046+0331	F	(Roman capital letter F with line below)	el_R
0047+0303	G	(Roman capital letter G with tilde)	tl
0047+0323	G	(Roman capital letter G with dot below)	pa_R, ur_R
0047+0331	G	(Roman capital letter G with line below)	3
0048+0324	H	(Roman capital letter H with diaeresis below)	wo
0048+0331	H	(Roman capital letter H with line below)	ar_R3, be_R, he_R
0049+0331	I	(Roman capital letter I with line below)	am_R1, el_R
004A+0331	J	(Roman capital letter J with line below)	as_R, bn_R
004C+0324	L	(Roman capital letter L with diaeresis below)	as_R
004C+032E	L	(Roman capital letter L with breve below)	ta_R
004D+0302	M	(Roman capital letter M with circumflex)	lb
004D+0304	M	(Roman capital letter M with macron)	bi
004D+0310	M̃	(Roman capital letter M with candrabindu)	or_R, te_R
004D+0327	M̂	(Roman capital letter M with cedilla)	mars
004E+0302	N	(Roman capital letter N with circumflex)	lb
004E+0304	N̄	(Roman capital letter N with macron)	mars
004E+0308	N̈	(Roman capital letter N with diaeresis)	mg
004F+0307	O	(Roman capital letter O with dot above)	livo, ulit
004F+0307 +0304	Ö	(Roman capital letter O with dot above and macron)	livo
004F+0327	Q	(Roman capital letter O with cedilla)	mars
004F+0331	Q	(Roman capital letter O with line below)	el_R
0050+0304	P̄	(Roman capital letter P with macron)	bi
0053+0304	S̄	(Roman capital letter S with macron)	fa_R, ps_R
0053+0331	S̅	(Roman capital letter S with line below)	ar_R3
0053+0332 +0048+0332	SH	(Roman capital letters SH with line below)	ps_R
0053+0332	Sh	(Roman capital letter S and small)	ps_R

+0068+0332	letter h with line below)	
0054+0324	T (Roman capital letter T with diaeresis below)	wo
0055+0331	U (Roman capital letter U with line below)	be_R
0056+0331	V (Roman capital letter V with line below)	el-R
0058+0331	X (Roman capital letter X with line below)	3
005A+0304	Z (Roman capital letter Z with macron)	fa_R, ps_R
005A+0327	Z (Roman capital letter Z with cedilla)	ar_R1, fa_R, ps_R
005A+0332	ZH (Roman capital letters ZH with line below)	ps_R
+0048+0332		
005A+0332	Zh (Roman capital letter Z and small letter h with line below)	ps_R
+0068+0032		
0061+0307	a (Roman small letter a with dot above)	ulit
0065+0331	e (Roman <i>small</i> letter e with line below)	am_R1
0066+0331	f (Roman small letter f with line below)	el-R
0067+0303	g (Roman <i>small</i> letter g with tilde)	gn, tl
0067+0323	g (Roman <i>small</i> letter g with dot below)	pa_R, ur_R
0067+0331	g (Roman small letter g with line below)	3
0068+0324	h (Roman small letter h with diaeresis below)	wo
0069+0331	i (Roman small letter i with line below)	am_R1, el_R
006A+0331	j (Roman small letter j with line below)	as_R, bn_R
006C+0324	l (Roman small letter l with diaeresis below)	as_R
006C+032E	l (Roman small letter l with breve below)	ta_R
006D+0302	m (Roman small letter m with circumflex)	lb
006D+0304	m (Roman small letter m with macron)	bi
006D+0310	m (Roman small letter m with candrabindu)	or-R, te_R
006D+0327	m (Roman small letter m with cedilla)	mars
006E+0302	n (Roman small letter n with circumflex)	lb
006E+0304	n (Roman small letter n with macron)	mars
006E+0308	n (Roman <i>small</i> letter n with diaeresis)	mg
006F+0307	o (Roman small letter o with dot above)	livo, ulit
006F+0307	o (Roman small letter o with dot above and macron)	livo
+0304		
006F+0327	o (Roman small letter o with cedilla)	mars
006F+0331	o (Roman <i>small</i> letter o with line below)	el_R
0070+0304	p (Roman <i>small</i> letter p with macron)	bi
0073+0304	s (Roman small letter s with macron)	fa_R, ps_R
0073+0331	s (Roman small letter s with line below)	ar_R3
0073+0332	sh (Roman small letters sh with line below)	ps_R
+0068+0332		
0074+0324	t (Roman small letter t with diaeresis below)	wo
0075+0331	u (Roman <i>small</i> letter u with line below)	be_R
0076+0331	v (Roman small letter v with line below)	el_R
0078+0331	x (Roman small letter x with line below)	3
007A+0304	z (Roman small letter z with macron)	fa_R, ps_R
007A+0327	z (Roman small letter z with cedilla)	ar_R1, fa_R, ps_R

007A+0332	zh	(Roman small letters zh with line below)	ps_R
+0068+0332			
00A2	¢	Cent sign	tk
00A3	£	Pound sign	tk
00A5	¥	Yen sign	tk
00B4	´	Acute accent	et
00C0	À	Roman capital letter A with grave	ca, ^{tl} , fr, it, mt, pt, rm, sámí2, tl, vi, zh_R1
00C1	Á	Roman capital letter A with acute	ar_R1, cs, cy, el-R, es, fa-R, fo, ga, gd, gl, gn, hu, is, km_R, mg, nl, pt, sámí1, sk, tl, vi, zh_R1
00C2	Â	Roman capital letter A with circumflex	ar_R2, ar_R4, br, cham, cy, fr, fy, km_R, nl, pt, ro, tl, tr, vi, wo
00C3	Ã	Roman capital letter A with tilde	pt, vi
00C4	Ä	Roman capital letter A with diaeresis	az, cy, de, et, fi, kk_R, lb, livo, sk, sv, tk, tk_R, yape
00C5	Å	Roman capital letter A with ring above	da, no, sv
00C6	Æ	Roman capital letter AE (ash)	da, fo, is, no
00C7	Ç	Roman capital letter C with cedilla	ar_R4, az, ca, eu, fr, sq, pt, tk, tr
00C8	Ē	Roman capital letter E with grave	af, ar_R4, be-R, ca, fr, gd, it, lo_R, mt, my_R, nl, pt, rm, ru_R1, sámí2, tl, vi, wo, zh_R1
00C9	É	Roman capital letter E with acute	af, ar_R4, bi, ca, cs, cy, da, el_R, es, fr, ga, gd, gl, gn, he_R, hu, id, is, it, km_R, lb, lo-R, nl, no, pt, rm, sk, sv, tl, vi, wo, zh_R1
00CA	Ê	Roman capital letter E with circumflex	af, ar_R4, br, cy, fr, fy, km_R, nl, ps_R, pt, tl, tn, vi, zh_R2, zh_R1
00CB	Ë	Roman capital letter E with diaeresis	af, be_R, cy, fr, lb, nl, ru_R1, ru_R2, sq, yape
00CC	Ï	Roman capital letter I with grave	gd, it, mt, rm, tl, vi, zh_R1
00CD	Í	Roman capital letter I with acute	ca, cs, el-R, es, fo, ga, gl, gn, hu, is, it, nl, pt, sk, tg_R, tl, vi, zh_R1
00CD+0331	Ĭ	(Roman capital letter I with line below and acute)	el-R
00CE	Î	Roman capital letter I with circumflex	af, ar_R2, ar_R4, cy, fy, fr, mt, rm, ro, tl, tr
00CF	Ï	Roman capital letter I with diaeresis	af, ar_R4, bi, ca, cy, el-R, fr, mn_R, nl, uk_R
00CF+0301	İ	(Roman capital letter I with diaeresis and acute)	el-R
00D0	Ð	Roman capital letter ETH	fo, is
00D1	Ñ	Roman capital letter N with tilde	as_R, bn_R, br, cham, es, eu, gl, gn, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, tk, tl
00D2	Ò	Roman capital letter O with grave	ca, gd, it, mt, no, pt, rm, sámí2, tl, vi, zh_R1
00D3	Ó	Roman capital letter O with acute	ca, cs, fo, el_R, es, ga, gd, gl, gn, hu, is, it, km_R, nl, pl, pt, sorb2, sk, tl, vi, zh_R1
00D3+0331	Ô	(Roman capital letter O with line below and acute)	el-R
00D4	Ö	Roman capital letter O with circumflex	af, ar_R2, ar_R4, br, cy, fr, fy,

			ja_R2, km_R, lo_R, mg, my_R, nl, no, pt, sk, tn, vi et, livo, pt, vi livo
00D5	Õ	Roman capital letter O with tilde	
00D5+0304	Ō	(Roman capital letter O with tilde and macron)	
00D6	O	Roman capital letter O with diaeresis	az, cy, de, et, fi, hu, is, kk_R, ky_R, lb, mn_R, nl, sv, tk, tk_R, tr, yape
00D8	O	Roman capital letter O with stroke	da, fo, no
00D9	U	Roman capital letter U with grave	br, cy, fr, gd, it, mt, rm, tl, vi, zh_R1
00DA	U	Roman capital letter U with acute	ca, cs, el_R, es, fo, fy, ga, gl, gn, hu, is, it, nl, pt, sk, tl, vi, zh_R1
00DB	Û	Roman capital letter U with circumflex	af, ar_R2, ar_R4, cy, fr, fy, ja_R2, nl, tl, tr
00DC	U	Roman capital letter U with diaeresis	az, bi, br, ca, cham, cy, de, et, es, eu, fr, gl, hu, kk_R, ky_R, lb, mn_R, pt, tk, tk_R, tr, zh_R2, zh_R1
00DD	Y	Roman capital letter Y with acute	cs, el-R, fo, is, sk, vi
00DE	Þ	Roman capital letter Thorn (Icelandic)	is
00DF	ß	Roman small letter sharp s (German)	de
00E0	à	Roman small letter a with grave	ca, gd, fr, it, mt, pt, rm, sámí2, tl, vi, zh_R1
00E1	á	Roman small letter a with acute	ar_R1, cs, cy, el-R, es, fa-R, fo, ga, gd, gl, gn, hu, is, km_R, mg, nl, pt, sámí1, sk, tl, vi, zh_R1
00E2	â	Roman small letter a with circumflex	ar_R2, ar_R4, br, cham, cy, fr, fy, km_R, nl, pt, ro, tl, tr, vi, wo
00E2+0328	â	(Roman small letter a with circumflex and ogonek)	3
00E3	ã	Roman small letter a with tilde	gn, pt, vi
00E4	ä	Roman small letter a with diaeresis	az, cy, de, et, fi, kk_R, lb, livo, sk, sv, tk, tk_R, yape
00E5	å	Roman small letter a with ring above	da, no, sv
00E6	æ	Roman small letter æ (ash)	da, fo, is, no
00E7	ç	Roman small letter c with cedilla	ar_R4, az, ca, eu, fr, sq, pt, tk, tr
00E8	è	Roman small letter e with grave	af, ar_R4, be-R, ca, fr, gd, it, lo_R, mt, my_R, nl, pt, rm, ru_R1, sámí2, tl, vi, wo, zh_R1
00E9	é	Roman small letter e with acute	af, ar_R4, bi, ca, cs, cy, da, el-R, es, fr, ga, gd, gl, gn, he_R, hu, id, is, it, km_R, lb, lo-R, nl, no, pt, rm, sk, sv, tl, wo, vi, zh_R1
00EA	ê	Roman small letter e with circumflex	af, ar_R4, br, cy, fr, fy, km_R, nl, ps_R, pt, tl, tn, vi, zh_R2, zh_R1
00EB	ë	Roman small letter e with diaeresis	af, be_R, cy, fr, lb, nl, ru_R1, ru_R2, sq, yape
00EC	ì	Roman small letter i with grave	gd, it, mt, rm, tl, vi, zh_R1
00ED	í	Roman small letter i with acute	ca, cs, el_R, es, fo, ga, gl, gn, hu, is, it, nl, pt, sk, tg_R, tl, vi, zh_R1
00ED+0331	ï	(Roman small letter i with <u>lira</u> below and acute)	el_R
00EE	î	Roman small letter i with circumflex	af, ar_R2, ar_R4, cy, fy, fr, mt, rm,

00EF	ı	Roman small letter i with diaeresis	ro, tl, tr af, ar_R4, bi, ca, cy, el-R, fr, mn_R, nl, uk_R el-R
00EF+0301	í	(Roman small letter i with diaeresis and acute)	
00F0	ð	Roman small letter eth	fo, is
00F1	ñ	Roman small letter n with tilde	as_R, bn_R, br, cham, es, eu, gl, gn, gu_R, hi_R, kn_R, ml_R, ne_R, or-R, pa_R, ta_R, te_R, tk, tl ca, gd, it, mt, no, pt, rm, sámí2, tl, vi, zh_R1
oom	ò	Roman small letter o with grave.	ca, cs, fo, el_R, es, ga, gd, gl, gn, hu, is, it, km_R, nl, pl, pt, sorb2, sk, tl, vi, zh_R1
00F3	ó	Roman small letter o with acute	el-R
00F3+0331	ó̇	(Roman small letter o with line below and acute)	
00F4	ô	Roman <i>small</i> letter o with circumflex	af, ar_R2, ar_R4, br, cy, fr, fy, ja_R2, km_R, lo_R, mg, my_R, nl, no, pt, sk, tn, vi et, gn, livo, pt, vi livo
00F5	õ	Roman small letter o with tilde	
00F5+0304	ȭ	(Roman small letter o with tilde and macron)	
00F6	ö	Roman small letter o with diaeresis	az, cy, de, et, fi, hu, is, kk_R, ky_R, lb, mn_R, nl, sv, tk, tk_R, tr, yape
00F8	ȯ	Roman small letter o with stroke	da, fo, no
00F9	ù	Roman small letter u with grave	br, cy, fr, gd, it, mt, rm, tl, vi, zh_R1
00FA	ú	Roman small letter u with acute	ca, cs, el_R, es, fo, fy, ga, gl, gn, hu, is, it, nl, pt, sk, tl, vi, zh_R1
00FB	û	Roman small letter u with circumflex	af, ar_R2, ar_R4, cy, fr, fy, ja_R2, nl, tl, tr
00FC	ü	Roman small letter u with diaeresis	az, bi, br, ca, cham, cy, de, et, es, eu, fr, gl, hu, kk_R, ky_R, lb, mn_R, pt, tk, tk_R, tr, zh_R2, zh_R1
00FD	ý	Roman small letter y with acute	cs, el-R, fo, is, sk, vi
00FE	þ	Roman <i>small</i> letter thorn (Icelandic)	is
00FF	ÿ	Roman small letter y with diaeresis	cy, fr, tk
0100	Ā	Roman capital letter A with macron	2, am_R2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hawa, hi_R, kn_R, livo, lv, mars, mi, ml_R, or-R, pa-R, ps_R, ta_R, te_R, ur_R, zh_R1
0100+0306	Ā̇	(Roman capital letter A with macron and breve)	hi-R
0101	ā	Roman <i>small</i> letter a with macron	2, 3, am_R2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hawa, hi_R, kn_R, livo, lv, mars, ml_R, mi, or-R, pa-R, ps_R, ta_R, te_R, ur_R, zh_R1
0101+0306	ā̇	(Roman <i>small</i> letter A with macron and breve)	hi_R
0102	Ȁ	Roman capital letter A with breve	km_R, ro, sámí2, vi, wo
0103	ȁ	Roman <i>small</i> letter a with breve	km_R, ro, sámí2, vi, wo
0104	Ą	Roman capital letter A with ogonek	lt, pl
0105	ą	Roman <i>small</i> letter a with ogonek	lt, pl

0106	Ć	Roman capital letter C with acute	bosn, hr, mk_R, pl, sorbl, sorb2, sr_R
0107	ć	Roman small letter c with acute	bosn, hr, mk_R, pl, sorbl, sorb2, sr_R
010A	Ć	Roman capital letter C with dot above	mt
010B	ċ	Roman small letter c with dot above	mt
010C	Č	Roman capital letter C with caron	bg_R1, be_R, bosn, cs, hr, lt, lv, mk_R, ru_R1, ru_R3, sámil, sorbl, sorb2, sk, sl, sr_R
010D	č	Roman small letter c with caron	bg_R1, be_R, bosn, cs, hr, lt, lv, mk_R, ru_R1, ru_R3, sámil, sorbl, sorb2, sk, sl, sr_R
010E	Ď	Roman capital letter D with caron	cs, sk
010F	ď	Roman small letter d with caron	cs, sk
0110	Đ	Roman capital letter D with stroke	bosn, hr, mk_R, sámil, sr_R, vi
0111	đ	Roman small letter d with stroke	bosn, hr, mk_R, sámil, sr_R, vi
0112	Ē	Roman capital letter E with macron	am_R2, livo, lv, mi, zh_R1
0113	ē	Roman <i>small</i> letter e with macron	3, am_R2, livo, lv, mi, zh_R1
Of 14	Ĕ	Roman capital letter E with breve	gu_R, hi_R, km_R, kn_R, ml_R, sámi2, ta_R, te_R
0115	ĕ	Roman small letter e with breve	gu_R, hi_R, km_R, kn_R, ml_R, sámi2, ta_R, te_R
0116	Ė	Roman capital letter E with dot above	kk_R, lt, tg_R, ulit
0117	ė	Roman small letter e with dot above	kk_R, lt, tg_R, ulit
0118	Ę	Roman capital letter E with ogonek	lt, pl
0119	ę	Roman small letter e with ogonek	lt, pl
011A	Ě	Roman capital letter E with caron	cs, sorbl, sorb2, zh_R1
011B	ě	Roman small letter e with caron	cs, sorbl, sorb2, zh_R1
011E	Ĝ	Roman capital letter G with breve	az, tr
011F	ğ	Roman small letter g with breve	az, tr
0120	Ġ	Roman capital letter G with dot above	mt
0121	ġ	Roman small letter g with dot above	mt
0122	Ģ	Roman capital letter G with cedilla	lv
0123	ģ	Roman small letter g with turned <i>comma</i> above	lv
0126	Ĥ	Roman capital letter H with stroke	mt
0127	ĥ	Roman small letter h with stroke	mt
0128	İ	Roman capital letter I with tilde	vi
0129	ĩ	Roman <i>small</i> letter i with tilde	gn, vi
012A	Ī	Roman capital letter I with macron	2, am_R2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hi_R, kk_R, kn_R, livo, lv, mi, ml_R, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
012B	ī	Roman small letter i with macron	2, am_R2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hi_R, kk_R, kn_R, livo, lv, mi, ml_R, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
012c	Ȫ	Roman capital letter I with breve	km_R, mn_R
012D	ȫ	Roman small letter i with breve	km_R, mn_R
012E	Į	Roman capital letter I with ogonek	lt
012F	į	Roman small letter i with ogonek	3, lt
0130	İ̇	Roman capital letter I with dot above	az, tr
0131	ı	Roman <i>small</i> letter i with no dot	az, tr
0132	Ŭ	Roman capital ligature U	nl

0133	ij	Roman small ligature ij	ni
0136	Ɔ	Roman capital letter K with cedilla	lv
0137	ƙ	Roman small letter k with cedilla	lv
0139	Ľ	Roman capital letter L with acute	sk
013A	ł	Roman small letter l with acute	sk
013B	Ł	Roman capital letter L with cedilla	livo, lv, mars
013C	ł̣	Roman small letter l with cedilla	livo, lv, mars
013D	Ľ	Roman capital letter L with caron	sk
013E	Ľ	Roman small letter l with caron	sk
013F	Ḷ̌	Roman capital letter L with middle dot	ca
0140	ł̣	Roman small letter l with middle dot	ca
0141	Ł	Roman capital letter L with stroke	3, pl, sorbl, sorb2
0142	ł	Roman small letter l with stroke	3, pl, sorbl, sorb2
0143	Ń	Roman capital letter N with acute	ne_R, pl, sorbl, sorb2
0144	ń	Roman small letter n with acute	ne_R, pl, sorbl, sorb2
0145	Ń	Roman capital letter N with cedilla	livo, lv, mars
0146	ņ	Roman small letter n with cedilla	livo, lv, mars
0147	Ń	Roman capital letter N with caron	cs, sk
0148	ň	Roman small letter n with caron	cs, sk
014A	Ŋ	Roman capital letter ENG (Sami)	sámi1
014B	ŋ	Roman small letter eng (Sami)	sámi1
014C	Ō	Roman capital letter O with macron	hawa, ja_R1, livo, mars, mi, zh_R1
014D	ō	Roman small letter o with macron	3, hawa, ja_R1, livo, mars, mi, zh_R1
014E	Ō	Roman capital letter O with breve	gu_R, hí_R, km_R, kn_R, ko_R1, ko_R2, ml_R, sámi2, ta_R, te_R
014F	o	Roman small letter o with breve	gu_R, hí_R, km_R, kn_R, ko_R1, ko_R2, ml_R, sámi2, ta_R, te_R
0150	Ŏ	Roman capital letter O with double acute	hu
0151	ô	Roman small letter o with double acute	hu
0152	Œ	Roman capital ligature OE	fr, km_R
0152+0306	Œ̣	(Roman capital ligature OE with breve)	km_R
0153	œ	Roman small ligature œ	fr, km_R
0153+0306	œ̣	(Roman small ligature œ with breve)	km_R
0154	Ɔ	Roman capital letter R with acute	sorbl, sk
0155	ƙ	Roman small letter r with acute	sorbl, sk
0156	Ŕ	Roman capital letter R with cedilla	livo
0157	ŕ	Roman small letter r with cedilla	livo
0158	Ř	Roman capital letter R with caron	cs, sorb2
0159	ř	Roman small letter r with caron	cs, sorb2
015A	Ś	Roman capital letter S with acute	ne_R, pl, sorbl
015B	ś	Roman small letter s with acute	ne_R, pl, sorbl
015E	Ş	Roman capital letter S with cedilla	ar_R1, az, fa-R, ps_R, ro, tr
015F	ş	Roman small letter s with cedilla	ar_R1, az, fa-R, ps_R, ro, tr
0160	Ŝ	Roman capital letter S with W o n	bg_R1, be_R, bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámi1, sorbl, sorb2, sk, sl, sr_R
0161	š	Roman small letter s with w o n	bg_R1, be_R, bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámi1, sorbl, sorb2, sk, sl, sr_R
0162	Ƨ	Roman capital letter T with cedilla	ar_R1, livo, ro
0163	Ƨ̣	Roman small letter t with cedilla	ar_R1, livo, M
0164	Ƨ̣	Roman capital letter T with caron	cs, sk
0165	Ƨ̣	Roman small letter t with w o n	cs, sk

0166	Ʀ	Roman capital letter T with stroke	sámi1
0167	Ƨ	Roman small letter t with stroke	sámi1
0168	Ū	Roman capital letter U with tilde	vi
0169	ũ	Roman small letter u with tilde	gn, vi
016A	Ū	Roman capital letter U with macron	2, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hi_R, ja_R1, kk_R, kn_R, livo, lt, lv, mars, mi, ml_R, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
016B	ū	Roman small letter u with macron	2, 3, ar_R1, ar_R3, as_R, bn_R, fa_R, gu_R, hi_R, ja_R1, kk_R, kn_R, livo, lt, lv, mars, mi, ml_R, or_R, pa_R, ps_R, ta_R, te_R, ur_R, zh_R1
016C	Ŭ	Roman capital letter U with breve	bg_R2, km_R, ko_R1, ko_R2, tg_R, uz_R, zh_R2
016D	ű	Roman small letter u with breve	bg_R2, km_R, ko_R1, ko_R2, tg_R, uz_R, zh_R2
016E	Ů	Roman capital letter U with ring above	cs
016F	ů	Roman small letter u with ring above	cs
0170	Ű	Roman capital letter U with double acute	hu
0171	ű	Roman small letter u with double acute	hu
0172	Ų	Roman capital letter U with ogonek	lt
0173	ų	Roman small letter u with ogonek	lt
0174	Ŵ	Roman capital letter W with circumflex	chic, cy
0175	ŵ	Roman small letter w with circumflex	chic, cy
0176	Ŷ	Roman capital letter Y with circumflex	cy
0177	ŷ	Roman small letter y with circumflex	cy
0178	ÿ	Roman capital letter Y with diaeresis	cy, fr
0179	Ž	Roman capital letter Z with acute	pl, sorbl, sorb2
017A	ž	Roman small letter z with acute	pl, sorbl, sorb2
017B	Ẑ	Roman capital letter Z with dot above	mt, pl
017C	ẑ	Roman small letter z with dot above	mt, pl
017D	Ž	Roman capital letter Z with caron	bg_R1, be-R, bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámi1, sorbl, sorb2, sk, sl, sr_R
017E	ž	Roman small letter z with caron	bg_R1, be-R, bosn, cs, et, hr, livo, lt, lv, mk_R, ru_R1, ru_R3, sámi1, sorbl, sorb2, sk, sl, sr_R
0181	Ɓ	Roman capital letter B with hook	ha
018A	Ɗ	Roman capital letter D with hook	ha
018F	Ɔ	Roman capital letter schwa	az
0198	Ƙ	Roman capital letter K with hook	ha
0199	ƙ	Roman small letter k with hook	ha
01A0	Ɔ	Roman capital letter O with horn	vi
01A1	o'	Roman small letter o with horn	vi
01AF	Ɔ	Roman capital letter U with horn	vi
01B0	u'	Roman small letter u with horn	vi
01B7	Ɔ	Roman capital letter EZH	sámi2
01C7	Ɔ	Roman capital letter LJ	sr_R
01C8	Ɔ	Roman capital letter L with small letter j	sr_R
01C9	Ɔ	Roman small letter lj	sr_R
01CA	Ɔ	Roman capital letter NJ	sr_R
01CB	Ɔ	Roman capital letter N with small	sr_R

		letter j	
01CC	n̄j	Roman small letter n̄j	sr_R
01CD	Ā	Roman capital letter A with caron	bg_R1, zh_R1
01CE	ā	Roman small letter a with caron	bg_R1, zh_R1
01CF	Ī	Roman capital letter I with caron	zh_R1
01D0	ī	Roman small letter i with caron	zh_R1
01D1	Ō	Roman capital letter O with caron	zh_R1
01D2	ō	Roman small letter o with caron	zh_R1
01D3	Ū	Roman capital letter U with diaeresis and caron	zh_R1
01D3	Ů	Roman capital letter U with caron	zh_R1
01D4	ů	Roman small letter u with diaeresis and caron	zh_R1
01D4	ü	Roman <i>small</i> letter u with caron	zh_R1
01D5	Û	Roman capital letter U with diaeresis and macron	zh_R1
01D6	ū	Roman small letter u with diaeresis and macron	zh_R1
01D7	Ú	Roman capital letter U with diaeresis and acute	zh_R1
01D8	ú	Roman small letter u with diaeresis and acute	zh_R1
01DB	Û	Roman capital letter U with diaeresis and grave	zh_R1
01DC	ù	Roman small letter u with diaeresis and grave	zh_R1
01DD	ə	Roman <i>small</i> letter schwa	az
01DE	Ā	Roman capital letter A with diaeresis and macron	livo, sámi2
01DF	ā	Roman small letter a with diaeresis and macron	livo, sámi2
01E0	Ā	Roman capital letter A with dot above and macron	sámi2
01E1	ā	Roman small letter a with dot above and macron	sámi2
01E2	Æ	Roman capital letter AE (ash) with macron	sámi2
01E3	æ	Roman small letter ae (ash) with macron)	sámi2
01E4	Ġ	Roman capital letter G with stroke	sámi2
01E5	ġ	Roman small letter g with stroke	sámi2
01E6	Ĝ	Roman capital letter G with caron	sámi2
01E7	ĝ	Roman <i>small</i> letter g with caron	sámi2
01E8	Ķ	Roman capital letter K with caron	sámi2
01E9	ķ	Roman <i>small</i> letter k with caron	sámi2
01EA	Ō	Roman capital letter O with ogonek	sámi2
01EB	ȯ	Roman <i>small</i> letter o with ogonek	sámi2
01EC	Ō̄	Roman capital letter O with ogonek and macron	sámi2
01ED	ȱ	Roman small letter o with ogonek and macron	sámi2
01EE	Ǻ	Roman capital letter EZH with caron	sámi2
01EF	ǻ	Roman small letter ezh with caron	sámi2
0218	Ŧ	Roman capital letter S with comma	ro

		below	
0219	ſ	Roman small letter s with comma below	ro
021A	Ŧ	Roman capital letter T with comma below	ro
021B	ŧ	Roman small letter t with comma below	ro
0253	ḅ	Roman small letter b with hook	ha
0257	ḏ	Roman small letter d with hook	ha
0259	ə	Roman s mall letter schwa	az
0292	ʒ	Roman small letter ezh	sámi2
02BB	‘	Inverted apostrophe	am_R2, ar_R1, fa_R, hawa, he_R, ps_R, sm, to
02BD	’	Reversed apostrophe	ar_R2
1E0C	Ḑ	Roman capital letter D with dot below	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, ur_R
1E0D	ḑ	Roman s mall letter d with dot below	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, ur_R
1E0E	Ḍ	Roman capital letter D with line below	ar_R3, ps_R
1E0F	ḍ	Roman small letter d with line below	ar_R3, ps_R
1E10	Ḑ̃	Roman capital letter D with cedilla	ar_R1, livo
1E11	ḑ̃	Roman small letter d with cedilla	ar_R1, livo
1E12	Ḑ̂	Roman capital letter D with circumflex below	vend
1E13	ḑ̂	Roman small letter d with circumflex below	vend
1E24	Ḑ̇	Roman capital letter H with dot below	ar_R2, ar_R4, as_R, bn_R, hi_R, kn_R, ml_R, or_R, te_R
1E25	ḑ̇	Roman small letter h with dot below	ar_R2, ar_R4, as_R, bn_R, hi_R, kn_R, ml_R, or_R, te_R
1E28	Ḑ̃̄	Roman capital letter H with cedilla	ar_R1, fa_R, ps_R
1E29	ḑ̃̄	Roman small letter h with cedilla	ar_R1, fa_R, ps_R
1E32	Ḑ̇̄	Roman capital letter K with dot below	pa_R, ur_R
1E33	ḑ̇̄	Roman small letter k with dot below	pa_R, ur_R
1E34	Ḑ̇̄̄	Roman capital letter K with line below	3
1E35	ḑ̇̄̄	Roman s mall letter k with line below	3
1E36	Ḑ̇̄̄̄	Roman capital letter L with dot below	gu_R, kn_R, ml_R, mr_R, or_R, ta_R, te_R
1E37	ḑ̇̄̄̄	Roman small letter l with dot below	gu_R, kn_R, ml_R, mr_R, or_R, ta_R, te_R
1E3A	Ḑ̇̄̄̄̄	(Roman capital letter L with–line below)	ml_R
1E3B	ḑ̇̄̄̄̄	(Roman s mall letter l with–line below)	ml_R
1E3C	Ḑ̂̄̄̄̄	Roman capital letter L with circumflex below	vend
1E3D	ḑ̂̄̄̄̄	Roman s mall letter l with circumflex below	vend
1E40	Ḑ̇̄̄̄̄̄	Roman capital letter M with <i>dot</i> above	as_R, hi_R, kn_R, ml_R, or_R, te_R
1E41	ḑ̇̄̄̄̄̄	Roman s mall letter m with <i>dot</i> above	as_R, hi_R, kn_R, ml_R, or_R, te_R
1E44	Ḑ̇̄̄̄̄̄̄	Roman capital letter N with <i>dot</i> above	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, or_R, pa_R, ta_R, te_R, vend
1E45	ḑ̇̄̄̄̄̄̄	Roman s mall letter n with <i>dot</i> above	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, or_R, pa_R, ta_R, te_R, vend
1E46	Ḑ̇̄̄̄̄̄̄̄	Roman capital letter N with <i>dot</i> below	as_R, bn_R, gu_R, hi_R, kn_R, ml_R,

1E X7	n	Roman small letter n with dot below	ne_R, or_R, pa_R, ta_R, te_R as_R, bn_R, gu_R, hi_R, kn_R, ml_R,
1E48	N	Roman capital letter N with line below	ne_R, or_R, pa_R, ta_R, te_R
1E49	n	Roman small letter n with line below	el_R, ps_R, ta_R
1E XA	N	Roman capital letter N with circumflex below	el_R, ps_R, ta_R
1E4B	n	Roman small letter n with circumflex below	vend
1E53	ó	Roman small letter o with macron and acute	3
1E53+0328	ô	(Roman small letter o with macron, acute and ogonek)	3
1E58	R	Roman capital letter R with dot above	as_R, bn_R, hi_R, kn_R, or_R, pa_R, ur_R
1E59	ř	Roman small letter r with dot above	as_R, bn_R, hi_R, kn_R, or_R, pa_R, ur_R
1E5A	R	Roman capital letter R with dot below	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, te_R
1E5B	r	Roman small letter r with dot below	as_R, bn_R, gu_R, hi_R, kn_R, ml_R, te_R
1E5C	Ř	(Roman capital letter R with dot below and macron)	te_R
1E5D	r	(Roman small letter R with dot below and macron)	te_R
1E5E	R	Roman capital letter R with line below	ml_R, ps_R, ta_R
1E5F	r	Roman small letter r with line below	ml_R, ps_R, ta_R
1E62	S	Roman capital letter S with dot below	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, ta_R, te_R, yo
1E63	s	Roman small letter s with dot below	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, ta_R, te_R, yo
1E6C	T	Roman capital letter T with dot below	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, ur_R
1E6C+0308	T	(Roman capital letter T with diaeresis and dot below)	ml_R
1E6D	t	Roman small letter t with dot below	ar_R2, ar_R4, as_R, bn_R, gu_R, hi_R, kn_R, ml_R, ne_R, or_R, pa_R, ta_R, te_R, ur_R
1E6D+0308	t	(Roman small letter t with diaeresis and dot below)	ml_R
1E6E	T	Roman capital letter T with line below	ar_R3, ps_R
1E6F	t	Roman small letter t with line below	ar_R3, ps_R
1E70	T	Roman capital letter T with circumflex below	vend
1E71	t	Roman small letter t with circumflex below	vend
1E84	W	Roman capital letter W with diaeresis	cy
1E85	w	Roman small letter w with diaeresis	cy
1E88	W	Roman capital letter W with dot below	ur_R
1E89	w	Roman small letter w with dot below	ur_R
1E92	Z	Roman capital letter Z with dot below	ar_R2, ar_R4, ta_R

1E93	ẓ	Roman small letter z with dot below	ar_R2, ar_R4, ta_R
1E94	Ẓ	Roman capital letter Z with line below	fa-R, he-R, ps_R
1E95	Z̄	Roman small letter z with line below	fa-R, he-R, ps_R
1E96	h̄	Roman small letter h with line below	ar_R3, be_R, he_R
1E97+0323	ṭ	(Roman small letter t with diaeresis and dot below)	ml_R
1EAO	À	Roman capital letter A with dot below	vi
1EA1	à	Roman small letter a with dot below	vi
1EA2	Ā	Roman capital letter A with hook above	vi
1EA3	ā	Roman small letter a with hook above	vi
1EA4	Ả	Roman capital letter A with circumflex and acute	vi
1EA5	ả	Roman small letter A with circumflex and acute	vi
1EA6	Ã	Roman capital letter A with circumflex and grave	vi
1EA7	ã	Roman small letter a with circumflex and grave	vi
1EA8	Ȧ	Roman capital letter A with circumflex and hook above	vi
1EA9	ȧ	Roman small letter a with circumflex and hook above	vi
1EAA	Ã	Roman capital letter A with circumflex and tilde	vi
1EAB	ã	Roman small letter a with circumflex and tilde	vi
1EAC	Â	Roman capital letter A with circumflex and dot below	vi
1EAD	â	Roman small letter a with circumflex and dot below	vi
1EAE	Ȧ	Roman capital letter A with breve and acute	vi
1EAF	ȧ	Roman small letter a with breve and acute	vi
1EBO	Ȧ	Roman capital letter A with breve and grave	vi
1EB1	ȧ	Roman small letter a with breve and grave	M
1EB2	Ȧ	Roman capital letter A with breve and hook above	vi
1EB3	ȧ	Roman small letter a with breve and hook above	vi
1EB4	Ȧ	Roman capital letter A with breve and tilde	vi
1EB5	ȧ	Roman small letter a with breve and tilde	vi
1EB6	Ȧ	Roman capital letter A with breve and dot below	vi
1EB7	ȧ	Roman small letter a with breve and dot below	vi

1EB8	Ē	Roman capital letter E with dot below	vi, yo
1EB9	ē	Roman small letter e with dot below	vi, yo
1EBA	Ē	Roman capital letter E with hook above	vi
1EBB	è	Roman small letter e with hook above	vi
1EBC	Ē	Roman capital letter E with tilde	vi
1EBD	ē	Roman small letter e with tilde	gn, vi
1EBE	É	Roman capital letter E with circumflex and acute	vi
1EBF	é	Roman small letter e with circumflex and acute	vi
1ECO	È	Roman capital letter E with circumflex and grave	vi
1EC1	è	Roman small letter e with circumflex and grave	vi
1EC2	Ē	Roman capital letter E with circumflex and hook above	vi
1EC3	é	Roman small letter e with circumflex and hook above	vi
1EC4	Ē	Roman capital letter E with circumflex and tilde	vi
1EC5	ē	Roman small letter e with circumflex and tilde	vi
1EC6	E	Roman capital letter E with circumflex and dot below	vi
1EC7	é	Roman small letter e with circumflex and dot below	vi
1EC8	ĭ	Roman capital letter I with hook above	vi
1EC9	ĭ	Roman small letter i with hook above	vi
1ECA	I	Roman capital letter I with dot below	ibo, vi
1ECB	ı	Roman small letter i with dot below	ibo, vi
1ECC	Ō	Roman capital letter O with dot below	ibo, vi, yo
1ECD	o	Roman small letter o with dot below	ibo, vi, yo
1ECE	Ō	Roman capital letter O with hook above	vi
1ECF	o	Roman small letter o with hook above	vi
1ED0	Ŏ	Roman capital letter O with circumflex and acute	vi
1ED1	ó	Roman small letter o with circumflex and acute	vi
1ED2	Ò	Roman capital letter O with circumflex and grave	vi
1ED3	o	Roman small letter o with circumflex and grave	vi
1ED4	Ŏ	Roman capital letter O with circumflex and hook above	vi
1ED5	ó	Roman small letter o with circumflex and hook above	vi
1ED6	Ŏ	Roman capital letter O with circumflex and tilde	vi
1ED7	õ	Roman small letter o with circumflex	vi

		and tilde	
1ED8	Ô	Roman capital letter O with circumflex and dot below	vi
1ED9	ô	Roman small letter o with circumflex and dot below	vi
1EDA	Ó	Roman capital letter O with horn and acute	vi
1EDB	ó	Roman small letter o with horn and acute	vi
1EDC	Ò	Roman capital letter O with horn and grave	vi
1EDD	ò	Roman small letter o with horn and grave	vi
1EDE	Û	Roman capital letter O with horn and hook above	vi
1EDF	û	Roman small letter o with horn and hook above	vi
1EE0	Õ	Roman capital letter O with horn and tilde	vi
1EE1	õ	Roman small letter o with horn and tilde	vi
1EE2	Ȯ	Roman capital letter O with horn and dot below	vi
1EE3	ȯ	Roman small letter o with horn and dot below	vi
1EE4	Ȯ	Roman capital letter U with dot below	ibo, vi
1EE5	ȯ	Roman small letter u with dot below	ibo, vi
1EE6	Û	Roman capital letter U with hook above	vi
1EE7	û	Roman small letter u with hook above	vi
1EE8	Ū	Roman capital letter U with horn and acute	vi
1EE9	ū	Roman small letter u with horn and acute	vi
1EEA	Ū	Roman capital letter U with horn and grave	vi
1EEB	ū	Roman small letter u with horn and grave	vi
1EEC	Û	Roman capital letter U with horn and hook above	vi
1EED	û	Roman small letter u with horn and hook above	vi
1EEE	Ū	Roman capital letter U with horn and tilde	vi
1EEF	ū	Roman small letter u with horn and tilde	vi
1EF0	Ȯ	Roman capital letter U with horn and dot below	vi
1EF1	ȯ	Roman small letter U with horn and dot below	vi
1EF2	ÿ	Roman capital letter Y with grave	vi
1EF3	ÿ	Roman small letter y with grave	vi

1EF4	Ỳ	Roman capital letter Y with dot below	or-R, ur_R, vi
1EF5	ỳ	Roman small letter y with dot below	or-R, ur_R, vi
1EF6	Ỳ̇	Roman capital letter Y with hook above	vi
1EF7	ỳ̇	Roman small letter y with hook above	vi
1EF8	Ȳ	Roman capital letter Y with tilde	vi
1EF9	ȳ	Roman small letter y with tilde	gn, vi
2018		Inverted apostrophe	am _R2, ar _R1, fa-R, hawa, he-R, ps_R, sm , to
201B	,	Reversed apostrophe	
	Ă	(Roman capital letters AI with breve)	hi-R
	ă	(Roman small letters ai with breve)	hi-R
	Ɔ	(Roman capital letter G with inverted apostrophe)	uz
	Ɔ'	(Roman small letter g with inverted apostrophe)	uz
	Ɔ'	(Roman capital letter O with inverted apostrophe)	uz
	o'	(Roman small letter o with inverted apostrophe)	uz
	ƶ	(Turkmen small letter zh)	tk

List of Countries, Languages and Writing Systems

Key to Columns below:

Entity Name (Column 1) - Country names in English as given in the United Nations *Terminology Bulletin No. 347* (1995 + Corrigenda) and the list of country names prepared by the UNGEGN Working Group on Country Names. Member countries of the United Nations and members of the UN specialized agencies are included.

Language (Column 2) - languages in which geographical names of a particular country are likely to occur are listed here, regardless of the status of the language. Where available, guidance has been sought from the toponymic guidelines issued by the individual countries. Languages are arranged as per in the list of country names compiled by the UNGEGN Working Group on Country Names, then all other languages.

Writing System(s) (Column 3) - writing systems used for particular languages are listed. All Roman-script languages are assumed to use all or part of the Basic Roman alphabet (BR) which is given in part 2. Extensions to that alphabet including characters with diacritical marks and other variations are given in separate tables for each of the languages

Romanization System(s) (Column 4) - lists romanization systems encountered in toponymy for non-Roman script names. Order for systems: UN, national, and other. n/a = not applicable. Where a UN-approved system exists, this is given first, quoting the appropriate Conference resolution in brackets.

Table Reference (Column 5) - refers to tables listing characters of particular languages in Part 2. Table references contain the abbreviation of the language name according to ISO 639 (1988). If such an abbreviation is not available, the first four letters of the English name of the language are used instead. References of tables listing romanized characters contain additionally a letter R and a number if multiple romanization systems are listed. In case of non-Roman scripts reference to the ISO 10646/Unicode encoding of the script is made. An asterisk denotes that the allocation of codes is provisional pending final adoption of these by both ISO and the Unicode consortium.

Entity Name	Language	Writing System(s)	Romanization System(s)	Table Ref. and Notes
Afghanistan	Dari	Perso-Arabic		0600-06FF
			UN 1967 (I/13)	fa-R
	Pashto	Perso-Arabic + Pashto extensions		0600-06FF
			BGN/PCGN 1968	ps_R
Albania	Albanian	BR + Albanian extensions	n/a	sq
Algeria	Arabic	Perso-Arabic		0600-06FF
			a) UN 1972 (II/8)	ar_R1
			b) I.G.N System 1973	ar_R4
	French	BR + French extensions	n/a	fr
Andorra	Catalan	BR + Catalan extensions	n/a	ca
	French	BR + French extensions	n/a	fr

	Spanish	BR + Spanish extensions	n/a	es
Angola	Portuguese	BR + Portuguese extensions	n/a	pt
Antigua and Barbuda	English	Basic Roman	n/a	1
Argentina	Spanish	BR + Spanish extensions	n/a	es
Armenia	Armenian	Armenian	‡	0530-058F
			BGN/PCGN 1981	1
Australia	English	Basic Roman	n/a	1
Austria	German	BR + German extensions	n/a	de
	Slovene	BR + Slovene extensions	n/a	sl
Azerbaijan	Azerbaijani	BR + Azerbaijani extensions	n/a	az
		Cyrillic + Azerbaijani extensions		0400-04FF
			BGN/PCGN 1993 Agreement	az
Bahamas (the)	English	Basic Roman	n/a	1
Bahrain	Arabic	Perso-Arabic		0600-06FF
			UN 1972(II/8)	ar_R 1
Bangladesh	Bengali	Bengali		0980-09FF
			a) UN 1972(II/11), 1977 (III/12)	bn_R
			b) National: Hunterian System	2
Barbados	English	Basic Roman	n/a	1
Belarus	Belarusian	Cyrillic + Belarusian extensions		0400-04FF
			a) National, unofficial (romanized maps 1992/93)	be_R
			b) BGN/PCGN 1979	1
	Russian	Cyrillic	<i>see Russian Federation (the)</i>	
Belgium	Dutch	BR + Dutch extensions	n/a	nl
	French	BR + French extensions	n/a	fr

	German	BR + German extensions	n/a	de
Belize	English	Basic Roman	n/a	1
	Spanish	BR + Spanish extensions	n/a	es
Benin	French	BR + French extensions	n/a	fr
Bhutan	Dzongkha	Tibetan		0F00-0FBF
Bolivia	Spanish	BR + Spanish extensions	n/a	es
	Aymara	Basic Roman	n/a	1
	Quechua	Basic Roman	n/a	1
Bosnia and Herzegovina	Bosnian	BR + Bosnian extensions	n/a	bosn
	Croatian	BR + Croatian extensions	n/a	hr
	Serbian	Cyrillic		0400-04FF
			UN 1977 (III/11)	sr_R
Botswana	English	Basic Roman	n/a	1
	Tswana	BR + Tswana extensions	n/a	tn
Brazil	Portuguese	BR + Portuguese extensions	n/a	pt
Brunei Darussalam	English	Basic Roman	n/a	1
	Malay	Basic Roman	n/a	1
Bulgaria	Bulgarian	Cyrillic		0400-04FF
			a) UN 1977 (III/10)	bg_R1
			b) BGN/PCGN 1952	bg_R2
Burkina Faso	French	BR + French extensions	da	fr
Burundi	French	BR + French extensions	n/a	
	Kirundi	Basic Roman	n/a	
Cambodia	Khmer	Khmer		1780-17FF*
			UN 1972 (II/10)	km_R
Cameroon	English	Basic Roman	n/a	1
	French	BR + French extensions	n/a	fr
Canada	English	Basic Roman	n/a	1
	French	BR + French extensions	n/a	fr

	<i>(various indigenous languages)</i> ¹	BR + Canadian extensions	n/a	3
Cape Verde	Portuguese	BR + Portuguese extensions	n/a	pt
Central African Republic (the)	French	BR + French extensions	n/a	fr
Chad	Arabic	Perso-Arabic		0600-06FF
			UN 1972(II/8)	ar_R1
	French	BR + French extensions	n/a	fr
Chile	Spanish	BR + Spanish extensions	n/a	es
<i>china</i>	Chinese	Chinese		Catered for
			a) UN 1977(III/8)	zh_R1
			b) Modified Wade-Giles (1912)	zh_R2
Colombia	Spanish	BR + Spanish extensions	n/a	es
Comoros (the)	Arabic	Perso-Arabic		0600-06FF
			UN 1972(II/8)	ar_R1
	French	BR + French extensions	n/a	fr
Congo (the)	French	BR + French extensions	n/a	fr
Congo, Democratic Republic of the	French	BR + French extensions	n/a	fr
costa Rica	Spanish	BR + Spanish extensions	n/a	es
Côte d'Ivoire	French	BR + French extensions	n/a	fr
Croatia	Croatian	BR + Croatian extensions	n/a	hr
Cuba	Spanish	BR + Spanish extensions	n/a	es
Cyprus	Greek	Greek		0370-03FF, 0F00-1F7F
			UN 1987(V/19)	el-R
	Turkish	BR + Turkish extensions	n/a	tr

¹ - includes Chipewyan, Inuktitut, Kaska, Northern Tutchone, Tlingit which use extended Roman alphabet characters, the inventory of which is not complete at the present stage of research

Czech Republic (the)	Czech	BR + Czech extensions	n/a	cs
Democratic People's Republic of Korea (the)	Korean	Hangul		1100-11FF
			McCune Reischauer 1939	ko_R2
Denmark	Danish	BR + Danish extensions	n/a	da
	Faroese	BR + Faroese extensions	n/a	fo
	German	BR + German extensions	n/a	de
	Greenlandic	Basic Roman	n/a	l
Djibouti	Arabic	Perso-Arabic		0600-06FF
	French	BR + French extensions	n/a	fr
Dominica	English	Basic Roman	n/a	l
Dominican Republic (the)	Spanish	BR + Spanish extensions	n/a	es
Ecuador	Spanish	BR + Spanish extensions	n/a	es
Egypt	Arabic	Perso-Arabic		0600-06FF
			a) UN 1972 (II/8)	ar_R1
			b) National: Survey of Egypt	ar_R2
El Salvador	Spanish	BR + Spanish extensions	n/a	es
Equatorial Guinea	Spanish	BR + Spanish extensions	n/a	es
Eritrea	Tigrinya	Ethiopic		1200-137F
			BGN/PCGN 1994	
Estonia	Estonian	BR + Estonian extensions	n/a	et
	Swedish	BR + Swedish extensions	n/a	sv
Ethiopia	Amharic	Ethiopic		1200-137F
			a) UN 1967 (I/17)	am_R1
			b) BGN/PCGN 1967	am_R2
Fiji	English	Basic Roman	n/a	l
	Fijian	Basic Roman	n/a	l
Finland	Finnish	BR + Finnish extensions	n/a	fi

	Swedish	BR + Swedish extensions		sv
	Sámi (Northern, Inari)	BR + Sámi extensions	n/a n/a	sámi1
	Sámi (Skolt)	BR + Skolt Sámi extensions	n/a	sámi2
France	French	BR + French extensions	n/a _i	fr
	Breton	BR + Breton extensions	n/a	br
	Occitan	BR + French extensions	n/a	fr
Gabon	French	BR + French extensions	n/a	fr
Gambia (the)	English	Basic Roman	n/a	1
Georgia	Georgian	Mkhedruli		10A0-10FF
			BGN/PCGN 1981	1
Germany	German	BR + German extensions	n/a	de
	Sorbian (Lower)	BR + Sorbian extensions 1	n/a	sorb1
	Sorbian (Upper)	BR + Sorbian extensions 2	n/a	sorb2
Ghana	English	Basic Roman	n/a	1
Greece	Greek	Greek		0370-03FF, 1F00-1FFF
			UN 1987(V/19)	el_R
Grenada	English	Basic Roman	n/a	1
Guatemala	Spanish	BR + Spanish extensions	n/a	es
Guinea	French	BR + French extensions	n/a	fr
Guinea-Bissau	Portuguese	BR + Portuguese extensions	n/a	pt
Guyana	English	Basic Roman	n/a	1
Haiti	French	BR + French extensions	n/a	fr
	Haitian Creole	BR + French extensions	n/a	fr
Holy See (the)	Italian	BR + Italian extensions	n/a	it
	Latin	Basic Roman	n/a	1
Honduras	Spanish	BR + Spanish extensions	n/a	es

Hungary	Hungarian	BR + Hungarian extensions	n/a	hu
Iceland	Icelandic	BR + Icelandic extensions	n/a	is
India	English	Basic Roman	n/a	I
	Hindi	Devanagari		0900-097F
			a) UN 1972(II/11), 1977 (III/12);	hi-R
			b) National: Hunterian System ²	2
	Assamese	Assamese		-
			a) UN 1972(II/11), 1977 (III/12)	as_R
	Bengali	Bengali	<i>see Bangladesh</i>	
	Gujarati	Gujarati		0A80-0AFF
			a) UN 1972(II/11), 1977 (III/12)	gu_R
	Kannada	Kannada (Dravidian)		0C80-0CFF
			a) UN 1972(II/11), 1977 (III/12)	kn_R
	Malayalam	Malayalam (Dravidian)		0D00-0D7F
			a) UN 1972(II/11), 1977 (III/12)	ml_R
	Marathi	Devanagari + Marathi extensions		0900-097F
			a) UN 1972(II/11), 1977 (III/12)	mr_R
	Oriya	Oriya		0B00-0B7F
			a) UN 1972(II/11), 1977 (III/12)	or-R
	Punjabi	Punjabi		0A80-0A7F
			a) UN 1972(II/11), 1977 (III/12)	pa_R
	Tamil	Tamil (Dravidian)		0B80-0BFF

² - although **not indicated** below, it is assumed that **the National system** is applied to names **from** other languages of India as well.

			a) UN 1972 (II/11), 1977 (III/12)	ta_R
	Telugu	Telugu (Dravidian)		0C00-0C7F
			a) UN 1972 (II/11), 1977 (III/12)	te_R
	Urdu	Perso-Arabic	<i>see Pakistan</i>	
Indonesia	Indonesian	BR + Indonesian extensions	n/a	id
Iran (Islamic Republic of)	Persian	Perso-Arabic		0600-06FF
			UN 1967 (I/13)	fa_R
Iraq	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1
Ireland	English	Basic Roman	n/a	ga
	Irish	BR + Irish extensions	n/a	
Israel	Arabic	Perso-Arabic		0600-06FF
	Hebrew	Hebrew		05B0-05F0
			UN 1977 (III/13)	he-R
Italy	Italian	BR + Italian extensions	n/a	it
	French	BR + French extensions	n/a	fr
	German	BR + German extensions	n/a	de
Jamaica	English	Basic Roman	n/a	l
Japan	Japanese	Sino-Japanese		Catered for
		Kana		3040-318F
			a) National: Modified Hepburn	ja_R1
			b) National: Kunrei	ja_R2
Jordan	Arabic	Perso-Arabic		0600-06FF
			a) UN 1972 (II/8)	ar_R1
			b) National	ar_R3
Kazakhstan	Kazakh	Cyrillic + Kazakh extensions		0400-04FF
			BGN/PCGN 1979	kk_R

	Russian	Cyrillic	<i>see Russian Federation (the)</i>	
Kenya	English	Basic Roman	n/a	1
	Swahili	Basic Roman	n/a	1
Kiribati	English	Basic Roman	n/a	1
	Kiribati	Basic Roman	n/a	1
Kuwait	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1
Kyrgyzstan	Kyrgyz	Cyrillic+Kyrgyz extensions		0400-04FF
			BGN/PCGN 1979	ky_R
	Russian	Cyrillic	<i>see Russian Federation (the)</i>	
Lao People's Democratic Republic (the)	Laotian	Laotian		0E80-0EFO
			National	lo_R
Latvia	Latvian	BR + Latvian extensions	n/a	lv
	Livonian	BR + Livonian extensions	n/a	livo
Lebanon	Arabic	Perso-Arabic		0600-06FF
			a) UN 1972 (II/8)	ar_R1
			b) I.G.N system 1973	ar_R4
	French	BR + French extensions	n/a	fr
Lesotho	English	Basic Roman	n/a	1
	Sesotho	Basic Roman	n/a	1
Liberia	English	Basic Roman	n/a	
Libyan Arab Jamahiriya (the)	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1
Liechtenstein	German	BR + German extensions	n/a	de
Lithuania	Lithuanian	BR + Lithuanian extensions	n/a	lt
	Polish	BR + Polish extensions	n/a	pl
Luxembourg	French	BR + French extensions	n/a	fr

	German	BR + German extensions	n/a	de
	Luxembourgian	BR + Luxembourgian extensions	n/a	lb
Madagascar	French	BR + French extensions	n/a	fr
	Malagasy	BR + Malagasy extensions	n/a	mg
Malawi	Chichewa	BR + Chichewa extensions	n/a	chic
	English	Basic Roman	n/a	1
Malaysia	Malay	Basic Roman	n/a	1
Maldives	Maldivian	Divehi		-
			National 1987	1
Mali	French	BR + French extensions	n/a	fr
Malta	English	Basic Roman	n/a	1
	Maltese	BR + Maltese extensions	n/a	mt
Marshall Islands (the)	English	Basic Roman	n/a	1
	Marshallese	BR + Marshallese extensions	n/a	mars
Mauritania	Arabic	Perso-Arabic		0600-06FF
			a) UN 1972 (II/8)	ar_R1
			b) I.G.N. System 1973	ar_R4
	French	BR + French extensions	n/a	fr
	Soninke	Basic Roman	n/a	1
	Toucouleur	Basic Roman	n/a	1
	Wolof	BR + Wolof extensions	n/a	wo
Mauritius	English	Basic Roman	n/a	1
Mexico	Spanish	BR + Spanish extensions	n/a	es
Micronesia (Federated States of)	English	Basic Roman	n/a	1
	Kosraean	Basic Roman	n/a	1
	Nukuoro	Basic Roman	n/a	1
	Ponapean	Basic Roman	n/a	1

List of Countries, Languages and Writing Systems

	Trukese	Basic Roman	n/a	1
	Ulithian	BR + Ulithian extensions	n/a	ulit
	Woleaian	Basic Roman	n/a	1
	Yapese	BR + Yapese extensions	n/a	yape
Monaco	French	BR + French extensions	n/a	fr
Mongolia	Mongolian	Cyrillic + Mongolian extensions		0400-04FF
			BGN/PCGN 1964	mn_R
		Mongolian		1000-109F*
Morocco	Arabic	Perso-Arabic		0600-06FF
				ar_R1
			b) I.G.N. System 1973	ar_R4
	French	BR + French extensions	n/a	fr
Mozambique	Portuguese	BR + Portuguese extensions	n/a	pt
Myanmar	Burmese	Burmese		1700-177F*
			BGN/PCGN 1970	my_R
Namibia	Afrikaans	BR + Afrikaans extensions	n/a	af
	English	Basic Roman	n/a	1
Nauru	English	Basic Roman	n/a	1
	Nauruan	Basic Roman	n/a	1
Nepal	Nepalese	De extensions		0900-097F
			a) UN 1972 (II/11), 1977 (III/12)	hi_R
			b) BGN/PCGN 1964	ne_R
Netherlands (the)	Dutch	BR + Dutch extensions	n/a	nl
	Frisian	BR + Frisian extensions	n/a	fy
	Papiamentu	Basic Roman	n/a	1
New Zealand	English	Basic Roman	n/a	
	Maori	BR + Maori extensions	n/a	mi
Nicaragua	Spanish	BR + Spanish extensions	n/a	es

Niger (the)	French	BR + French extensions	n/a	fr
Nigeria	English	Basic Roman	n/a	l
	Hausa	BR + Hausa extensions	n/a	h a —
	Ibo	BR + Ibo extensions	n/a	ibo
	Yoruba	BR + Yoruba extensions	n/a	yo —
		BR + Norwegian extensions	n/a †	no
		BR+ Finnish extensions	n/a	fi
		BR + Sámi extensions	n/a	sámi l
Oman	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar R1
Pakistan	English	Basic Roman	n/a	l
	Urdu	Perso-Arabic + Urdu extensions		0600-06FF
			a) UN 1972 (II/11), 1977 (III/12)	ur_R
			b) National: Hunterian System	2
Palau	English	Basic Roman	n/a	l
	Palauan	Basic Roman	n/a	l
Panama	Spanish	BR + Spanish extensions	n/a	l
Papua New Guinea	English	Basic Roman	n/a	l
	Neomelanesian (Tok-Pisin)	Basic Roman	n/a	l
Paraguay	Guaraní	BR + Guaraní extensions	n/a	gn
	Spanish	BR + Spanish extensions	n/a	es
Peru	Spanish	BR + Spanish extensions	n/a	es
	Aymara	Basic Roman	n/a	l
	Quechua	Basic Roman	n/a	l
Philippines (the)	English	Basic Roman	n/a	l
	Pilipino	BR + Pilipino extensions	n/a	tl
Poland	Polish	BR + Polish extensions	n/a	pl

Portugal	Portuguese	BR + Portuguese extensions	n/a	pt
Qatar	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1
Republic of Korea (the)	Korean	Hangul		Catered for
			a) National: Ministry of Education ¹ 1984	ko_R1
			b) McCune Reischauer	ko_R2
Republic of Moldova (the)	Romanian	BR + Romanian extensions	n/a	ro
Romania	Romanian	BR + Romanian extensions	n/a	ro
	Hungarian	BR + Hungarian extensions	n/a	hu
Russian Federation (the)	Russian	Cyrillic		0400-04FF
			a) UN 1987 (V/18)	ru_R1
			b) BGN/PCGN 1947	ru_R2
			d) Academy of Sciences System	ru_R3
Rwanda	English	Basic Roman	n/a	1
	French	BR + French extensions	n/a	fr
	Kinyarwanda	Basic Roman	n/a	1
Saint Kitts and Nevis	English	Basic Roman	n/a	1
Saint Lucia	English	Basic Roman	n/a	1
Saint Vincent and the Grenadines	English	Basic Roman	n/a	1
Samoa	English	Basic Roman	n/a	1
	Samoan	BR + Samoan extensions	n/a	sm
San Marino	Italian	BR + Italian extensions	n/a	it
Sao Tome and Principe	Portuguese	BR + Portuguese extensions	n/a	pt

Saudi Arabia	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1
Senegal	French	BR + French extensions	n/a	fr
Seychelles	Creole (Seselwa)	Basic Roman	n/a	1
	English	Basic Roman	n/a	1
	French	BR + French extensions	n/a	fr
Sierra Leone	English	Basic Roman	n/a	1
Singapore	Chinese	Chinese		Catered for
	English	Basic Roman	n/a	1
	Malay	Basic Roman	n/a	1
	Tamil	Tamil		0B80-0BFF
			UN 1972 (II/11), 1977 (III/12)	ta_R
Slovakia	Slovak			
	Hungarian	BR + Hungarian extensions	n/a	hu
Slovenia	Slovene	BR + Slovene extensions	n/a	sl
	Hungarian	BR + Hungarian extensions	n/a	hu
	Italian	BR + Italian extensions	n/a	it
Solomon Islands	English	Basic Roman	n/a	1
Somalia	Arabic	Perso-Arabic		0600-06FF
	Somali	Basic Roman	n/a	1
South Africa	Afrikaans	BR+ Afrikaans extensions	n/a	af
	English	Basic Roman	n/a	1
	Ndebele	Basic Roman	n/a	1
	Pedi	Basic Roman	n/a	1
	Sotho	Basic Roman	n/a	1
	siSwati	Basic Roman	n/a	1
	Tsonga	Basic Roman	n/a	1
	Tswana	BR + Tswana extensions	n/a	tn

	Venda	BR + Venda extensions	n/a	vend
	Xhosa	Basic Roman	n/a	l
	Zulu	Basic Roman	n/a	l
Spain	Spanish	BR + Spanish extensions	n/a	es
	Basque	BR + Basque extensions	n/a	eu
	Catalan	BR + Catalan extensions	n/a	ca
	Galician	BR + Galician extensions	n/a	gl
Sri Lanka	Sinhalese	Sinhalese		0D80-0DFF
	Tamil	Tamil (Dravidian)		0B80-0BFF
			UN 1972 (II/11), 1977 (III/12)	ta_R
	English	Basic Roman	n/a	l
Sudan (the)	Arabic	Perso-Arabic		0600-06FF
			BGN/PCGN 1956	ar_R1
	English	Basic Roman	n/a	l
Suriname	Dutch	BR + Dutch extensions	n/a	nl
Swaziland	English	Basic Roman	n/a	l
	siSwati	Basic Roman	n/a	l
Sweden	Swedish	BR + Swedish extensions	n/a	sv
	Finnish	BR + Finnish extensions	n/a	fi
	Sámi	BR + Sámi extensions	n/a	sámi1
Switzerland	French	BR + French extensions	n/a	fr
	German	BR + German extensions	n/a	de
	Italian	BR + Italian extensions	n/a	it
	Rhaeto-Romance	BR + Rhaeto-Romance extensions	n/a	rm
Syrian Arab Republic (the)	Arabic	Perso-Arabic		0600-06FF
			BGN/PCGN 1956	ar_R1
Tajikistan	Tajik	Cyrillic + Tajik extensions		0400-04FF
			BGN/PCGN 1994	tg_R

	Russian	Cyrillic	see <i>Russian Federation (the)</i>	
Thailand	Thai	Thai		0E00-0E7F
			UN 1%7 (I/14)	1
The former Yugoslav Republic of Macedonia	Macedonian	Cyrillic		0400-04FF
			UN 1977 (III/11)	mk_R
Togo	French	BR + French extensions	n/a	fr
Tonga	English	Basic Roman	n/a	1
	Tongan	BR + Tongan extensions	n/a	to
Trinidad and Tobago	English	Basic Roman	n/a	1
Tunisia	Arabic	Perso-Arabic	n/a	0600-06FF
			a) UN 1972(II/8)	ar_R1
			b) I.G.N. System 1973	ar_R4
	French	BR + French extensions	n/a	fr
Turkey	Turkish	BR + Turkish extensions	n/a	tr
Turkmenistan	Turkmen	Cyrillic + Turkmen extensions		0400-04FF
			BGN/PCGN 1979	tk_R
		BR + Turkmen extensions	n/a	tk
Tuvalu	English	Basic Roman	n/a	1
	Tuvaluan	Basic Roman	n/a	1
Uganda	English	Basic Roman	n/a	1
Ukraine	Ukrainian	Cyrillic + Ukrainian extensions		0400-04FF
			a) National 1993	uk_R
			b) National 1996	1
			c) BGN/PCGN 1965	1
United Arab Emirates (the)	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1

United Kingdom of Great Britain and Northern Ireland (the)	English	Basic Roman	n/a	1
	Gaelic (Scots)	BR + Gaelic extensions	n/a	gd
	Irish	BR + Irish extensions	n/a	ga
	Welsh	BR + Welsh extensions	n/a	cy
United Republic of Tanzania (the)	English	Basic Roman	n/a	1
	Swahili	Basic Roman	n/a	1
United States of America (the)	English	Basic Roman	n/a	1
	Chamorro	BR + Chamorro extensions	n/a	cham
	French	BR + French extensions	n/a	fr
	Hawaiian	BR + Hawaiian extensions	n/a	hawa
	Samoan	Basic Roman	n/a	1
	Spanish	BR + Spanish extensions	n/a	es
Uruguay	Spanish	BR + Spanish extensions	n/a	es
Uzbekistan	Russian	Cyrillic	<i>see Russian Federation (the)</i>	
	Uzbek	Cyrillic + Uzbek extensions		0400-04FF
			BGN/PCGN 1979	uz_R
		BR + Uzbek extensions	n/a	uz
Vanuatu	Bislama	BR + Bislama extensions	n/a	bi
	English	Basic Roman	n/a	1
	French	BR + French extensions	n/a	fr
Venezuela	Spanish	BR + Spanish extensions	n/a	es
Viet Nam	Vietnamese	BR + Vietnamese extensions	n/a	vi
Yemen	Arabic	Perso-Arabic		0600-06FF
			UN 1972 (II/8)	ar_R1
Yugoslavia	Serbian	Cyrillic		0400-04FF

			UN 1977 (III/1 1)	sr_R
	Albanian	BR+ Albanian extensions	n/a	sq
	Hungarian	BR + Hungarian extensions	n/a	hu
Zambia	English	Basic Roman	n/a	1
Zimbabwe	English	Basic Roman	d a	I

Tables of Characters

Key to Columns below:

Table headings - Table names contain the language abbreviation according to ISO 639 (1988). If such an abbreviation is not available, the first four letters of the **English name** of the language are used instead. References of tables listing romanized characters contain additionally a letter **R** and a number if **necessary**. Tables **connected** to multiple languages are numbered. If the **name** of the language does not coincide with that of the writing system, the language **name** is added in brackets. Notes to tables refer to the analysis of the practical **aspects** of the implementation of the ISO **standards**, described in para 20 of the main report and in the notes given below.

1. Languages covered by Table 1 or Basic Roman (Languages where only one or several character **sets** out of the total are covered by Table 1 are marked with an asterisk; those languages are to be found also in the tables below):

Armenian	Malay	Somali
Aymara	Maldivian	Soninke
Belarusian*	Nauruan	Swahili
Creole (Seselwa)	Ndebele	Thai
English	Neomelanesian (Tok Pisin)	Toucouleur
Fijian	Nukuoro	Trukese
Georgian	Palauan	Tsonga
Greenlandic	Papiamentu	Tuvaluan
Ikiribati	Pedi	Ukrainian*
Kinyarwanda	Ponapean	Woleaian
Kirundi	Quechua	Xhosa
Kosraean	Sesotho (Sotho)	Zulu
Latin	siSwati (Swati)	

2. Abbreviations of languages covered by tables of extensions:

Afrikaans	af	Frisian	fy	Luxembourgian	lb
Albanian	sq	Gaelic (Scots)	gd	Macedonian	<i>mk</i>
Amharic	am	Galician	gl	Malagasy	mg
Arabic	ar	G e m	de	Maltese	mt
Azerbaijani	az	Greek	el	Maori	mi
Basque	eu	Guarani	gn	Marshallese	mars
Belarusian	be	Hausa	ha	Mongolian	mn
Bengali	bn	Hawaiian	hawa	Nepali	ne
Bislama	bi	Hebrew	he	Norwegian	no
Bosnian	bosn	Hĩĩĩ	hi	Pashto	ps
Breton	br	Hungarian	hu	Persian	fa
Bulgarian	bg	Ibo	ibo	Pilipino	tl
Burmese	my	Icelandic	is	Polish	pl
Catalan	ca	Indonesian	id	Portuguese	pt
Chamorro	Cham	Irish	ga	Romance	rm
Chichewa	chic	Italian	it	Romanian	ro
Chinese	zh	Japanese	ja	Russian	ru
Croatian	hr	Kazakh	kk	Sámi	sámi
Czech	cs	Khmer	km	Samoan	sm
Danish	da	Korean	ko	Serbian	sr
Dutch	nl	Kyrgyz	ky	Slovak	sk
Estonian	et	Laotian	lo	Slovene	sl
Faroese	fo	Latvian	lv	Sorbian	sorb
Finnish	fi	Lithuanian	lt	Spanish	es
French	fr	Livonian	livo	Swedish	sv

Tagalog: <i>see</i> Pilipino		Turkmen	tk	Vietnamese	vi
Tajik	tg	Ukrainian	uk	Welsh	cy
Tamil	ta	Ulithian	ulit	Wolof	wo
Tongan	to	Urdu	ur	Yapese	yape
Tswana	tn	Uzbek	uz	Yoruba	yo
Turkish	tr	Venda	vend		

3. Index of tables:

Page/tbl	Title				
4	1	Basic Roman	26	gn	Guarani
5	2	Hunterian System	27	gu_R	Gujarati
5	3	Canadian	27	ha	Hausa
6	af	Afrikaans	28	hawa	Hawaiian
6	am_R1	Amharic (UN)	28	he-R	Hebrew
7	am_R2	Amharic (BGN/PCGN)	28	hi-R	Hindi
7	ar_R1	Arabic (UN)	29	hr	Croatian
8	ar_R2	Arabic (Egypt)	30	hu	Hungarian
8	ar_R3	Arabic (Jordan)	30	ibo	Ibo
9	ar_R4	Arabic (I.G.N.)	31	id	Indonesian
10	as-R	Assamese	31	is	Icelandic
11	az	Azerbaijani	32	it	Italian
11	be_R	Belarusian	32	ja_R1	Japanese (Hepburn)
12	bg_R1	Bulgarian (UN)	32	ja_R2	Japanese (Kunrei)
12	bg_R2	Bulgarian (BGN/PCGN)	33	kk_R	Kazakh
12	bi	Bislama	33	km_R	Khmer
13	bn_R	Bengali	34	kn_R	Kannada
13	bosn	Bosnian	35	ko_R1	Korean (McCune Reischauer)
14	br	Breton	35	ko_R2	Korean (MOE)
14	ca	Catalan	35	ky_R	Kyrgyz
15	cham	Chamorro	36	lb	Luxembourgian
15	chic	Chichewa	36	livo	Livonian
16	cs	Czech	38	lo-R	Laotian
17	cy	Welsh	38	lt	Lithuanian
18	da	Danish	39	lv	Latvian
18	de	German	40	mars	Marshallese
19	el-R	Greek	40	mg	Malagasy
20	es	Spanish	40	mi	Maori
20	et	Estonian	41	mk_R	Macedonian
21	eu	Basque	41	ml_R	Malayalam
21	fa-R	Persian	42	mn_R	Mongolian
22	fi	Finnish	43	mr_R	Marathi
22	fo	Faroese	43	mt	Maltese
23	fr	French	43	my-R	Burmese
24	fy	Frisian	44	ne_R	Nepali
24	ga	Irish	44	nl	Dutch
25	gd	Gaelic (Scots)	45	no	Norwegian
25	gl	Galician	46	or-R	Oriya
			47	pa-R	Punjabi
			47	pl	Polish
			48	ps_R	Pashto
			49	pt	Portuguese
			50	rm	(Rhaeto-)Romance
			50	ro	Romanian
			51	ru_R1	Russian (UN)
			52	ru_R2	Russian (BGN/PCGN)
			52	ru_R3	Russian (Acad. of Sc.)
			52	sámi1	Sámi (general)
			53	sámi2	Sámi (Skolt)
			54	sk	Slovak
			55	sl	Slovene
			55	sm	Samoa
			55	sorb1	Lower Sorbian
			56	sorb2	Upper Sorbian
			57	sq	Albanian
			57	sr_R	Serbian
			58	sv	Swedish
			58	ta_R	Tamil
			59	te_R	Telugu
			60	tg_R	Tajik
			60	tk	Turkmen (Roman)
			61	tk_R	Turkmen (Cyrillic)
			61	tl	Pilipino (Tagalog)
			62	tn	Tswana
			62	to	Tongan
			62	tr	Turkish
			63	uk_R	Ukrainian
			63	ulit	Ulithian
			64	ur_R	Urdu
			64	uz	Uzbek (Roman)
			65	uz_R	Uzbek (Cyrillic)
			65	vend	Venda
			65	vi	Vietnamese
			70	wo	Wolof
			70	yape	Yapese
			70	yo	Yoruba
			71	zh_R1	Chinese (UN)
			72	zh_R2	Chinese (Wade- Giles)

4. Notes to Tables

Note 1. Cases where romanisation systems have been approved by the United Nations but have not been implemented by the **country** concerned.

Note 2. Cases where **non-Roman** writing systems have not yet been catered for in the ISO10646/Unicode standard.

Note 3. Instances where **Roman** characters are not available in: (a) ISO 8859, (b) ISO 10646.

Note 4. Instances where characters could not be encoded as a single unified character but only in conjunction with a combining diacritical **mark**.

Note 5. Cases where the word processing software **and** fonts used by the WG could not display or print the character. Formatting commands (e.g. for underline) or, as a last resort, manuscript additions had to be made to the hardcopy.

Note 6. Multiple **coding**. There are 3 ~~cases~~ identified by the survey where the character appears more than once in the ISO 10646/Unicode standard with different **code** references.

Note 7. Cases where codes are available in ISO 8859 but are difficult to implement as they occur in different parts of the standard.

Char. (Column 1) - characters are in alphabetical order **A** to **Z** with upper case characters before lower case characters. The order of diacritics is taken from the Unicode Table 8 Row 03 (Combining Diacritical Marks, see **Annex A**). Ligatures are given in the order of the first letter of the ligature, after combinations of the same letter with diacritical **marks**. Other variant characters are given next to their closest counterpart in the Basic Roman character table.

Character Name (Column 2) - **name** of character as listed in ISO 10646/Unicode with "Roman" replacing "Latin". Names given in brackets follow the *same* pattern, but **describe** characters which cannot actually be found in a unified form in this standard

ISO 8859 (Column 3) - two digit code is given with the part number(s) in brackets. Dash equals no code. For some characters different codes are given in various parts of the ISO 8859 standard. At present the ISO 8859 consists of 10 parts (8859-1 – Latin 1; 8859-2 – Latin 2; 8859-3 – Latin 3; 8859-4 – Latin 4; 8859-5 – Latin/Cyrillic; 8859-6 – Latin/Arabic; 8859-7 – Latin/Greek; 8859-8 – Latin/Hebrew; 8859-9 – Latin 5; 8859-10 – Latin 6) with another 4 parts in draft form only (marked with an asterisk: 8859-11* – Latin/Thai; 8859-12* – Latin 7 (Celtic); 8859-13* – Latin 8 (Baltic Rim); 8859-14* – Latin 9 (Sámi)). For individual characters the codes of all parts of the standard are included, to make it possible to choose a particular part of the standard that **suits** best the given language.

ISO/Unic. (Column 4) - four digit code of the ISO 10646/Unicode is given unless the character is not unified. In the latter case two or three codes with † sign are given, consisting usually of the code of the basic character and that of the combining diacritical **mark**.

Table 1
Writing System: Basic Roman

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Apostrophe	39 (1..14)	0039	
A	Roman capital letter A	41 (1..14)	0041	
a	Roman small letter a	61 (1..14)	0061	
B	Roman capital letter B	42 (1..14)	0042	
b	Roman small letter b	62 (1..14)	0062	
C	Roman capital letter C	43 (1..14)	0043	
c	Roman small letter c	63 (1..14)	0063	
D	Roman capital letter D	44 (1..14)	0044	
d	Roman small letter d	64 (1..14)	0064	
E	Roman capital letter E	45 (1..14)	0045	
e	Roman small letter e	65 (1..14)	0065	
F	Roman capital letter F	46 (1..14)	0046	
f	Roman small letter f	66 (1..14)	0066	
G	Roman capital letter G	47 (1..14)	0047	
g	Roman small letter g	67 (1..14)	0067	
H	Roman capital letter H	48 (1..14)	0048	
h	Roman small letter h	68 (1..14)	0068	
I	Roman capital letter I	49 (1..14)	0049	
i	Roman small letter i	69 (1..14)	0069	
J	Roman capital letter J	4A (1..14)	004A	
j	Roman small letter j	6A (1..14)	006A	
K	Roman capital letter K	4B (1..14)	004B	
k	Roman small letter k	6B (1..14)	006B	
L	Roman capital letter L	4C (1..14)	004C	
l	Roman small letter l	6C (1..14)	006C	
M	Roman capital letter M	4D (1..14)	004D	
m	Roman small letter m	6D (1..14)	006D	
N	Roman capital letter N	4E (1..14)	004E	
n	Roman small letter n	6E (1..14)	006E	
O	Roman capital letter O	4F (1..14)	004F	
o	Roman small letter o	6F (1..14)	006F	
P	Roman capital letter P	50 (1..14)	0050	
p	Roman small letter p	70 (1..14)	0070	
Q	Roman capital letter Q	51 (1..14)	0051	
q	Roman small letter q	71 (1..14)	0071	
R	Roman capital letter R	52 (1..14)	0052	
r	Roman small letter r	72 (1..14)	0072	
S	Roman capital letter S	53 (1..14)	0053	
s	Roman small letter s	73 (1..14)	0073	
T	Roman capital letter T	54 (1..14)	0054	
t	Roman small letter t	74 (1..14)	0074	
U	Roman capital letter U	55 (1..14)	0055	
u	Roman small letter u	75 (1..14)	0075	
V	Roman capital letter V	56 (1..14)	0056	
v	Roman small letter v	76 (1..14)	0076	
W	Roman capital letter W	57 (1..14)	0057	
w	Roman small letter w	77 (1..14)	0077	
X	Roman capital letter X	58 (1..14)	0058	
x	Roman small letter x	78 (1..14)	0078	
Y	Roman capital letter Y	59 (1..14)	0059	

Y	Roman small letter y	79 (1..14)	0079
Z	Roman capital letter Z	5A (1..14)	005A
z	Roman small letter z	7A (1..14)	007A
Number of characters: 53			

Table 2
Writing Systems of the Languages of Bangladesh, India and Pakistan
Romanization System: Munterian System

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	00 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	

Number of characters: 6

Table 3
Writing Systems of the Indigenous Languages of Canada¹
Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
&	Ampersand		0038	
ǎ	(Roman small letter a with circumflex and ogonek)		00E2 +0328	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
ē	Roman small letter e with macron	BA (4), B2 (10), E7 (13*)	0113	
Ɔ	(Roman capital letter G with line below)		0047 +0331	
ǰ	(Roman small letter g with line below)		0067 +0331	
ı	Roman small letter i with ogonek	E7 (4,10), E1 (13*)	012F	
Ɔ	Roman capital letter K with line below		1E34	
Ʒ	Roman small letter k with line below		1E35	
Ł	Roman capital letter L with stroke	A3 (2), D9 (13*)	0141	
ł	Roman small letter l with stroke	B3 (2), F9 (13*)	0142	
ō	Roman small letter o with macron	F2 (4,10), F4 (13*)	014D	
ó	Roman small letter o with macron and acute		1E53	
ȯ	(Roman small letter o with macron acute and ogonek)		1E53 +0328	

¹ the full inventory of these languages (Chipewyan, Inuktitut, Kaska, Northern Tutchone, Tlingit, etc.) needs further research. At present the table lists only modified extended Roman alphabet characters (also referred to as "hard-to-construct characters") found in the geographical names of Yukon. In addition the inventory of the languages may contain "ordinary" characters with diacritical marks of the Latin-1 code table: Àá Êè Éé Êê Ìí Úú Ûû etc.

υ	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B
X̄	(Roman capital letter X with line below)		0058+0331
ẋ	(Roman small letter x with line below)	-	0078+0331

Number of characters: 17

Table: af

Writing System: Afrikaans Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
E	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
e	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
E	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
E	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10,12*,14*)	00CB	
e	Roman small letter e with diaeresis	EB (1,2,3,4,9,10,12*,14*)	WEB	
ï	Roman capital letter I with circumflex	CE (1,2,3,4,9,10,12*,14*)	00CE	
ï	Roman small letter i with circumflex	EE (1,2,3,4,9,10,12*,14*)	00EE	
Ï	Roman capital letter I with diaeresis	CF (1,3,9,10,12*,14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*,14*)	CQEF	
O	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
o	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	WF4	
U	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*,14*)	00DB	
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*,14*)	00FB	

Number of characters: 16

Table: am_R1

Writing System: Ethiopic [Amharic]

Romanization System: UN 1967 (I/17)

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
E	(Roman capital letter E with line below)	-	0045+0331	
ē	(Roman <i>small</i> letter e with line below)	-	0065+0331	
I	(Roman capital letter I with line below)	-	0049+0331	
ı	(Roman <i>small</i> letter i with line below)	-	0069+0331	

Number of characters: 4

Table: am_R2
 Writing System: Ethiopic [Amharic]
 Romanization System: BGN/PCGN 1967
 Notes: 3a, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Inverted apostrophe		02BB (2018)	
À	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ĕ	Roman capital letter E with macron	AA (4), A2 (10), C7 (13*)	0112	
ē	Roman small letter e with macron	BA (4), B2 (10), E7 (13*)	0113	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ı̇	Roman small letter i with macron	EF (4), IA (10), EE (13*)	012B	

Number of characters: 7

Table: ar_R1
 Writing System: Perso-Arabic [Arabic]
 Romanization System: UN 1972 (II/8)
 Notes: 3a, 4, 5, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Inverted apostrophe		02BB (2018)	
À	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
Ā	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ď	Roman capital letter D with cedilla		1E10	
ď	Roman small letter d with cedilla		1E11	
Ĥ	Roman capital letter H with cedilla		1E28	
ĥ	Roman small letter h with cedilla	-	1E29	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ı̇	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ş	Roman capital letter S with cedilla	AA (2,3), DE (9)	015E	
ş	Roman small letter s with cedilla	BA (2,3), FE (9)	015F	
Ť	Roman capital letter T with cedilla	DE (2)	0162	
ť	Roman small letter t with cedilla	FE (2)	0163	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Ț	(Roman capital letter Z with cedilla)	-	005A+0327	
ț	(Roman small letter z with cedilla)	-	007A+0327	

Number of characters: 19

Table: ar_R2

Writing System: Perso-Arabic [Arabic]

Romanization System: Survey of Egypt

Notes: 3a, 5, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
'	Reversed apostrophe		02BD (201B)	
A	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	
D	Roman capital letter D with dot below		1E0C	
d	Roman small letter d with dot below		1E0D	
H	Roman capital letter H with dot below		1E24	
h	Roman small letter h with dot below		1E25	
i	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE	
î	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE	
Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4	
S	Roman capital letter S with dot below		1E62	
s	Roman small letter s with dot below	-	1E63	
T	Roman capital letter T with dot below		1E6C	
t	Roman small letter t with dot below		1E6D	
U	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*, 14*)	00DB	
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*, 14*)	00FB	
Z	Roman capital letter Z with dot below		1E92	
z	Roman small letter z with dot below		1E93	

Number of characters: 19

Table: ar_R3

Writing System: Perso-Arabic [Arabic]

Romanization System: National System of Jordan

Notes: 3a, 4, 5, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
D	Roman capital letter D with line below		1E0E	
ḏ	Roman small letter d with line below	-	1E0F	
H	(Roman capital letter H with line below)	-	0048+0331	
ḥ	Roman small letter h with line below		1E96	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE	012B	

Š	(Roman capital letter S with line below)	(13*)	0053+033 1
š	(Roman small letter s with line below)		0073+033 1
Ť	Roman capital letter T with line below		1E6E
ť	Roman small letter t with line below		1E6F
Ů	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ů	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B

Number of characters: 14

Table: ar R4
Writing System: Perso-Arabic [Arabic]
Romanization System: I.G.N. System 1973
Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
Đ	Roman capital letter D with dot below		1E0C	
đ	Roman small letter d with dot below		1E0D	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ĥ	Roman capital letter H with dot below		1E24	
ĥ	Roman small letter h with dot below		1E25	
Î	Roman capital letter I with circumflex	CE (1,2,3,4,9,10,12*,14*)	00CE	
î	Roman small letter i with circumflex	EE (1,2,3,4,9,10,12*,14*)	00EE	
Ï	Roman capital letter I with diaeresis	CF (1,3,9,10,12*,14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*,14*)	00EF	
Ï	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	
Š	Roman capital letter S with dot below		1M2	
š	Roman small letter s with dot below		1E63	
Ť	Roman capital letter T with dot below		1E6C	
ť	Roman small letter t with dot below		1E6D	
Ů	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*,14*)	00DB	

ú	Roman small letter u with circumflex	FB (1,3,4,9,10,12*,14*)	00FB
Z	Roman capital letter Z with dot below	-	1E92
z	Roman small letter z with dot below	-	1E93
Number of characters: 28			

Table: **as_R****Writing System: Assamese****Romanization System: UN 1972 (II/11), 1977 (III/12)**

Notes: 1, 2, 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
a	Roman small letter a with macron	EO (4,10), E2 (13*)	0101	
D	Roman capital letter D with dot below		1E0C	
d	Roman small letter d with dot below		1E0D	
H	Roman capital letter H with dot below		1E24	
h	Roman small letter h with dot below		1E25	
I	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ı	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
J	(Roman capital letter J with line below)		004A + 0331	
ı	(Roman small letter j with line below)		006A + 0331	
L	(Roman capital letter L with diaeresis below)		004C + 0324	
ı	(Roman small letter l with diaeresis below)		006C + 0324	
M	Roman capital letter M with dot above	B6 (12*)	1E40	
m	Roman small letter m with dot above	B7 (12*)	1E41	
N	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
A	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
N	Roman capital letter N with dot above		1E44	
n	Roman small letter n with dot above		1E45	
N	Roman capital letter N with dot below		1E46	
n	Roman small letter n with dot below		1E47	
R	Roman capital letter R with dot above		1E58	
r	Roman small letter r with dot above		1E59	
R	Roman capital letter R with dot below		1E5A	
r	Roman small letter r with dot below		1E5B	
S	Roman capital letter S with dot below		1E62	
s	Roman small letter s with dot below		1E63	
T	Roman capital letter T with dot below		1E6C	
t	Roman small letter t with dot below		1E6D	
U	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Number of characters: 30				

Table: az
Writing System: Azerbaijani Extensions to Basic Roman
Notes: 3a, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,12*,13*,14*)	00C4	Variant of schwa
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,12*,13*,14*)	00E4	Variant of schwa
ä	Roman capital letter schwa		018F	
a	Roman small letter schwa		0259 (01DD)	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
Ġ	Roman capital letter G with breve	AB (3), D0 (9)	011E	
ğ	Roman small letter g with breve	BB (3), F0 (9)	011F	
İ	Roman capital letter I with dot above	A9 (3), DD (9)	0130	
ı	Roman small letter i with no dot	B9 (3), FD (9)	0131	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10,12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10,12*,13*,14*)	00F6	
Ş	Roman capital letter S with cedilla	AA (2,3), DE (9)	015E	
ş	Roman small letter s with cedilla	BA (2,3), FE (9)	015F	
U	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10,12*,13*,14*)	00DC	
u	Roman small letter u with diaeresis	FC (1,2,3,4,9,10,12*,13*,14*)	00FC	

Number of characters: 16

Table: be-R
Writing System Cyrillic [Belarusian]
Romanization System: National, unofficial 1992/93
Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
E	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10,12*,14*)	00CB	
e	Roman small letter e with diaeresis	EB (1,2,3,4,9,10,12*,14*)	00EB	
H̄	(Roman capital letter H with line below)		0048+0331	
h̄	Roman small letter h with line below	-	1E96	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ū	(Roman capital letter U with line below)		0055+0331	

	below)		
u	(Roman small letter u with line below)	-	0075+0331
Z	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E
Number of characters: 14			

Table: bg_R1**Writing System:** Cyrillic [Bulgarian]**Romanization System:** UN 1977 (III/10)

Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ă	Roman capital letter A with caron		01CD	
ă	Roman small letter a with caron		01CE	
Ĉ	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
ĉ	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Z	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	
Number of characters: 8				

Table: bg_R2**Writing System:** Cyrillic [Bulgarian]**Romanization System:** BGN/PCGN 1952

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ū	Roman capital letter U with breve	DD (3)	016C	
ū	Roman small letter u with breve	FD (3)	016D	
Number of characters: 2				

Table: bi**Writing System:** Bislama Extensions to Basic Roman

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
İ	Roman capital letter I with diaeresis	CF (1,3,9,10,12*, 14*)	00CF	
ı	Roman small letter i with diaeresis	EF (1,3,9,10,12*, 14*)	00EF	
Ṁ	(Roman capital letter M with macron)	-	004D+0304	

m̄	(Roman small letter m with macron)		006D+0304
P̄	(Roman capital letter P with macron)		0050+0304
p̄	(Roman small letter p with macron)		0070+0304
Ū	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC

Number of characters: 10

Table: bn_R
Writing System: Bengali
Romanization System: UN 1972 (II/11), 1977 (III/12)
Notes: 1, 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
D	Roman capital letter D with dot below		1E0C	
d	Roman small letter d with dot below		1E0D	
H	Roman capital letter H with dot below		1E24	
h	Roman small letter h with dot below		1E25	
I	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
i	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
J	(Roman capital letter J with line below)		004A+0331	
j	(Roman small letter j with line below)		006A+0331	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ñ	Roman capital letter N with dot above		1E44	
n	Roman small letter n with dot above		1E45	
N	Roman capital letter N with dot below		1E46	
n	Roman small letter n with dot below		1E47	
Ř	Roman capital letter R with dot above		1E58	
ř	Roman small letter r with dot above		1E59	
R	Roman capital letter R with dot below		1E5A	
r	Roman small letter r with dot below		1E5B	
S	Roman capital letter S with dot below	-	1E62	
s	Roman small letter s with dot below	-	1E63	
T	Roman capital letter T with dot below	-	1E6C	
t	Roman small letter t with dot below		1E6D	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
u	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	

Number of characters: 26

Table: bosn
Writing System: Bosnian Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ć	Roman capital letter C with acute	C6 (2), C3 (13*)	0106	

c	Roman small letter c with acute	E6 (2), E3 (13*)	0107
Ĉ	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C
ċ	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D
Đ	Roman capital letter D with stroke	D0 (2), A9 (10), A3 (14*)	0110
đ	Roman small letter d with stroke	F0 (2), B9 (10), A4 (14*)	0111
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	016Q
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E

Number of characters: 10

Table: br

Writing System: Breton Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	CQEA	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	FI (1,3,9,12*,14*)	00F1	
Ó	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ó	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	
Û	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9	
u	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10,12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10,12*,13*,14*)	00FC	

Number of characters: 12

Table: ca

Writing System: Catalan Extensions to Basic Roman

Notes: 3a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	C0 (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	

è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8
Ē	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED
Ī	Roman capital letter I with diaeresis	CF (1,3,9,10,12*, 14*)	00CF
ī	Roman small letter i with diaeresis	EF (1,3,9,10,12*, 14*)	00EF
Ĺ	Roman capital letter L with middle dot		013F
ĺ	Roman small letter l with middle dot		0140
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
Û	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC

Number of characters: 22

Table: cham

Writing System: Chamorro Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ă	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
ă	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	
Ȧ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ȧ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ū	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	

Number of characters: 6

Table: chic

Writing System: Chichewa Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ŵ	Roman capital letter W with circumflex	D0 (12*)	0174	

w **Roman** small letter w with circumflex F0 (12*) 0175
 Number of characters: 2

Table: csWriting System: Czech Extensions to Basic **Roman**

Notes: 4

Char.	Character Name	ISO 8859	ISO/Unic.	Notes/Exceptions
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Ď	Roman capital letter D with caron	CF (2)	010E	
ď	Roman small letter d with caron	EF (2)	010F	Usual typesetting is ě (0064+0315)
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
Ě	Roman capital letter E with caron	CC (2)	011A	
ě	Roman small letter e with caron	EC (2)	011B	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10,12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10,12*,14*)	00ED	
Ň	Roman capital letter N with caron	D2 (2)	0147	
ň	Roman small letter n with caron	F2 (2)	0148	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*,13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*,13*,14*)	00F3	
Ř	Roman capital letter R with caron	D8 (2)	0158	
ř	Roman small letter r with caron	F8 (2)	0159	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ť	Roman capital letter T with caron	AB (2)	0164	
ť	Roman small letter t with caron	BB (2)	0165	Usual typesetting is ě (0074+0315)
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10,12*,14*)	00DA	
u	Roman small letter u with acute	FA (1,2,3,4,9,10,12*,14*)	CQFA	
Ů	Roman capital letter U with ring above	D9 (2)	016E	
ů	Roman small letter u with ring above	F9 (2)	016F	
Y	Roman capital letter Y with acute	DD (1,2,10,12*,14*)	O0DD	
ý	Roman small letter y with acute	FD (1,2,10,12*,14*)	O0FD	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	

ž **Roman** small letter z with caron BE (2,4), BC (10), 017E
FE (13*), BA (14*)

Number of characters: 30

Table: cy
Writing System: Welsh Extensions to Basic Roman

Notes: 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Ā	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,12*,13*,14*)	00E4	
Ē	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
Ĕ	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ë	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10,12*,14*)	00CB	
ë	Roman small letter e with diaeresis	EB (1,2,3,4,9,10,12*,14*)	00EB	
Ī	Roman capital letter I with circumflex	CE (1,2,3,4,9,10,12*,14*)	00CE	
ï	Roman small letter i with circumflex	EE (1,2,3,4,9,10,12*,14*)	00EE	
Ï	Roman capital letter I with diaeresis	CF (1,3,9,10,12*,14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*,14*)	00EF	
Ď	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ó	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10,12*,13*,14*)	00D6	
ö	Roman <i>small</i> letter o with diaeresis	F6 (1,2,3,4,9,10,12*,13*,14*)	00F6	
Û	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9	
ù	Roman <i>small</i> letter u with grave	F9 (1,3,9,12*,14*)	00F9	
Ū	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*,14*)	00DB	
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*,14*)	00FB	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10,	00DC	

		12*,13*,14*)	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	OOFC
Ŵ	Roman capital letter W with circumflex	D0 (12*)	0174
w	Roman small letter w with circumflex	F0 (12*)	0175
Ẁ	Roman capital letter W with diaeresis	BD (12*)	1E84
w	Roman small letter w with diaeresis	BE (12*)	1E85
Ỳ	Roman capital letter Y with circumflex	DE (12*)	0176
ỳ	Roman small letter y with circumflex	FE (12*)	0177
Ỳ	Roman capital letter Y with diaeresis	AF (12*)	0178
ÿ	Roman small letter y with diaeresis	FF (1,9,12*)	00FF

Number of characters: 34

Table: da

Writing System: Danish Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with ring above	C5 (1,4,9,10,12*, 13*,14*)	00C5	
á	Roman small letter a with ring above	E5 (1,4,9,10,12*, 13*,14*)	00E5	
Æ	Roman capital letter AE (ash)	C6 (1,4,9,10,12*, 14*), AF (13*)	00C6	
æ	Roman small letter ae (ash)	E6 (1,4,9,10,12*, 14*), BF (13*)	00E6	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ø	Roman capital letter O with stroke	D8 (1,4,9,10,12*, 14*)	00D8	
ø	Roman small letter o with stroke	F8 (1,4,9,10,12*, 14*)	00F8	

Number of characters: 8

Table: de

Writing System: German Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10, 12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10, 12*,13*,14*)	00E4	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
ß	Roman small letter <i>sharp s</i> (German)	DF (1,2,3,4,9,10, 12*,13*,14*)	00DF	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	

12*,13*,14*)

Number of characters: 7

Table: el-R
 Writing System: Greek
 Romanization System: UN 1987²
 Notes: 3a, 4, 5

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
F	(Roman capital letter F with line below)		0046+0331	
f	(Roman small letter f with line below)		0066+0331	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
Ï	Roman capital letter I with diaeresis	CF (1,3,9,10,12*, 14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*, 14*)	00EF	
İ	(Roman capital letter I with diaeresis and acute)		00CF+0301	
ı	(Roman small letter i with diaeresis and acute)	-	00EF+0301	
Ɔ	(Roman capital letter I with line below)		0049+0331	
Ɔ	(Roman small letter i with line below)		0069+0331	
İ	(Roman capital letter I with line below and acute)		00CD+0331	
ı	(Roman small letter i with line below and acute)	-	00ED+0331	
Ɔ	Roman capital letter N with line below		1E48	
Ɔ	Roman small letter n with line below		1E49	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
Ɔ	(Roman capital letter O with line below)		004F+0331	
Ɔ	(Roman small letter o with line below)		006F+0331	
Ɔ	(Roman capital letter O with line below and acute)		00D3+0331	
Ɔ	(Roman small letter o with line below)	-	00F3+0331	

² the acute accent used to represent the Greek tonos (stress mark) has been included in the above table

	and acute)		
U	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	OODA
u	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	CQFA
V	(Roman capital letter V with line below)		0056+0331
v	(Roman small letter v with line below)		0076+0331
Y	Roman capital letter Y with acute	DD (1,2,10,12*,14*)	OODD
Y	Roman small letter y with acute	FD (1,2,10,12*,14*)	OOFD
Number of characters: 30			

Table: es**Writing System: Spanish Extensions to Basic Roman**

Notes: n/a

<i>Char.</i>	Character Name	ISO 8859	ISO/Unic.	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
a	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
N	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
O	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
U	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	OODA	
u	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	OQFA	
U	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
u	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	OQFC	

Number of characters: 14

Table: et**Writing System: Estonian Extensions to Basic Roman**

Notes: n/a

<i>Char.</i>	Character Name	ISO 8859	ISO/Unic.	<u>Notes/Exceptions</u>
	Acute accent			
A	Roman capital letter A with diaeresis	B4 (1,2,3,4,9,14*) C4 (1,2,3,4,9,10, 12*,13*,14*)	00B4 00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,	00E4	

O	Roman capital letter O with tilde	12*,13*,14*) D5 (1,4,9,10,12*, 13*,14*)	00D5
ø	Roman small letter o with tilde	F5 (1,4,9,10,12*, 13*,14*)	00F5
Ō	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6
o	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6
Š	Roman capital letter S with caron	A9 (2,4), AA (10), DO (13*), B2 (14*)	0160
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E

Number of characters: 13

Table: eu
Writing System: Basque Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	

Number of characters: 6

Table: fa-R
Writing System: Perso-Arabic [Persian]
Romanization System: UN 1%7 (I/13)
Notes: 3a, 4, 5, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Inverted apostrophe	-	MBB (2018)	
À	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ĥ	Roman capital letter H with cedilla		1E28	
ĥ	Roman small letter h with cedilla	-	1E29	

İ	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A
ı	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B
Š	(Roman capital letter S with macron)		0053+0304
š	(Roman small letter s with macron)		0073+0304
Ş	Roman capital letter S with cedilla	AA (2,3), DE (9)	015E
ş	Roman small letter s with cedilla	BA (2,3), FE (9)	015F
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B
Ț	(Roman capital letter Z with macron)		005A + 0304
ț	(Roman small letter z with macron)		007A + 0304
Ț	(Roman capital letter Z with cedilla)	-	005A + 0327
ț	(Roman small letter z with cedilla)	-	007A + 0327
Z	Roman capital letter Z with line below	-	1E94
z	Roman small letter z with line below	-	1E95

Number of characters: 21

Table: fi**Writing System: Finnish Extensions to Basic Roman**

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,12*,13*,14*)	00C4	
a	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,12*,13*,14*)	00E4	
O	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10,12*,13*,14*)	00D6	
o	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10,12*,13*,14*)	00F6	

Number of characters: 4

Table: fo**Writing System: Faroese Extensions to Basic Roman**

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Æ	Roman capital letter AE (ash)	C6 (1,4,9,10,12*,14*), AF (13*)	00C6	
æ	Roman small letter ae (ash)	E6 (1,4,9,10,12*,14*), BF (13*)	00E6	
Ð	Roman capital letter ETH	D0 (1,4,10,14*)	00D0	
ð	Roman small letter eth	F0 (1,4,10,14*)	00F0	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10,12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10,12*,14*)	00ED	

Œ	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
o	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Œ	Roman capital letter O with stroke	D8 (1,4,9,10,12*, 14*)	00D8
ø	Roman small letter o with stroke	F8 (1,4,9,10,12*, 14*)	00F8
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	OODA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	OOFA
Ÿ	Roman capital letter Y with acute	DD (1,2,10,12*,14*)	OODD
Ÿ	Roman small letter y with acute	FD (1,2,10,12*,14*)	OOFD

Number of characters: 16

Table: fr
Writing System: French Extensions to Basic Roman
Notes: n/a

<i>Char.</i>	Character Name	ISO 8859	ISO/Unic.	Notes/Exceptions
A	Roman capital letter A with grave	CO (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
À	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
Ē	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ē	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ë	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10, 12*,14*)	00CB	
e	Roman small letter e with diaeresis	EB (1,2,3,4,9,10, 12*,14*)	00EB	
Î	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE	
î	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE	
Ï	Roman capital letter I with diaeresis	CF (1,3,9,10,12*, 14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*, 14*)	00EF	
Œ	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4	
ø	Roman small letter o with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4	
Œ	Roman capital ligature OE	8C (1,14*)	0152	

œ	Roman small ligature œ	9C (1,14*)	0153
Û	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9
u	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9
Ů	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*,14*)	00DB
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*,14*)	00FB
Ů	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10,12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10,12*,13*,14*)	00FC
Ÿ	Roman capital letter Y with diaeresis	AF (12*)	0178
ÿ	Roman small letter y with diaeresis	FF (1,9,12*)	00FF

Number of characters: 30

Table: fy

Writing-System: **Frisian** Extensions to **Basic Roman**

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ă	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
Ĕ	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ī	Roman capital letter I with circumflex	CE (1,2,3,4,9,10,12*,14*)	00CE	
î	Roman small letter i with circumflex	EE (1,2,3,4,9,10,12*,14*)	00EE	
Ŏ	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10,12*,14*)	00DA	
ü	Roman small letter u with acute	FA (1,2,3,4,9,10,12*,14*)	00FA	
Ů	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*,14*)	00DB	
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*,14*)	00FB	

Number of characters: 12

Table: ga

Writing System: **Irish** Extensions to **Basic Roman**

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
a	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10,	00C9	

é	Roman small letter e with acute	12*,13*,14*) E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9
í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED
ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA

Number of characters: 10

Table: gd

Writing System: Gaelic (Scots) Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	CO (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ì	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC	
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
Ù	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9	
ù	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9	

Number of characters: 16

Table: gl

Writing System: Galician Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	

E	Roman capital letter E with acute	12*,14*) C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED
N	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1 _i
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1 _i
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
U	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
U	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC

Number of characters' 14

Table: gn

Writing System: **Guarani Extensions to Basic Roman**

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
a	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
a	Roman small letter a with tilde	E3 (1,4,9,10,12*, 14*)	00E3	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
ẽ	Roman small letter e with tilde		1EBD	
ĝ	(Roman small letter g with tilde)		0067 +0303	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
ı	Roman small letter i with tilde	B5 (4,10)	0129	
N	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
A	Roman <i>small</i> letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
õ	Roman small letter o with tilde	F5 (1,4,9,10,12*, 13*,14*)	00F5	

U	Roman capital letter U with acute	13*,14*) DA (1,2,3,4,9,10, 12*,14*)	OODA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	OOFA
ũ	Roman small letter u with tilde	FD (4), F7 (10)	0169
ÿ	Roman small letter y with tilde		1EF9
Number of characters: 19			

Table: gu_R
Writing System: Gujarati
Romanization System: UN 1972 (II/11), 1977 (III/12)
Notes: 1, 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
#	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
D	Roman capital letter D with dot below		1E0C	
d	Roman small letter d with dot below		1E0D	
Ē	Roman capital letter E with breve		0114	
ē	Roman small letter e with breve		0115	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ı	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
L	Roman capital letter L with dot below		1E36	
l	Roman small letter l with dot below		1E37	
N	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	OOD1	
A	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	OOF1	
N	Roman capital letter N with dot above		1E44	
n	Roman small letter n with dot above	-	1E45	
N	Roman capital letter N with dot below		1E46	
n	Roman small letter n with dot below		1E47	
Ö	Roman capital letter O with breve		014E	
ö	Roman small letter o with breve		014F	
R	Roman capital letter R with dot below		1E5A	
r	Roman small letter r with dot below	-	1E5B	
S	Roman capital letter S with dot below	-	1E62	
s	Roman small letter s with dot below		1E63	
T	Roman capital letter T with dot below		1E6C	
t	Roman small letter t with dot below		1E6D	
U	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
u	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Number of characters: 26				

Table: ha
Writing System: Hausa Extensions to Basic Roman
Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
B	Roman capital letter B with hook		0181	
6	Roman small letter b with hook		0253	

Đ	Roman capital letter Đ with hook	018A
đ	Roman small letter đ with hook	0257
K	Roman capital letter K with hook	0198
k	Roman small letter k with hook	0199
Number of characters: 6		

Table: hawa

Writing System: **Hawaiian** Extensions to Basic Roman

Notes: 3a, 6

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Inverted apostrophe		02BB (2018)	
A	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ō	Roman capital letter O with macron	D2 (4,10), D4 (13*)	014C	
ō	Roman small letter o with macron	F2 (4,10), F4 (13*)	014D	
Number of characters: 5				

Table: he-R

Writing System: Hebrew

Romanization System: **UN 1977 (III/13)**

Notes: 3a, 4, 5

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Inverted apostrophe		02BB (2018)	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
H	(Roman capital letter H with line below)		0048+0331	
h	Roman small letter h with line below		1E96	
Z	Roman capital letter Z with line below		1E94	
z	Roman small letter z with line below		1E95	
Number of characters: 7				

Table: hi-R

Writing System: Devanagari [**Hindi**]Romanization System: **UN 1972 (II/11), 1977 (III/12)**

Notes: 1, 3a, 3b, 4, 5

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ā	(Roman capital letter A with macron and breve)		0100+0306	
ạ̄	(Roman small letter A with macron and breve)	-	0101+0306	
ĀI	Roman capital letters AI with breve			
āi	Roman small letters ai with breve			
Ḍ	Roman capital letter D with dot below	-	1E0C	
ḍ	Roman small letter d with dot below		1E0D	

Ě	Roman capital letter E with breve		0114
ě	Roman small letter e with breve		0115
Ḥ	Roman capital letter H with dot below		1E24
ḥ	Roman small letter h with dot below		1E25
İ	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A
ı̇	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B
Ṁ	Roman capital letter M with dot above	B6 (12*)	1E40
ṁ	Roman small letter m with dot above	B7 (12*)	1E41
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1
Ñ	Roman capital letter N with dot above		1E44
ṅ	Roman small letter n with dot above		1E45
Ṇ	Roman capital letter N with dot below		1E46
ṇ	Roman small letter n with dot below		1E47
Ö	Roman capital letter O with breve		014E
ö	Roman small letter o with breve		014F
Ṙ	Roman capital letter R with dot above		1E58
ṙ	Roman small letter r with dot above		1E59
Ṛ	Roman capital letter R with dot below		1E5A
ṛ	Roman small letter r with dot below		1E5B
Ṣ	Roman capital letter S with dot below	-	1E62
ṣ	Roman small letter s with dot below	-	1E63
Ṫ	Roman capital letter T with dot below		1E6C
ṫ	Roman small letter t with dot below		1E6D
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
u̇	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B

Number of characters: 34

Table: hr

Writing System: Croatian Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ć	Roman capital letter C with acute	C6 (2), C3 (13)	0106	
ć	Roman small letter c with acute	E6 (2), E3 (13*)	0107	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Đ	Roman capital letter D with stroke	D0 (2), A9 (10), A3 (14*)	0110	
đ	Roman small letter d with stroke	F0 (2), B9 (10), A4 (14*)	0111	
Š	Roman capital letter S with w o n	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10),	017E	

FE (13*), BA (14*)

Number of characters: 10

Table: hu**Writing System: Hungarian Extensions to Basic Roman**

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
a	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
i	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
o	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
Ő	Roman capital letter O with double acute	D5 (2)	0150	
ó	Roman small letter o with double acute	F5 (2)	0151	
U	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA	
u	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	
Û	Roman capital letter U with double acute	DB (2)	0170	
ű	Roman small letter u with double acute	FB (2)	0171	

Number of characters: 18

Table: ibo**Writing System: Ibo Extensions to Basic Roman**

Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
I	Roman capital letter I with dot below	-	1ECA	
ı	Roman small letter i with dot below	-	1ECB	
Ɔ	Roman capital letter O with dot below		1ECC	
o	Roman small letter o with dot below		1ECD	
U	Roman capital letter U with dot below		1EE4	

u **Roman** small letter u with dot below
Number of characters: 6

IEE5

Table: id
Writing System: Indonesian Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	

Number of characters: 2

Table: is
Writing System: Icelandic Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
Æ	Roman capital letter AE (ash)	C6 (1,4,9,10,12*, 14*), AF (13*)	00C6	
æ	Roman small letter ae (ash)	E6 (1,4,9,10,12*, 14*), BF (13*)	00E6	
Ð	Roman capital letter ETH	D0 (1,4,10,14*)	00D0	
ð	Roman small letter eth	F0 (1,4,10,14*)	00F0	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
Þ	Roman capital letter Thorn (Icelandic)	DE (1,10,14*)	00DE	
þ	Roman small letter thorn (Icelandic)	FE (1,10,14*)	00FE	
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA	
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA	
Ý	Roman capital letter Y with acute	DD (1,2,10,12*,14*)	00DD	
ý	Roman small letter y with acute	FD (1,2,10,12*,14*)	00FD	

Number of characters: 20

Table: it

Writing System: Italian Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	00 (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
Ē	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9 ⁱ	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ì	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC	
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC	
Ī	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
Ū	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9	
ù	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9	
Ū	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA	
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA	

Number of characters: 18

Table: ja_R1

Writing System: Sino-Japanese [Japanese]

Romanization System: Modified Hepburn

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ō	Roman capital letter O with macron	D2 (4,10), D4 (13*)	014C	
ō	Roman small letter o with macron	F2 (4,10), F4 (13*)	014D	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
u	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	

Number of characters: 4

Table: ja_R2

Writing System: Sino-Japanese [Japanese]

Romanization System: Kunrei

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ō	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,	00D4	

6	Roman small letter o with circumflex	12*,14*) F4 (1,2,3,4,9,10,	00F4
U	Roman capital letter U with circumflex	12*,14*) DB (1,3,4,9,10,12*,	OODB
	Roman small letter u with circumflex	14*) FB (1,3,4,9,10,12*,	00FB
		14*)	

Number of characters: 4

Table: kk_R
Writing System: Cyrillic [Kazakh]
Romanization System: BGN/PCGN 1979
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,	00C4	
ä	Roman small letter a with diaeresis	12*,13*,14*) E4 (1,2,3,4,9,10,	00E4	
É	Roman capital letter E with dot above	12*,13*,14*) CC (4,10), CB (13*)	0116	
é	Roman small letter e with dot above	EC (4,10), EB (13*)	0117	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE	012A	
ī	Roman small letter i with macron	(13*) EF (4), B4 (10), EE	012B	
Ö	Roman capital letter O with diaeresis	(13*) D6 (1,2,3,4,9,10,	00D6	
ö	Roman small letter o with diaeresis	12*,13*,14*) F6 (1,2,3,4,9,10,	00F6	
Ū	Roman capital letter U with macron	12*,13*,14*) DE (4), AE (10), DB	016A	
ū	Roman small letter u with macron	(13*) FE (4), BE (10), FB	016B	
Û	Roman capital letter U with diaeresis	(13*) DC (1,2,3,4,9,10,	00DC	
ü	Roman small letter u with diaeresis	12*,13*,14*) FC (1,2,3,4,9,10,	00FC	
		12*,13*,14*)		

Number of characters: 12

Table: km_R
Writing System: Khmer
Romanization System: UN 1972 (II/10)
Notes: 2, 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with acute	C1 (1,2,3,4,9,10,	00C1	
ā	Roman small letter a with acute	12*,14*) E1 (1,2,3,4,9,10,	00E1	
Ă	Roman capital letter A with circumflex	12*,14*) C2 (1,2,3,4,9,10,	00C2	
ă	Roman small letter a with circumflex	12*,14*) E2 (1,2,3,4,9,10,	00E2	
Ȧ	Roman capital letter A with breve	12*,14*) C3 (2)	0102	

ă	Roman small letter a with breve	E3 (2)	0103
Ě	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA
Ĕ	Roman capital letter E with breve		0114
ĕ	Roman small letter e with breve		0115
Ĭ	Roman capital letter I with breve		012C
ī	Roman small letter i with breve		012D
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4
Ö	Roman capital letter O with breve		014E
ö	Roman small letter o with breve		014F
Œ	Roman capital ligature OE	8C (1,14*)	0152
œ	Roman small ligature oe	9C (1,14*)	0153
Ě	(Roman capital ligature OE with breve)		0152+0306
œ	(Roman small ligature oe with breve)		0153+0306
Ū	Roman capital letter U with breve	DD (3)	016C
ū	Roman small letter u with breve	FD (3)	016D

Number of characters: 26

Table: kn_R

Writing System: Kannada

Romanization System: UN 1972 (II/11), 1977 (III/12)

Notes: 1, 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
D	Roman capital letter D with dot below		1E0C	
d	Roman small letter d with dot below	-	1E0D	
Ĕ	Roman capital letter E with breve		0114	
ĕ	Roman small letter e with breve		0115	
H	Roman capital letter H with dot below		1E24	
h	Roman small letter h with dot below		1E25	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
L	Roman capital letter L with dot below		1E36	
l	Roman small letter l with dot below		1E37	
M	Roman capital letter M with dot above	B6 (12*)	1E40	
ṁ	Roman small letter m with dot above	B7 (12*)	1E41	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ṅ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ñ	Roman capital letter N with dot above		1E44	

ñ	Roman small letter n with dot above		1E45
N	Roman capital letter N with dot below		1E46
ņ	Roman small letter n with dot below		1E47
Ō	Roman capital letter O with breve		014E
ō	Roman small letter o with breve		014F
Ṙ	Roman capital letter R with dot above		1E58
ř	Roman small letter r with dot above		1E59
Ṛ	Roman capital letter R with dot below		1E5A
ṛ	Roman small letter r with dot below		1E5B
Ṡ	Roman capital letter S with dot below		1E62
ṡ	Roman small letter s with dot below		1E63
Ṫ	Roman capital letter T with dot below		1E6C
ṫ	Roman small letter t with dot below		1E6D
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B

Number of characters: 32

Table: ko_R1
Writing System: Hangul [Korean]
Romanization System: McCune Reischauer 1939
Notes: 3a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ō	Roman capital letter O with breve		014E	
ō	Roman small letter o with breve		014F	
Ū	Roman capital letter U with breve	DD (3)	016C	
ū	Roman small letter u with breve	FD (3)	016D	

Number of characters: 4

Table: ko_R2
Writing System: Hangul [Korean]
Romanization System: Ministry of Education (Rep. of Korea) System 1984
Notes: 3a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ō	Roman capital letter O with breve	-	014E	
ō	Roman small letter o with breve	-	014F	
Ū	Roman capital letter U with breve	DD (3)	016C	
	Roman small letter u with breve	FD (3)	016D	

Number of characters: 4

Table: ky_R
Writing System: Cyrillic [Kyrgyz]
Romanization System: BGN/PCGN 1979
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10,	00DC	

ü	Roman small letter u with diaeresis	12*,13*,14*) FC (1,2,3,4,9,10, 12*,13*,14*)	OOFC
---	-------------------------------------	---	------

Number of characters: 4

Table: lb

Writing System: Luxembourgian Extensions to Basic Roman

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10, 12*,13*,14*)	00C4	
a	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10, 12*,13*,14*)	00E4	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
E	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10, 12*,14*)	00CB	
e	Roman small letter e with diaeresis	EB (1,2,3,4,9,10, 12*,14*)	00EB	
Ŋ	(Roman capital letter M with circumflex)		004D+0302	
rn	(Roman small letter m with circumflex)		006D+0302	
Ñ	(Roman capital letter N with circumflex)	-	004E+0302	
ñ	(Roman small letter n with circumflex)		006E+0302	
O	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	0006	
o	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
U	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	

Number of characters: 14

Table: livo

Writing System: Livonian Extensions to Basic Roman

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10, 12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10, 12*,13*,14*)	00E4	
Ǟ	Roman capital letter A with diaeresis and macron		01DE	
ā̄	Roman small letter a with diaeresis and macron		01DF	
Ḑ	Roman capital letter D with cedilla		1E10	

đ	Roman small letter d with cedilla		1E11
Ě	Roman capital letter E with macron	AA (4), A2 (10), C7 (13*)	0112
ē	Roman small letter e with macron	BA (4), B2 (10), E7 (13*)	0113
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B
Ľ	Roman capital letter L with cedilla	A6 (4), A8 (10), CF (13*)	013B
ĺ	Roman small letter l with cedilla	B6 (4), B8 (10), EF (13*)	013C
Ń	Roman capital letter N with cedilla	D1 (4,10), D2 (13*)	0145
ņ	Roman small letter n with cedilla	F1 (4,10), F2 (13*)	0146
Õ	Roman capital letter O with tilde	D5 (1,4,9,10,12*,13*,14*)	00D5
ó	Roman small letter o with tilde	F5 (1,4,9,10,12*,13*,14*)	00F5
Ō	(Roman capital letter O with tilde and macron)		00D5+0304
ō	(Roman small letter o with tilde and macron)	-	00F5+0304
Ŏ	Roman capital letter O with macron	D2 (4,10), D4 (13*)	014C
ō	Roman <i>small</i> letter o with macron	F2 (4,10), F4 (13*)	014D
Ȫ	(Roman capital letter O with dot above)		004F+0307
ȫ	(Roman small letter o with dot above)	-	006F+0307
Ȭ	(Roman capital letter O with dot above and macron)	-	004F+0307+0304
ȭ	(Roman small letter o with dot above and macron)		006F+0307+0304
Ŕ	Roman capital letter R with cedilla	A3 (4), AA (13*)	0156
ŗ	Roman small letter r with cedilla	B3 (4), BA (13*)	0157
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161
Ţ	Roman capital letter T with cedilla	DE (2)	0162
ţ	Roman small letter t with cedilla	FE (2)	0163
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D
ž	Roman small letter z with caron	BE (2,4), BC (10), FE? (13*), BA (14*)	017E

Number of characters: 36

Table: **lo-R****Writing System: Laotian****Romanization System: National**

Notes: n/a

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
E	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ĕ	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4	

Number of characters: 6

Table: It

Writing System: Lithuanian Extensions to Basic Roman

Notes: n/a

Ą	Roman capital letter A with ogonek	A1 (2,4,10), Ć (13*)	0104	
ą	Roman small letter a with ogonek	B1 (2,4,10), E0 (13*)	0105	
Ć	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Ė	Roman capital letter E with dot above	CC (4,10), CB (13*)	0116	
ė	Roman small letter e with dot above	EC (4,10), EB (13*)	0117	
Ę	Roman capital letter E with ogonek	CA (2,4,10), C6 (13*)	0118	
ę	Roman small letter e with ogonek	EA (2,4,10), E6 (13*)	0119	
Į	Roman capital letter I with ogonek	C7 (4,10), C1 (13*)	012E	
į	Roman small letter i with ogonek	E7 (4,10), E1 (13*)	012F	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Ų	Roman capital letter U with ogonek	D9 (4,10), D8 (13*)	0172	
ų	Roman small letter u with ogonek	F9 (4,10), F8 (13*)	0173	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 18

Table: lv
 Writing System: **Latvian** Extensions to Basic **Roman**
 Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Ē	Roman capital letter E with macron	AA (4), A2 (10), C7 (13*)	0112	
ē	Roman small letter e with macron	BA (4), B2 (10), E7 (13*)	0113	
Ģ	Roman small letter g with <i>turned comma</i> above	BB (4), B3 (10), EC (13*)	0123	
Ģ	Roman capital letter G with cedilla	AB (4), A3 (10), CC (13*)	0122	Usual typesetting is Ģ (0047 + 0326)
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ķ	Roman capital letter K with cedilla	D3 (4), A6 (10), CD (13*)	0136	Usual typesetting is Ķ (004B + 0326)
ķ	Roman small letter k with cedilla	F3 (4), B6 (10), ED (13*)	0137	Usual typesetting is ķ (006B + 0326)
Ļ	Roman capital letter L with cedilla	A6 (4), A8 (10), CF (13*)	013B	Usual typesetting is Ļ (004C + 0326)
ļ	Roman small letter l with cedilla	B6 (4), B8 (10), EF (13*)	013C	Usual typesetting is ļ (006C + 0326)
Ņ	Roman capital letter N with cedilla	D1 (4,10), D2 (13*)	0145	Usual typesetting is Ņ (004E + 0326)
ņ	Roman small letter n with cedilla	F1 (4,10), F2 (13*)	0146	Usual typesetting is ņ (006E + 0326)
Š	Roman capital letter S with caron	A9 (2,4), AA (10), DO (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), FO (13*), B3 (14*)	0161	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 22

Table: **mars**Writing System: **Marshallese Extensions to Basic Roman**

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ļ	Roman capital letter L with cedilla	A6 (4), A8 (10), CF (13*)	013B	
ļ	Roman small letter l with cedilla	B6 (4), B8 (10), EF (13*)	013d	
Ṁ	(Roman capital letter M with cedilla)		004D+0327	
ṁ	(Roman small letter m with cedilla)		006D+0327	
Ñ	(Roman capital letter N with macron)		004E+0304	
ñ	(Roman small letter n with macron)		006E+0304	
Ņ	Roman capital letter N with cedilla	D1 (4,10), D2 (13*)	0145	
ņ	Roman small letter n with cedilla	F1 (4,10), F2 (13*)	0146	
Ō	Roman capital letter O with macron	D2 (4,10), D4 (13*)	014C	
ō	Roman small letter o with macron	F2 (4,10), F4 (13*)	014D	
Ȯ	(Roman capital letter O with cedilla)		004F+0327	
ȯ	(Roman small letter o with cedilla)		006F+0327	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	

Number of characters: 16

Table: **mg**Writing System: **Malagasy Extensions to Basic Roman**

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Ñ	(Roman capital letter N with diaeresis)		W E+0308	
ñ	(Roman small letter n with diaeresis)		006E+0308	
Ȫ	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	

Number of characters: 6

Table: **mī**Writing System: **Maori Extensions to Basic Roman**

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ē	Roman capital letter E with macron	AA (4), A2 (10), C7 (13*)	0112	

ē	Roman small letter e with macron	BA (4), B2 (10), E7 (13*)	0113
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B
Ō	Roman capital letter O with macron	D2 (4,10), D4 (13*)	014C
ō	Roman small letter o with macron	F2 (4,10), F4 (13*)	014D
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B

Number of characters: 10

Table: mk_R
 Writing System: Cyrillic [Macedonian]
 Romanization System: UN 1977 (III/11)
 Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ĉ	Roman capital letter C with acute	C6 (2), C3 (13*)	0106	
ĉ	Roman small letter c with acute	E6 (2), E3 (13*)	0107	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
Ď	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
D̄	Roman capital letter D with stroke	D0 (2), A9 (10), A3 (14*)	0110	
d̄	Roman small letter d with stroke	F0 (2), B9 (10), A4 (14*)	0111	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 10

Table: ml_R
 Writing System: Malayalam
 Romanization System: UN 1972 (II/11), 1977 (III/12)
 Notes: 1, 3a, 4

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
Ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Đ	Roman capital letter D with dot below		1E0C	
đ	Roman small letter d with dot below		1E0D	
Ē	Roman capital letter E with breve		0114	
ē	Roman small letter e with breve		0115	
Ĥ	Roman capital letter H with dot below		1E24	
ĥ	Roman small letter h with dot below		1E25	

Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B
L	Roman capital letter L with dot below	-	1E36
l	Roman small letter l with dot below	-	1E37
Ł	(Roman capital letter L with-line below)		1E3A
ł	(Roman small letter l with-line below)		1E3B
Ṁ	Roman capital letter M with dot above	B6 (12*)	1E40
ṁ	Roman small letter m with dot above	B7 (12*)	1E41
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1
Ñ̇	Roman capital letter N with dot above		1E44
ṅ	Roman small letter n with dot above		1E45
Ṇ	Roman capital letter N with dot below	-	1E46
ṇ	Roman small letter n with dot below	-	1W7
Ŏ	Roman capital letter O with breve	-	014E
ö	Roman small letter o with breve		014F
Ṙ	Roman capital letter R with dot below		1E5A
ṙ	Roman small letter r with dot below		1E5B
Ṛ	Roman capital letter R with line below		1E5E
ṛ	Roman small letter r with line below		1E5F
Ṡ	Roman capital letter S with dot below	-	1E62
ṡ	Roman small letter s with dot below	-	1E63
Ṫ	(Roman capital letter T with diaeresis and dot below)		1E6C+0308
ṽ	(Roman small letter t with diaeresis and dot below)	-	1E6D+0308 (1E97+0323)
Ṫ	Roman capital letter T with dot below		1E6C
ṫ	Roman small letter t with dot below		1E6D
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B

Number of characters: 36

Table: mn_R

Writing System: Cyrillic [Mongolian]

Romanization System: BGN/PCGN 1964

Notes: 3a

Char.	Character Name	ISO 8859	ISO/Unic.	Notes/Exceptions
Ī	Roman capital letter I with breve		012C	
ī	Roman small letter i with breve		012D	
İ	Roman capital letter I with diaeresis	CF (1,3,9,10,12*,14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*,14*)	00EF	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10,12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	€6(1,2,3,4,9,10,12*,13*,14*)	00F6	
Ū	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10,	00DC	

ii	Roman small letter u with diaeresis	12*,13*,14*) FC (1,2,3,4,9,10, 12*,13*,14*)	OOFC
----	-------------------------------------	---	------

Number of characters: 8

Table: mr_R
Writing System: Marathi Extensions to Devanagari
Romanization System: UN 1972 (II/11), 1977 (III/12)
Notes: 1, 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
L	Roman capital letter L with dot below	-	1E36	
l	Roman small letter l with dot below		1E37	

Number of characters: 2

Table: mt
Writing System: Maltese Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with grave	CO (1,3,9,12*,14*)	00C0	
a	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Ā	Roman capital letter C with dot above	C5 (3), A5 (12*)	010A	
ċ	Roman small letter c with dot above	E5 (3), A6 (12*)	010B	
E	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
Ġ	Roman capital letter G with dot above	D5 (3), B4 (12*)	0120	
ġ	Roman small letter g with dot above	F5 (3), B5 (12*)	0121	
H	Roman capital letter H with stroke	A1 (3)	0126	
h	Roman small letter h with stroke	B1 (3)	0127	
Ī	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC	
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC	
Ī	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE	
ï	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
U	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9	
u	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9	
Ž	Roman capital letter Z with dot above	AF (2,3), DD (13*)	017B	
ž	Roman small letter z with dot above	BF (2,3), FD (13*)	017C	

Number of characters: 20

Table: my_R
Writing System: Burmese
Romanization System: BGN/PCGN 1970
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
Ō	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4	

6 **Roman** small letter o with circumflex F4 (1,2,3,4,9,10, 12*,14*) 00F4

Number of characters: 4

Table: ne_R

Writing System: **Devanagari** [Nepali]

Romanization System: **BGN/PCGN 1964**

Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
D	Roman capital letter D with dot below		1E0C	
d	Roman <i>small</i> letter d with dot below		1E0D	
N	Roman capital letter N with acute	D1 (2,13*)	0143	
n	Roman <i>small</i> letter n with acute	F1 (2,13*)	0144	
N	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman <i>small</i> letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
N	Roman capital letter N with dot below		1E46	
n	Roman <i>small</i> letter n with dot below		1E47	
Š	Roman capital letter S with acute	A6 (2), DA (13*)	015A	
s	Roman <i>small</i> letter s with acute	B6 (2), FA (13*)	015B	
S	Roman capital letter S with dot below		1E62	
s	Roman <i>small</i> letter s with dot below		1E63	
T	Roman capital letter T with dot below		1E6C	
t	Roman <i>small</i> letter t with dot below	-	1E6D	

Number of characters: 14

Table: nl

Writing System: Dutch Extensions **to** Basic Roman

Notes: 3a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
À	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	
E	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	WE9	
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman <i>small</i> letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
E	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10, 12*,14*)	00CB	
ë	Roman <i>small</i> letter e with diaeresis	EB (1,2,3,4,9,10, 12*,14*)	00EB	
İ	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
ı	Roman <i>small</i> letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	

I	Roman capital letter I with diaeresis	12*,14*) CF (1,3,9,10,12*, 14*)	00CF
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*, 14*)	00EF
IJ	Roman capital ligature IJ		0132
ij	Roman small ligature ij		0133
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6
U	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
u	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
Û	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*, 14*)	00DB
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*, 14*)	00FB

Number of characters: 28

Table: no
Writing System: Norwegian (bokmål, nynorsk) Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with ring above	C5 (1,4,9,10,12*, 13*,14*)	00C5	
á	Roman small letter a with ring above	E5 (1,4,9,10,12*, 13*,14*)	00E5	
Æ	Roman capital letter AE (ash)	C6 (1,4,9,10,12*, 14*), AF (13*)	00C6	
æ	Roman small letter ae (ash)	E6 (1,4,9,10,12*, 14*), BF (13*)	00E6	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4	
Ø	Roman capital letter O with stroke	D8 (1,4,9,10,12*, 14*)	00D8	

o **Roman** small letter o with stroke F8 (1,4,9,10,12*, 14*) 00F8

Number of characters: 12

Table: or-R

Writing System: Oriya

Romanization System: UN 1972 (II/11), 1977 (III/12)

Notes: 1, 3a, 5

<i>Char.</i>	Character Name	ISO 8859	ISO/Unic.	Notes/Exceptions
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ḑ	Roman capital letter D with dot below	-	1E0C	
ḑ	Roman small letter d with dot below	-	1E0D	
Ḥ	Roman capital letter H with dot below	-	1E24	
ḥ	Roman small letter h with dot below	-	1E25	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ĺ	Roman capital letter L with dot below		1E36	
ĺ	Roman small letter l with dot below		1E37	
Ṁ	Roman capital letter M with dot above	B6 (12*)	1E40	
ṁ	Roman small letter m with dot above	B7 (12*)	1E41	
Ṃ	(Roman capital letter M with candrabindu)		004D+0310	
ṃ	(Roman small letter m with candrabindu)		006D+0310	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ñ̇	Roman capital letter N with dot above		1E44	
ṅ	Roman small letter n with dot above	-	1E45	
Ṇ̃	Roman capital letter N with dot below		1E46	
ṇ	Roman small letter n with dot below		1E47	
Ŕ	Roman capital letter R with dot above		1E58	
ŕ	Roman small letter r with dot above	-	1E59	
Ŗ	Roman capital letter S with dot below		1E62	
ŗ	Roman small letter s with dot below		1E63	
Ṭ	Roman capital letter T with dot below		1E6C	
ṭ	Roman small letter t with dot below	-	1E6D	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Ÿ	Roman capital letter Y with dot below	-	1EF4	
ÿ	Roman small letter y with dot below		1EF5	

Number of characters: 30

Table: pa-R
Writing System: Punjabi
Romanization System: UN 1972 (II/11), 1977 (III/12)
Notes: 1, 2, 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ḑ	Roman capital letter D with dot below		1E0C	
ḑ	Roman small letter d with dot below		1E0D	
Ḕ	(Roman capital letter G with dot below)		0047+0323	
ḕ	(Roman small letter g with dot below)		0067+0323	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ḳ	Roman capital letter K with dot below		1E32	
ḳ	Roman small letter k with dot below		1E33	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ṅ	Roman capital letter N with dot above		1E44	
ṅ	Roman small letter n with dot above		1E45	
Ṇ	Roman capital letter N with dot below		1E46	
ṇ	Roman small letter n with dot below		1E47	
Ṛ	Roman capital letter R with dot above		1E58	
ṛ	Roman small letter r with dot above		1E59	
Ṛ	Roman capital letter T with dot below		1E6C	
ṛ	Roman small letter t with dot below		1E6D	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	

Number of characters: 22

Table: pl
Writing System: Polish Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ą	Roman capital letter A with ogonek	A1 (2,4,10), CO (13*)	0104	
ą	Roman small letter a with ogonek	B1 (2,4,10), E0 (13*)	0105	
Ć	Roman capital letter C with acute	C6 (2), C3 (13*)	0106	
ć	Roman small letter c with acute	E6 (2), E3 (13*)	0107	
Ę	Roman capital letter E with ogonek	CA (2,4,10), C6 (13*)	0118	
ę	Roman small letter e with ogonek	EA (2,4,10), E6 (13*)	0119	
Ł	Roman capital letter L with stroke	A3 (2), D9 (13*)	0141	
ł	Roman small letter l with stroke	B3 (2), F9 (13*)	0142	
Ń	Roman capital letter N with acute	D1 (2,13*)	0143	
ń	Roman small letter n with acute	F1 (2,13*)	0144	

Ŏ	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
6	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Ś	Roman capital letter S with acute	A6 (2), DA (13*)	015A
s	Roman small letter s with acute	B6 (2), FA (13*)	015B
Ź	Roman capital letter Z with acute	AC (2), CA (13*)	0179
z	Roman small letter z with acute	BC (2), EA (13*)	017A
Ź	Roman capital letter Z with dot above	AF (2,3), DD (13*)	017B
ž	Roman small letter z with dot above	BF (2,3), FD (13*)	017C

Number of characters: 18

Table: ps_R

Writing System: Perso-Arabic [Pashto]

Romanization System: BGN/PCGN 1%8

Notes: 3a, 4, 5, 6

<i>Char.</i>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
	Inverted apostrophe		02BB (2018)	
A	Roman capital letter A with macron	CO (4,10), C2 (13*)	0100	
a	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
D	Roman capital letter D with line below		1E0E	
d	Roman small letter d with line below		1E0F	
E	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ĥ	Roman capital letter H with cedilla		1E28	
ĥ	Roman small letter h with cedilla		1E29	
ı	Roman capital letter ı with macron	CF (4), A4 (10), CE (13*)	012A	
ı̇	Roman small letter ı with macron	EF (4), B4 (10), EE (13*)	012B	
N	Roman capital letter N with line below		1E48	
n	Roman small letter n with line below	-	1E49	
R	Roman capital letter R with line below		1E5E	
r	Roman small letter r with line below		1E5F	
Š	(Roman capital letter S with macron)		0053+0304	
š	(Roman small letter s with macron)		0073+0304	
Ş	Roman capital letter S with cedilla	AA (2,3), DE (9)	015E	
ş	Roman small letter s with cedilla	BA (2,3), FE (9)	015F	
SH	(Roman capital letters SH with line below)		0053+0332 +0048 +0332	
Sh	(Roman capital letter S and small letter h with line below)	-	0053+0332 +0068+0332	
	(Roman small letters sh with line below)		0073+0332 +0068+0332	
T	Roman capital letter T with line below		1E6E	
t	Roman small letter t with line below	-	1E6F	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ü	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Z̄	(Roman capital letter Z with macron)		005A +0304	
z̄	(Roman small letter z with macron)		007A +0304	

Ẓ	(Roman capital letter Z with cedilla)	005A+0327
ẓ	(Roman small letter z with cedilla)	007A+0327
Z̅	Roman capital letter Z with line below	1E94
z̅	Roman small letter z with line below	1E95
Z̅H̅	(Roman capital letters ZH with line below)	005A+0332 +0048+0332
Z̅h̅	(Roman capital letter Z and small letter h with line below)	005A+0332 +0068+0332
z̅h̅	(Roman small letters zh with line below)	007A+0332 +0068+0332

Number of characters: 35

Table: pt
Writing System: Portuguese Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	CO (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Â	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
Ã	Roman capital letter A with tilde	C3 (1,4,9,10,12*,14*)	00C3	
ã	Roman small letter a with tilde	E3 (1,4,9,10,12*,14*)	00E3	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	WE9	
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10,12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10,12*,14*)	00ED	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*,13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*,13*,14*)	00F3	
Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	

Õ	Roman capital letter O with tilde	12*,14*) D5 (1,4,9,10,12*, 13*,14*)	00D5
õ	Roman small letter o with tilde	F5 (1,4,9,10,12*, 13*,14*)	00F5
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
U	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC

Number of characters: 30

Table: rm

Writing System: (Rhaeto-)Romance Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
A	Roman capital letter A with grave	CO (1,3,9,12*,14*)	00C0	
a	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
E	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
e	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
E	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
I	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC	
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC	
Î	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE	
ı	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE	
O	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
U	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9	
ù	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9	

Number of characters: 14

Table: ro

Writing System: Romanian Extensions to Basic Roman

Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ă	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	
Ȧ	Roman capital letter A with breve	C3 (2)	0102	
ă	Roman small letter a with breve	E3 (2)	0103	
Î	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE	
ı	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE	

		12*,14*)		
Ş	Roman capital letter S with comma below	-	0218	
ş	Roman small letter s with comma below	-	0219	
Ș	Roman capital letter S with cedilla	AA (2,3), DE (9)	015E	Correct typesetting: Ș (0053+0326)
ș	Roman small letter s with cedilla	BA (2,3), FE (9)	015F	Correct typesetting: ș (0073+0326)
Ț	Roman capital letter T with comma below	-	021A	
ț	Roman small letter t with comma below	-	021B	
Ț	Roman capital letter T with cedilla	DE (2)	0162	Correct typesetting: Ț (0054+0326)
ț	Roman small letter t with cedilla	FE (2)	0163	Correct typesetting: ț (0074+0326)

Number of characters: 14

*Note: at the request of Romania (accepted at the ISO/IEC JTC1/SC2/WG2 meeting on Crete, Greece, 30.06-04.07.1997) new ISO 10646 codes (marked with an asterisk) will be assigned and new character names designated for characters formerly associated with cedillas

Table: ru_R1
Writing System: Cyrillic [Russian]
Romanization System: UN 1987 (V/18)
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
"	Quotation mark	34 (1..14)	0034	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
Ë	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10,12*,14*)	00CB	
ë	Roman small letter e with diaeresis	EB (1,2,3,4,9,10,12*,14*)	00EB	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 11

Table: ru_R2**Writing System: Cyrillic [Russian]****Romanization System: BGN/PCGN 1947****Notes: n/a**

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
"	Quotation mark	34 (1..14)	0034	
Ë	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10, 12*,14*)	00CB	
ë	Roman small letter e with diaeresis	EB (1,2,3,4,9,10, 12*,14*)	00EB †	

Number of characters: 3

Table: ru_R3**Writing System: Cyrillic [Russian]****Romanization System: Russian Academy of Sciences System****Notes: n/a**

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
ċ	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 6

Table: sámI**Writing System: Sámi (Northern, Inari, Southern, Luleá) Extensions to Basic Roman****Notes: n/a**

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
Ā	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
ċ	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Ð	Roman capital letter D with stroke	D0 (2), A9 (10), A3 (14*)	0110	
đ	Roman small letter d with stroke	F0 (2), B9 (10), A4 (14*)	0111	
Ŋ	Roman capital letter ENG (Sami)	BD (4), AF (10,14*)	014A	
ŋ	Roman small letter eng (Sami)	BF (4,10), B1 (14*)	014B	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	

š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161
Ʀ	Roman capital letter T with stroke	AC (4), AB (10), B5 (14*)	0166
Ƨ	Roman small letter t with stroke	BC (4), BB (10), B8 (14*)	0167
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E

Number of characters: 14

Table: sámí2

Writing System: Skolt Sámi Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	C0 (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Ā	Roman capital letter A with breve	C3 (2)	0102	
ā	Roman small letter a with breve	E3 (2)	0103	
Ą	Roman capital letter A with dot above and macron	-	01E0	
ą	Roman small letter a with dot above and macron	-	01E1	
Ȁ	Roman capital letter A with diaeresis and macron	-	01DE	
ȁ	Roman small letter a with diaeresis and macron	-	01DF	
Æ	Roman capital letter AE (ash) with macron	-	01E2	
æ	Roman small letter ae (ash) with macron	-	01E3	
Ʒ	Roman capital letter EZH	BC (14*)	01B7	
Ʒ	Roman small letter ezh	BD (14*)	0292	
Š	Roman capital letter EZH with caron	BE (14*)	01EE	
š	Roman small letter ezh with caron	BF (14*)	01EF	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
Ē	Roman capital letter E with breve	-	0114	
ē	Roman small letter e with breve	-	0115	
Ġ	Roman capital letter G with caron	A8 (14*)	01E6	
ġ	Roman small letter g with caron	AA (14*)	01E7	
Ɔ	Roman capital letter G with stroke	A5 (14*)	01E4	
Ƨ	Roman small letter g with stroke	A6 (14*)	01E5	
Ķ	Roman capital letter K with caron	AC (14*)	01E8	
ķ	Roman small letter k with caron	AE (14*)	01E9	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
Ȫ	Roman capital letter O with breve	-	014E	
ȫ	Roman small letter o with breve	-	014F	
Ȭ	Roman capital letter O with ogonek	-	01EA	
ȭ	Roman small letter o with ogonek	-	01EB	
Ȯ	Roman capital letter O with ogonek	-	01EC	

and macron
 ō Roman small letter o with ogonek and macron - 01ED

Number of characters: 32

Table: sk

Writing System: Slovak Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,12*,13*,14*)	00E4	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Ď	Roman capital letter D with caron	CF (2)	010E	
ď	Roman small letter d with caron	EF (2)	010F	Usual typesetting is đ (0064+0315)
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10,12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10,12*,14*)	00ED	
Ĺ	Roman capital letter L with acute	C5 (2)	0139	
ĺ	Roman small letter l with acute	E5 (2)	013A	
Ľ	Roman capital letter L with caron	A5 (2)	013D	Usual typesetting is Ľ (004C+0315)
ľ	Roman small letter l with caron	B5 (2)	013E	Usual typesetting is ľ (006C+0315)
Ň	Roman capital letter N with caron	D2 (2)	0147	
ň	Roman small letter n with caron	F2 (2)	0148	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*,13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*,13*,14*)	00F3	
Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4	
ô	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4	
Ř	Roman capital letter R with acute	C0 (2)	0154	
ř	Roman small letter r with acute	E0 (2)	0155	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10),	0161	

Ŧ	Roman capital letter T with caron	F0 (13*), B3 (14*) AB (2)	0164	Usual typesetting is Ŧ (0074+0315)
ť	Roman small letter t with caron	BB (2)	0165	
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA	
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA	
Ý	Roman capital letter Y with acute	DD (1,2,10,12*,14*)	00DD	
ý	Roman small letter y with acute	FD (1,2,10,12*,14*)	00FD	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 34

Table: sl
Writing System: Slovene Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 6

Table: sm
Writing System: Samoan Extensions to Basic Roman
Notes: 3a, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
ˆ	Inverted apostrophe	-	02BB (2018)	

Number of characters: 1

Table: sorb1
Writing System: Lower Sorbian Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ć	Roman capital letter C with acute	C6 (2), C3 (13*)	0106	
ć	Roman small letter c with acute	E6 (2), E3 (13*)	0107	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2	010D	

		(14*)	
Ě	Roman capital letter E with caron	CC (2)	011A
ě	Roman small letter e with caron	EC (2)	011B
Ľ	Roman capital letter L with stroke	A3 (2), D9 (13*)	0141
ĺ	Roman small letter l with stroke	B3 (2), F9 (13*)	0142
Ň	Roman capital letter N with acute	D1 (2,13*)	0143
ň	Roman small letter n with acute	F1 (2,13*)	0144
Ř	Roman capital letter R with acute	C0 (2)	0154
ř	Roman small letter r with acute	E0 (2)	0155
Š	Roman capital letter S with acute	A6 (2), DA (13*)	015A
š	Roman small letter s with acute	B6 (2), FA (13*)	015B
Ŝ	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160
ŝ	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161
Ž	Roman capital letter Z with acute	AC (2), CA (13*)	0179
ž	Roman small letter z with acute	BC (2), EA (13*)	017A
Ẑ	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D
ẑ	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E

Number of characters: 20

Table: sorb2

Writing System: Upper Sorbian Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ć	Roman capital letter C with acute	C6 (2), C3 (13*)	0106	
ć	Roman small letter c with acute	E6 (2), E3 (13*)	0107	
Ć	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Ě	Roman capital letter E with caron	CC (2)	011A	
ě	Roman small letter e with caron	EC (2)	011B	
Ľ	Roman capital letter L with stroke	A3 (2), D9 (13*)	0141	
ĺ	Roman small letter l with stroke	B3 (2), F9 (13*)	0142	
Ň	Roman capital letter N with acute	D1 (2,13*)	0143	
ň	Roman small letter n with acute	F1 (2,13*)	0144	
Ŏ	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3	
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3	
Ř	Roman capital letter R with caron	D8 (2)	0158	
ř	Roman small letter r with caron	F8 (2)	0159	
Ŝ	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
ŝ	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with acute	AC (2), CA (13*)	0179	
ž	Roman small letter z with acute	BC (2), EA (13*)	017A	
Ẑ	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	

ž Roman small letter z with caron BE (2,4), BC (10), 017E
FE (13*), BA (14*)

Number of characters: 20

Table: sq
Writing System: Albanian Extensions to Basic Roman
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
Ē	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10, 12*,14*)	00CB	
ë	Roman small letter e with diaeresis	EB (1,2,3,4,9,10, 12*,14*)	00EB	

Number of characters: 4

Table: sr_R
Writing System: Cyrillic [Serbian]
Romanization System: UN 1977 (III/11)
Notes: 3a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ć	Roman capital letter C with acute	C6 (2), C3 (13*)	0106	
ć	Roman small letter c with acute	E6 (2), E3 (13*)	0107	
Č	Roman capital letter C with caron	C8 (2,4,10,13*), A1 (14*)	010C	
č	Roman small letter c with caron	E8 (2,4,10,13*), A2 (14*)	010D	
Đ	Roman capital letter D with stroke	D0 (2), A9 (10), A3 (14*)	0110	
đ	Roman small letter d with stroke	F0 (2), B9 (10), A4 (14*)	0111	
LJ	Roman capital letter LJ	-	01C7	
Lj	Roman capital letter L with small letter j	-	01C8	
lj	Roman small letter lj	-	01C9	
NJ	Roman capital letter NJ	-	01CA	
Nj	Roman capital letter N with small letter j	-	01CB	
nj	Roman small letter nj	-	01CC	
Š	Roman capital letter S with caron	A9 (2,4), AA (10), D0 (13*), B2 (14*)	0160	
š	Roman small letter s with caron	B9 (2,4), BA (10), F0 (13*), B3 (14*)	0161	
Ž	Roman capital letter Z with caron	AE (2,4), AC (10), DE (13*), B9 (14*)	017D	
ž	Roman small letter z with caron	BE (2,4), BC (10), FE (13*), BA (14*)	017E	

Number of characters: 16

Table: sv

Writing System: Swedish Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,12*,13*,14*)	00E4	
Å	Roman capital letter A with ring above	C5 (1,4,9,10,12*,13*,14*)	00C5	
å	Roman small letter a with ring above	E5 (1,4,9,10,12*,13*,14*)	00E5	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10,12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10,12*,13*,14*)	00F6	

Number of characters: 8

Table: ta_R

Writing System: Tamil

Romanization System: UN 1972 (II/11), 1977 (III/12)

Notes: 1, 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ḍ	Roman capital letter D with dot below	-	1E0C	
ḍ	Roman small letter d with dot below	-	1E0D	
Ē	Roman capital letter E with breve	-	0114	
ē	Roman small letter e with breve	-	0115	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ḳ	Roman capital letter L with dot below	-	1E36	
ḳ	Roman small letter l with dot below	-	1E37	
Ḷ	(Roman capital letter L with breve below)	-	004C+032E	
ḷ	(Roman small letter l with breve below)	-	006C+032E	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ṅ	Roman capital letter N with dot above	-	1E44	
ṅ	Roman small letter n with dot above	-	1E45	
Ṇ	Roman capital letter N with dot below	-	1E46	
ṇ	Roman small letter n with dot below	-	1E47	
Ṓ	Roman capital letter N with line below	-	1E48	
ṓ	Roman small letter n with line below	-	1E49	
Ō	Roman capital letter O with breve	-	014E	

ö	Roman small letter o with breve	-	014F
Ŕ	Roman capital letter R with line below	-	1E5E
ŕ	Roman small letter r with line below	-	1E5F
Ŗ	Roman capital letter S with dot below	-	1E62
ŗ	Roman small letter s with dot below	-	1E63
Ť	Roman capital letter T with dot below	-	1E6C
ť	Roman small letter t with dot below	-	1E6D
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B
Z	Roman capital letter Z with dot below	-	1E92
z	Roman small letter z with dot below	-	1E93

Number of characters: 32

Table: te_R
Writing System: Telugu
Romanization System: UN 1972 (II/11), 1977 (III/12)
Notes: 1, 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
D	Roman capital letter D with dot below	-	1E0C	
d	Roman small letter d with dot below	-	1E0D	
Ē	Roman capital letter E with breve	-	0114	
ē	Roman small letter e with breve	-	0115	
Ĥ	Roman capital letter H with dot below	-	1E24	
ĥ	Roman small letter h with dot below	-	1E25	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ĺ	Roman capital letter L with dot below	-	1E36	
ĺ	Roman small letter l with dot below	-	1E37	
Ṁ	Roman capital letter M with dot above	B6 (12*)	1E40	
ṁ	Roman small letter m with dot above	B7 (12*)	1E41	
Ṃ	(Roman capital letter M with candrabindu)	-	004D+0310	
ṃ	(Roman small letter m with candrabindu)	-	006D+0310	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ṅ	Roman capital letter N with dot above	-	1E44	
ṅ	Roman small letter n with dot above	-	1E45	
Ṇ	Roman capital letter N with dot below	-	1E46	
ṇ	Roman small letter n with dot below	-	1E47	
Ŏ	Roman capital letter O with breve	-	014E	
ö	Roman small letter o with breve	-	014F	
Ŕ	(Roman capital letter R with dot below and macron)	-	1E5C	
ŗ	(Roman small letter R with dot below and macron)	-	1E5D	
R	Roman capital letter R with dot below	-	1E5A	

ŕ	Roman small letter r with dot below	-	1E5B
Ŗ	Roman capital letter S with dot below	-	1E62
ŝ	Roman small letter s with dot below	-	1E63
Ť	Roman capital letter T with dot below	-	1E6C
ŧ	Roman small letter t with dot below	-	1E6D
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B

Number of characters: 34

Table: tg_R

Writing System: Cyrillic [Tajik]

Romanization System: BGN/PCGN 1994

Notes: 7

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ĕ	Roman capital letter E with dot above	CC (4,10), CB (13*)	0116	
ĕ	Roman small letter e with dot above	EC (4,10), EB (13*)	0117	
Ī	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
ī	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
Ū	Roman capital letter U with breve	DD (3)	016C	
ū	Roman small letter u with breve	FD (3)	016D	

Number of characters: 6

Table: tk

Writing System: Turkmen Extensions to Basic Roman

Notes: 3a, 3b, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10, 12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10, 12*,13*,14*)	00E4	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7	
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7	
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1	
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
\$	Dollar sign	24 (1..14)	0024	
¢	Cent sign	A2 (1,9,13*)	00A2	
Û	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	
ÿ	Roman small letter y with diaeresis	FF (1,9,12*)	00FF	
¥	Yen sign	A5 (1,9)	00A5	Glyph may also have one cross bar
£	Pound sign	A3 (1,3,9,13*)	00A3	

ƚ (Turkmen small letter zh)
Number of characters: 16

Table: tk_R
Writing System: Cyrillic [Turkmen]
Romanization System: BGN/PCGN 1979
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ä	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10, 12*,13*,14*)	00C4	
ä	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10, 12*,13*,14*)	00E4	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	

Number of characters: 6

Table: tl
Writing System: Pilipino (Tagalog) Extensions to Basic Roman
Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	C0 (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
Â	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ĝ	(Roman capital letter G with tilde)	-	0047+0303	
ĝ	(Roman small letter g with tilde)	-	0067+0303	
Ì	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC	
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	

Î	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE
î	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE
Ñ	Roman capital letter N with tilde	D1 (1,3,9,12*,14*)	00D1
ñ	Roman small letter n with tilde	F1 (1,3,9,12*,14*)	00F1
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Ù	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9
ù	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
Û	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*, 14*)	00DB
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*, 14*)	00FB

Number of characters: 32

Table: tn

Writing System: Tswana Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ë	Roman capital letter E with circumflex	D4 (1,2,3,4,9,10, 12*,14*)	00D4	
ë	Roman small letter e with circumflex	F4 (1,2,3,4,9,10, 12*,14*)	00F4	

Number of characters: 4

Table: to

Writing System: Tongan Extensions to Basic Roman

Notes: 3a, 6

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
'	Inverted apostrophe	-	02BB (2018)	

Number of characters: 1

Table: tr

Writing System: Turkish Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Â	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10, 12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10, 12*,14*)	00E2	

		12*,14*)	
Ç	Roman capital letter C with cedilla	C7 (1,2,3,9,12*,14*)	00C7
ç	Roman small letter c with cedilla	E7 (1,2,3,9,12*,14*)	00E7
Ġ	Roman capital letter G with breve	AB (3), D0 (9)	011E
ġ	Roman small letter g with breve	BB (3), F0 (9)	011F
Ī	Roman capital letter I with circumflex	CE (1,2,3,4,9,10, 12*,14*)	00CE
ī	Roman small letter i with circumflex	EE (1,2,3,4,9,10, 12*,14*)	00EE
İ	Roman capital letter I with dot above	A9 (3), DD (9)	0130
ı	Roman small letter i with no dot	B9 (3), FD (9)	0131
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10, 12*,13*,14*)	00D6
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10, 12*,13*,14*)	00F6
Ş	Roman capital letter S with cedilla	AA (2,3), DE (9)	015E
ş	Roman small letter s with cedilla	BA (2,3), FE (9)	015F
Û	Roman capital letter U with circumflex	DB (1,3,4,9,10,12*, 14*)	00DB
û	Roman small letter u with circumflex	FB (1,3,4,9,10,12*, 14*)	00FB
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC

Number of characters: 18

Table: uk_R
Writing System: Cyrillic [Ukrainian]
Romanization System: National 1993
Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ĭ	Roman capital letter I with diaeresis	CF (1,3,9,10,12*, 14*)	00CF	
ï	Roman small letter i with diaeresis	EF (1,3,9,10,12*, 14*)	00EF	

Number of characters: 2

Table: ulit
Writing System: Ulithian Extensions to Basic Roman
Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	(Roman capital letter A with dot above)	-	0041+0307	
á	(Roman small letter a with dot above)	-	0061+0307	
Ĕ	Roman capital letter E with dot above	CC (4,10), CB (13*)	0116	
ĕ	Roman small letter e with dot above	EC (4,10), EB (13*)	0117	
Ȫ	(Roman capital letter O with dot above)	-	004F+0307	
ȫ	(Roman small letter o with dot above)	-	006F+0307	

Number of characters: 6

Table: ur

Writing System: Perso-Arabic [Urdu]

Romanization System: UN 1972 (II/11), 1977 (III/12)

Notes: 1, 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ḑ	Roman capital letter D with dot below	-	1E0C	
ḑ	Roman small letter d with dot below	-	1E0D	
Ḓ	(Roman capital letter G with dot below)	-	0047+0323	
ḓ	(Roman small letter g with dot below)	-	0067+0323	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ḑ	Roman capital letter K with dot below	-	1E32	
ḑ	Roman small letter k with dot below	-	1E33	
Ṛ	Roman capital letter R with dot above	-	1E58	
ṛ	Roman small letter r with dot above	-	1E59	
Ṥ	Roman capital letter T with dot below	-	1E6C	
ṥ	Roman small letter t with dot below	-	1E6D	
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A	
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B	
Ẃ	Roman capital letter W with dot below	-	1E88	
ẃ	Roman small letter w with dot below	-	1E89	
Ỳ	Roman capital letter Y with dot below	-	1EF4	
ỳ	Roman small letter y with dot below	-	1EF5	

Number of characters: 20

Table: uz

Writing System: Uzbek Extensions to Basic Roman

Notes: 3a, 3b, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
G'	(Roman capital letter G with inverted apostrophe)	-	-	
g'	(Roman small letter g with inverted apostrophe)	-	-	
O'	(Roman capital letter O with inverted apostrophe)	-	-	
o'	(Roman small letter o with inverted apostrophe)	-	-	

Number of characters: 4

Table: uz_R
 Writing System: Cyrillic [Uzbek]
 Romanization System: BGN/PCGN 1979
 Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ū	Roman capital letter U with breve	DD (3)	016C	
ū	Roman small letter u with breve	FD (3)	016D	
Number of characters: 2				

Table: vend
 Writing System: Venda Extensions to Basic Roman
 Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ḑ	Roman capital letter D with circumflex below	-	1E12	
ḑ	Roman small letter d with circumflex below	-	1E13	
Ḑ	Roman capital letter L with circumflex below	-	1E3C	
ḑ	Roman small letter l with circumflex below	-	1E3D	
Ḑ	Roman capital letter N with dot above	-	1E44	
ḑ	Roman small letter n with dot above	-	1E45	
Ḑ	Roman capital letter N with circumflex below	-	1E4A	
ḑ	Roman small letter n with circumflex below	-	1E4B	
Ḑ	Roman capital letter T with circumflex below	-	1E70	
ḑ	Roman small letter t with circumflex below	-	1E71	
Number of characters: 10				

Table: vi
 Writing System: Vietnamese Extensions to Basic Roman
 Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	C0 (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10,12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10,12*,14*)	00E1	
Â	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
â	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
À	Roman capital letter A with circumflex and grave	-	1EA6	
à	Roman small letter a with circumflex and grave	-	1EA7	

Á	Roman capital letter A with circumflex and acute	-	1EA4
á	Roman small letter A with circumflex and acute	-	1EA5
Ã	Roman capital letter A with circumflex and tilde	-	1EAA
ã	Roman small letter a with circumflex and tilde	-	1EAB
Ȧ	Roman capital letter A with circumflex and hook above	-	1EA8
ȧ	Roman small letter a with circumflex and hook above	-	1EA9
Ạ	Roman capital letter A with circumflex and dot below	-	1EAC
ạ	Roman small letter a with circumflex and dot below	-	1EAD
Ā	Roman capital letter A with tilde	C3 (1,4,9,10,12*,14*)	00C3
ā	Roman small letter a with tilde	E3 (1,4,9,10,12*,14*)	00E3
Ȧ	Roman capital letter A with breve	C3 (2)	0102
ȧ	Roman small letter a with breve	E3 (2)	0103
À	Roman capital letter A with breve and grave	-	1EB0
à	Roman small letter a with breve and grave	-	1EB1
Ȧ	Roman capital letter A with breve and acute	-	1EAE
ȧ	Roman small letter a with breve and acute	-	1EAF
Ä	Roman capital letter A with breve and tilde	-	1EB4
ä	Roman small letter a with breve and tilde	-	1EB5
Ȧ	Roman capital letter A with breve and hook above	-	1EB2
ȧ	Roman small letter a with breve and hook above	-	1EB3
Ạ	Roman capital letter A with breve and dot below	-	1EB6
ạ	Roman small letter a with breve and dot below	-	1EB7
Ȧ	Roman capital letter A with hook above	-	1EA2
ȧ	Roman small letter a with hook above	-	1EA3
Ạ	Roman capital letter A with dot below	-	1EA0
ạ	Roman small letter a with dot below	-	1EA1
Đ	Roman capital letter D with stroke	D0 (2), A9 (10), A3 (14*)	0110
đ	Roman small letter d with stroke	F0 (2), B9 (10), A4 (14*)	0111
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8

è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA
Ë	Roman capital letter E with circumflex and grave	-	1EC0
ë	Roman small letter e with circumflex and grave	-	1EC1
Ê	Roman capital letter E with circumflex and acute	-	1EBE
é	Roman small letter e with circumflex and acute	-	1EBF
Ë	Roman capital letter E with circumflex and tilde	-	1EC4
ë	Roman small letter e with circumflex and tilde	-	1EC5
Ê	Roman capital letter E with circumflex and hook above	-	1EC2
ê	Roman small letter e with circumflex and hook above	-	1EC3
Ë	Roman capital letter E with circumflex and dot below	-	1EC6
ë	Roman small letter e with circumflex and dot below	-	1EC7
Ë	Roman capital letter E with tilde	-	1EBC
ë	Roman small letter e with tilde	-	1EBD
Ê	Roman capital letter E with hook above	-	1EBA
ê	Roman small letter e with hook above	-	1EBB
E	Roman capital letter E with dot below	-	1EB8
e	Roman small letter e with dot below	-	1EB9
Ì	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED
İ	Roman capital letter I with tilde	A5 (4,10)	0128
ı	Roman small letter i with tilde	B5 (4,10)	0129
İ	Roman capital letter I with hook above	-	1EC8
ı	Roman small letter i with hook above	-	1EC9
İ	Roman capital letter I with dot below	-	1ECA
ı	Roman small letter i with dot below	-	1ECB
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*, 13*,14*)	00D3
ó	Roman small letter o with acute	F3 (1,2,3,9,10,12*, 13*,14*)	00F3

Ô	Roman capital letter O with circumflex	D4 (1,2,3,4,9,10,12*,14*)	00D4
ó	Roman small letter o with circumflex	F4 (1,2,3,4,9,10,12*,14*)	00F4
Û	Roman capital letter O with circumflex and grave	-	1ED2
ô	Roman small letter o with circumflex and grave	-	1ED3
Ó	Roman capital letter O with circumflex and acute	-	1ED0
ó	Roman small letter o with circumflex and acute	-	1ED1
Õ	Roman capital letter O with circumflex and tilde	-	1ED6
õ	Roman small letter o with circumflex and tilde	-	1ED7
Ö	Roman capital letter O with circumflex and hook above	-	1ED4
ö	Roman small letter o with circumflex and hook above	-	1ED5
Ï	Roman capital letter O with circumflex and dot below	-	1ED8
ï	Roman small letter o with circumflex and dot below	-	1ED9
Ö	Roman capital letter O with tilde	D5 (1,4,9,10,12*,13*,14*)	00D5
ö	Roman small letter o with tilde	F5 (1,4,9,10,12*,13*,14*)	00F5
Ï	Roman capital letter O with hook above	-	1ECE
ï	Roman small letter o with hook above	-	1ECF
Ŏ	Roman capital letter O with horn	-	01A0
ȯ	Roman small letter o with horn	-	01A1
Ï	Roman capital letter O with horn and grave	-	1EDC
ï	Roman small letter o with horn and grave	-	1EDD
Ŏ	Roman capital letter O with horn and acute	-	1EDA
ȯ	Roman small letter o with horn and acute	-	1EDB
Ï	Roman capital letter O with horn and tilde	-	1EE0
ï	Roman small letter o with horn and tilde	-	1EE1
Ï	Roman capital letter O with horn and hook above	-	1EDE
ï	Roman small letter o with horn and hook above	-	1EDF
Ŏ	Roman capital letter O with horn and dot below	-	1EE2
ȯ	Roman small letter o with horn and dot below	-	1EE3

	dot below		
Ō	Roman capital letter O with dot below	-	1ECC
o	Roman small letter o with dot below	-	1ECD
Ū	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9
ù	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
u	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
Ů	Roman capital letter U with tilde	DD (4), D7 (10)	0168
ů	Roman small letter u with tilde	FD (4), F7 (10)	0169
Ů̆	Roman capital letter U with hook above	-	1EE6
ů̆	Roman small letter u with hook above	-	1EE7
Ů̇	Roman capital letter U with horn	-	01AF
ů̇	Roman small letter u with horn	-	01B0
Ů̈	Roman capital letter U with horn and grave	-	1EEA
ů̈	Roman small letter u with horn and grave	-	1EEB
Ů̉	Roman capital letter U with horn and acute	-	1EE8
ů̉	Roman small letter u with horn and acute	-	1EE9
Ů̊	Roman capital letter U with horn and tilde	-	1EEE
ů̊	Roman small letter u with horn and tilde	-	1EEF
Ů̋	Roman capital letter U with horn and hook above	-	1EEC
ů̋	Roman small letter u with horn and hook above	-	1EED
Ů̌	Roman capital letter U with horn and dot below	-	1EF0
ů̌	Roman small letter U with horn and dot below	-	1EF1
Ů̍	Roman capital letter U with dot below	-	1EE4
ů̍	Roman small letter u with dot below	-	1EE5
Ÿ	Roman capital letter Y with grave	AC (12*)	1EF2
ý	Roman small letter y with grave	BC (12*)	1EF3
Ÿ̇	Roman capital letter Y with acute	DD (1,2,10,12*,14*)	00DD
ý̇	Roman small letter y with acute	FD (1,2,10,12*,14*)	00FD
Ÿ̈	Roman capital letter Y with tilde	-	1EF8
ý̈	Roman small letter y with tilde	-	1EF9
Ÿ̉	Roman capital letter Y with hook above	-	1EF6
ý̉	Roman small letter y with hook above	-	1EF7
Ÿ̊	Roman capital letter Y with dot below	-	1EF4
ý̊	Roman small letter y with dot below	-	1EF5

Number of characters: 134

Table: wo

Writing System: Wolof Extensions to Basic Roman

Notes: 3a, 4, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ă	Roman capital letter A with circumflex	C2 (1,2,3,4,9,10,12*,14*)	00C2	
ă	Roman small letter a with circumflex	E2 (1,2,3,4,9,10,12*,14*)	00E2	
Ȧ	Roman capital letter A with breve	C3 (2)	0102	
ȧ	Roman small letter a with breve	E3 (2)	0103	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10,12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10,12*,13*,14*)	00E9	
Ĥ	(Roman capital letter H with diaeresis below)	-	0048+0324	
ĥ	(Roman small letter h with diaeresis below)	-	0068+0324	
Ț	(Roman capital letter T with diaeresis below)	-	0054+0324	
ț	(Roman small letter t with diaeresis below)	-	0074+0324	

Number of characters: 12

Table: yape

Writing System: Yapese Extensions to Basic Roman

Notes: n/a

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ȧ	Roman capital letter A with diaeresis	C4 (1,2,3,4,9,10,12*,13*,14*)	00C4	
ȧ	Roman small letter a with diaeresis	E4 (1,2,3,4,9,10,12*,13*,14*)	00E4	
Ë	Roman capital letter E with diaeresis	CB (1,2,3,4,9,10,12*,14*)	00CB	
ë	Roman small letter e with diaeresis	EB (1,2,3,4,9,10,12*,14*)	00EB	
Ö	Roman capital letter O with diaeresis	D6 (1,2,3,4,9,10,12*,13*,14*)	00D6	
ö	Roman small letter o with diaeresis	F6 (1,2,3,4,9,10,12*,13*,14*)	00F6	

Number of characters: 6

Table: yo

Writing System: Yoruba Extensions to Basic Roman

Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
Ɛ	Roman capital letter E with dot below	-	1EB8	
ɛ	Roman small letter e with dot below	-	1EB9	
Ɔ	Roman capital letter O with dot below	-	1ECC	

o	Roman small letter o with dot below	-	1ECD
S	Roman capital letter S with dot below	-	1E62
s	Roman small letter s with dot below	-	1E63
Number of characters: 6			

Table: zh_R1
 Writing System: Sino-Japanese [Chinese]
 Romanization System: UN 1977 (III/8)³
 Notes: 3a, 5

<u>Char.</u>	<u>Character Name</u>	<u>ISO 8859</u>	<u>ISO/Unic.</u>	<u>Notes/Exceptions</u>
À	Roman capital letter A with grave	C0 (1,3,9,12*,14*)	00C0	
à	Roman small letter a with grave	E0 (1,3,9,12*,14*)	00E0	
Á	Roman capital letter A with acute	C1 (1,2,3,4,9,10, 12*,14*)	00C1	
á	Roman small letter a with acute	E1 (1,2,3,4,9,10, 12*,14*)	00E1	
Ā	Roman capital letter A with macron	C0 (4,10), C2 (13*)	0100	
ā	Roman small letter a with macron	E0 (4,10), E2 (13*)	0101	
Ă	Roman capital letter A with caron	-	01CD	
ă	Roman small letter a with caron	-	01CE	
È	Roman capital letter E with grave	C8 (1,3,9,12*,14*)	00C8	
è	Roman small letter e with grave	E8 (1,3,9,12*,14*)	00E8	
É	Roman capital letter E with acute	C9 (1,2,3,4,9,10, 12*,13*,14*)	00C9	
é	Roman small letter e with acute	E9 (1,2,3,4,9,10, 12*,13*,14*)	00E9	
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Ē	Roman capital letter E with macron	AA (4), A2 (10), C7 (13*)	0112	
ē	Roman small letter e with macron	BA (4), B2 (10), E7 (13*)	0113	
Ĕ	Roman capital letter E with caron	CC (2)	011A	
ĕ	Roman small letter e with caron	EC (2)	011B	
Ĭ	Roman capital letter I with grave	CC (1,3,9,12*,14*)	00CC	
ì	Roman small letter i with grave	EC (1,3,9,12*,14*)	00EC	
Í	Roman capital letter I with acute	CD (1,2,3,4,9,10, 12*,14*)	00CD	
í	Roman small letter i with acute	ED (1,2,3,4,9,10, 12*,14*)	00ED	
Ī	Roman capital letter I with macron	CF (4), A4 (10), CE (13*)	012A	
ī	Roman small letter i with macron	EF (4), B4 (10), EE (13*)	012B	
Ĳ	Roman capital letter I with caron	-	01CF	
ı	Roman small letter i with caron	-	01D0	
Ò	Roman capital letter O with grave	D2 (1,3,9,12*,14*)	00D2	
ò	Roman small letter o with grave	F2 (1,3,9,12*,14*)	00F2	
Ó	Roman capital letter O with acute	D3 (1,2,3,9,10,12*,	00D3	

³ - for completeness, Chinese tonal marks have been included in the above table, although these would not normally be shown on geographical names

ó	Roman small letter o with acute	13*,14*) F3 (1,2,3,9,10,12*, 13*,14*)	00F3
Ō	Roman capital letter O with macron	D2 (4,10), D4 (13*)	014C
ō	Roman small letter o with macron	F2 (4,10), F4 (13*)	014D
Ȫ	Roman capital letter O with caron	-	01D1
ȯ	Roman small letter o with caron	-	01D2
Ù	Roman capital letter U with grave	D9 (1,3,9,12*,14*)	00D9
ù	Roman small letter u with grave	F9 (1,3,9,12*,14*)	00F9
Ú	Roman capital letter U with acute	DA (1,2,3,4,9,10, 12*,14*)	00DA
ú	Roman small letter u with acute	FA (1,2,3,4,9,10, 12*,14*)	00FA
Ū	Roman capital letter U with macron	DE (4), AE (10), DB (13*)	016A
ū	Roman small letter u with macron	FE (4), BE (10), FB (13*)	016B
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC
Û	Roman capital letter U with diaeresis and grave	-	01DB
ü	Roman small letter u with diaeresis and grave	-	01DC
Ů	Roman capital letter U with diaeresis and acute	-	01D7
ů	Roman small letter u with diaeresis and acute	-	01D8
Ů	Roman capital letter U with diaeresis and macron	-	01D5
ū	Roman small letter u with diaeresis and macron	-	01D6
Û	Roman capital letter U with diaeresis and caron	-	01D3
ů	Roman small letter u with diaeresis and caron	-	01D4
Ů	Roman capital letter U with caron	-	01D3
ů	Roman small letter u with caron	-	01D4

Number of characters: 52

Table: zh_R2

Writing System: Sino-Japanese [Chinese]

Romanization System: Modified Wade-Giles 1912

Notes: n/a

Char.	Character Name	ISO 8859	ISO/Unic.	Notes/Exceptions
Ê	Roman capital letter E with circumflex	CA (1,3,9,12*,14*)	00CA	
ê	Roman small letter e with circumflex	EA (1,3,9,12*,14*)	00EA	
Û	Roman capital letter U with breve	DD (3)	016C	
ü	Roman small letter u with breve	FD (3)	016D	
Ü	Roman capital letter U with diaeresis	DC (1,2,3,4,9,10, 12*,13*,14*)	00DC	
ü	Roman small letter u with diaeresis	FC (1,2,3,4,9,10, 12*,13*,14*)	00FC	

Number of characters: 6

Draft Toponymic Data Exchange Standard

Part 1 Data Elements

The following data elements, recommended in Resolution 4 of the first United Nations Conference on the Standardization of Geographical Names, comprise the minimum set of critical geographic names information for digital exchange.

1.0 GEOGRAPHIC NAME -- the spelling of a standardized name referring to a feature. Spellings of variant names, if any, follow the standardized name with suitable delimiters. An explanatory note accompanying the data set must define all conventions used in the presentation of GEOGRAPHIC NAME; e.g., sorting order, reversal of generic terms, meaning of special flags such as asterisks, etc.

Data Type: text. An explanatory note accompanying the data set must define the text format and encoding; e.g.,

char, ASCII
char, ISO 8859 (2)
char, KOI 8
wchar_t, ISO 10646

2.0 FEATURE TYPE -- a characterization of the kind of feature represented by GEOGRAPHIC NAME. An explanatory note accompanying the data set must define the characterization scheme employed.

Data Type: text

3.0 LOCATION -- the position associated with GEOGRAPHIC NAME or FEATURE TYPE expressed in latitude and longitude (based on the Prime Meridian, Greenwich). For most toponymic data sets, the preferred format would be positions specified by degrees and minutes, rounded down to the nearest integer minute. This format allows future update to include seconds or decimal minutes. An explanatory note accompanying the data set must identify alternate formats if applicable; e.g., decimal degrees, integer degrees and minutes, integer degrees and decimal minutes, etc.

3.1 LATITUDE -- the latitude value of LOCATION.

Data Type: integer
Domain: 90 00S <= LATITUDE <= 90 00N

3.2 LONGITUDE -- the longitude value of LOCATION.

Data Type: integer
Domain: 180 00E <= LONGITUDE <= 180 00W

4.0 ADMINISTRATIVE AREA -- an identification of the country and administrative division in which LOCATION falls.

Data Type: text
Domain: (the domain is the content of ISO/DIS 3166, Parts 1 and 2)

5.0 MAP SHEET REFERENCE -- an identification of the sheet (within a standard national map series) containing LOCATION. An explanatory note accompanying the data set must provide citation for the referenced map series.

Data Type: text

Part 2 Metadata

* The information presented here is summarized from *Content Standard for Digital Geospatial Metadata*, Federal Geographic Data Committee (USA), April, 1997.

The notes to Part 1 of this Annex refer to additional information required by the user of a data set to interpret competently the data contained therein. This additional information is generally referred to as *metadata*.

Metadata serves to answer four questions:

What sets of data exist for a geographic location?

Does a set of data meet a specific need?

How is a set of data acquired?

What information is required to process and use a set of data?

Metadata characterizing geographic names data sets would typically comprise the following subsets:

Identification information – basic information about the data, including publisher.

Text encoding standard - identification of the national, international, or proprietary standard used to represent the text digitally in the data set.

Data quality information – a general assessment of the quality (accuracy, currency) of the data set.

Spatial data organization (if applicable) – the mechanism (text, point, vector, raster) used to represent spatial data in the data set.

Spatial reference information (if applicable) – the description of the reference frame for coordinates in the data set.

Entity and attribute information – the description of the content of a data set including entity types, their attributes and domains (Part 1 of this Annex qualifies as entity and attribute metadata).

Software information - identification of the software configuration used to create the data set, including identification of operating system and application.

Distribution information – information about the distribution of and options for obtaining the data set.

Part 15 of ISO 15046 is a draft standard for metadata pertaining to geographic information. It is based largely on the Metadata Standard developed by the United States Federal Geographic Data Committee (FGDC). While it is beyond the scope of the present Working Group to define rigorously a metadata standard specific to toponymic data, it is appropriate and necessary to inform the Conference of the importance of describing the content of geographic names data sets in sufficient detail to ensure their utility in a program of exchange of digital information.

Part 3 Illustration

The following listing is an abbreviated example of a print-out of a digital toponymic data set adhering to the guidelines recommended by the Working Group. Two files are represented; the first is the toponymic data, the second is a text file (conventionally known as a “readme” file) containing the metadata necessary for a recipient to determine the data’s content and utility.

Data.txt

Aabla	PPL	5933N	2532E	EE	37	NO35-01		
Aadami	PPL	5815N	2652E	EE	78	NO35-05		
Aadma	PPL	58475N	2239E	EE	39	NO34-06		
Aadma Asundus > Aadma	PPL	5847N	2239E	EE	39	NO34-06		
Aakaru	PPL	5815N	2639E	EE	78	NO35-05		
Aakla > Haakla	PPL	5849N	2454E	EE	70	NO35-04		
Aakre	PPL	5806N	2611E	EE	82	NO35-05		
Aakre Asundus > Aakre	PPL	580605	261140	EE	82	NO35-05		
Aakre Peatus	RSTP	5804N	2616E	EE	82	NO35-05		
Aaksi Saar > Aksii	ISL	5935N	2505E	EE	37	NO35-01		
Aamse	PPLX	5855N	2342E	EE	57	NO34-06		
Aamsi > Aamse	PPLX	5855N	2342E	EE	57	NO34-06		
Aandu	PPL	5909N	2443E	EE	70	NO35-01		

Aardla	PPL	5818N	2646E	EE	78	NO35-05	
Aardla Järv	LK	5818N	2645E	EE	78	NO35-05	
Aardla Peatus	RSTP	5821N	2642E	EE	78	NO35-05	
Äärepere	PPLX	5833N	2338E	EE	57	NO34-06	
Aarna	PPL	5803N	2658E	EE	65	NO35-05	
Aasa	PPL	5841N	2430E	EE	67	NO35-04	
Ääsmäe	PPL	5914N	2430E	EE	37	NO35-01	
Ääsmäe Raba	BOG	5914N	2433E	EE	37	NO35-01	
Aaspere	PPL	5925N	2607E	EE	59	NO35-02	
Aaspere Asundus > Aaspere	PPL	5925N	2607E	EE	59	NO35-02	
Aasu	PPL	5924N	2619E	EE	59	NO35-02	
Aasukalda	PPL	5926N	2643E	EE	59	NO35-02	
Aasumetsa	PPL	5930N	2550E	EE	59	NO35-01	
Aasuvälja	PPL	5853N	2525E	EE	51	NO35-04	
Aasuvälja	PPL	5919N	2635E	EE	59	NO35-02	
Aavena Säär	SPIT	5830N	2338E	EE	67	NO34-06	
Aavere	PPL	5923N	2516E	EE	37	NO35-01	
Aavere	PPL	5904N	2603E	EE	59	NO35-02	
Aaviku	PPL	5929N	2617E	EE	59	NO35-02	
Aaviku	PPL	5920N	2454E	EE	37	NO35-01	
Aavre > Aavere	PPL	5923N	2516E	EE	37	NO35-01	
Abaja	PPL	5855N	2557E	EE	51	NO35-04	
Abbia > Abja-Paluoja	PPL	5807N	2520E	EE	84	NO35-04	
Abbul > Abula	PPL	5825N	2208E	EE	74	NO34-06	
Abia > Abja-Paluoja	PPL	5807N	2520E	EE	84	NO35-04	
Abissaare	PPL	5805N	2643E	EE	65	NO35-05	
Abja > Abjaku	PPL	5806N	2521E	EE	84	NO35-04	
Abja Asundus > Abjaku	PPL	5806N	2521E	EE	84	NO35-04	
Abjaku	PPL	5806N	2521E	EE	84	NO35-04	
Abja-Paluoja	PPL	5807N	2520E	EE	84	NO35-04	
Abro > Abruca	ISL	5808N	2230E	EE	74	NO34-06	
Abruk, Ostrov > Abruca	ISL	5808N	2230E	EE	74	NO34-06	
Abruca	ISL	5808N	2230E	EE	74	NO34-06	
Abruca	PPL	5809N	2230E	EE	74	NO34-06	
Abruca, Ostrov > Abruca	ISL	5808N	2230E	EE	74	NO34-06	
Abruca Island > Abruca	ISL	5808N	2230E	EE	74	NO34-06	
Abruca Saar > Abruca	ISL	580856	223059	EE	74	NO34-06	
Abula	PPL	5825N	2208E	EE	74	NO34-06	
Aburi	PPL	5910N	2616E	EE	59	NO35-02	
Ab'ya > Abja-Paluoja	PPL	5807N	2520E	EE	84	NO35-04	
Ab'ya-Paluoja > Abja-Paluoja	PPL	5807N	2520E	EE	84	NO35-04	
Açupe > Atse Jõgi	STM	5753N	2529E	EE	00	NO35-07	
Acup'ite > Atse Jõgi	STM	5753N	2529E	EE	00	NO35-07	
Adaka	PPL	5929N	2608E	EE	59	NO35-02	
Ädara > Ädara Jõgi	STM	5914N	2640E	EE	59	NO35-02	
Ädara Jõgi	STM	5914N	2640E	EE	59	NO35-02	
Adavere	PPL	5842N	2553E	EE	49	NO35-04	
Adavere Asundus > Adavere	PPL	5842N	25535E	EE	49	NO35-04	
Adila	PPL	5906N	2438E	EE	70	NO35-01	
Adila-Krimmi	PPLX	5907N	2436E	EE	70	NO35-01	
Adila-Tagaküla	PPLX	5909N	2436E	EE	70	NO35-01	
Adila-Tagakyula > Adila-Tagaküla	PPLX	5909N	2436E	EE	70	NO35-01	

Adiste	PPL	5806N	2707E	EE	67	NO35-05		
Adra	PPL	5922N	2422E	EE	37	NO35-01		
Adraku	PPL	5855N	2649E	EE	44	NO35-05		
Adraküla > Adra	PPL	5922N	2422E	EE	37	NO35-01		
Adrakyulya > Adra	PPL	5922N	2422E	EE	37	NO35-01		
Adrika	PPLX	5929N	2643E	EE	59	NO35-02		
Adriku > Adrika	PPLX	5929N	2643E	EE	59	NO35-02		
Ädu	PPL	5800N	2614E	EE	82	NO35-05		
Aedlinn Peatus > Klooga Aedlinna Peatus	RSTP	5919N	2412E	EE	37	NO35-01		
Aegna > Aegna Tulepaak	BCN	5936N	2444E	EE	37	NO35-01		
Aegna ISL	5934N	2444E	EE	37	NO35-01			
Aegna, Ostrov > Aegna ISL	5934N	2444E	EE	37	NO35-01			
Aegna Island > Aegna ISL	5934N	2444E	EE	37	NO35-01			
Aegna Saar > Aegna ISL	5934N	2444E	EE	37	NO35-01			
Aegna-Sar > Aegna ISL	5934N	2444E	EE	37	NO35-01			

Readme.txt

This file describes the content, lineage, and utility of the toponymic data contained in the file data.txt.

Data Identification: Place name spellings for features in Estonia.

Text encoding standard: Eight-bit ASCII (ISO 8859 Latin 1). Two characters (Roman small letter c with caron; Roman small letter i with macron) are encoded in a proprietary scheme, and may be viewed using the Monotype Special G1 family of fonts.

Data quality: Name spellings are current for most features as of 1993, and have been taken from the following authoritative Estonian sources:

Eesti Topograafiline Kaart, 1:200,000, 1993.

Eesti Ülevaatekaart, 1:400,000, 1993.

Eesti Sood, 1:400,000, 1993.

Eesti Veed, 1:400,000, 1991.

Settlement names were taken from a list provided by Eesti Keele Instituut, Tallinn, in 1994.

Data organization: Data is provided alphabetically by name spelling. Data elements within a record are tab delimited (ASCII character 09). Records are delimited by carriage return (ASCII character 0D) Single geographic coordinates are provided for each record.

Data reference: Coordinates are referenced to the World Geodetic System 1984.

Entity and attribute information: Each record contains the following data elements and associated attributes:

Geographic name. Name spellings conform to standard Estonian orthography with the exception that initial letters of generic terms are capitalized. Variant names are cross referenced to standard names by use of the greater than sign (ASCII character 3E).

Feature type. A five character designation denoting the type of feature to which the name refers. The designation scheme is that employed by the U.S. Board on Geographic Names in its data base of foreign geographic names. This data element is left justified, and space filled (ASCII character 20) in cases where the designation is fewer than five characters.

Latitude. The geographic latitude of the named feature, taken at the center of spot and areal features, and at one end of linear features. The data element is five characters in length;

the first two represent integer degrees, the second two integer minutes (rounded down), and the fifth the compass point N.

Longitude. The geographic longitude of the named feature, taken at the center of spot and areal features, and at one end of linear features. The data element is five characters in length; the first two represent integer degrees, the second two integer minutes (rounded down), and the fifth the compass point E.

Country code. The country code digraph taken from ISO 3166 "Codes for the representation of names of countries and their subdivisions."

Administrative division code. The two character administrative division code taken from ISO 3166 "Codes for the representation of names of countries and their subdivisions." The administrative division identified for each name is the division in which the geographic location for the record falls. Features that cross administrative or international boundaries are given the general code 00.

Map sheet reference. The map sheet number of the standard 1:250,000 map series on which the geographic coordinate of the record falls. The feature and its name may not necessarily actually appear on the referenced sheet.

Software requirements. This data set was compiled using Microsoft Word 7.0 for Windows 95. The only known special requirement is the use of the Monotype Special G1 font for display of certain characters, i.e., lower case Roman c with caron above, lower case Roman i with macron above.

Distribution information. There are no restrictions on the distribution and use of the information contained in these data sets. Additional copies of this data set may be obtained from the producer.

Producer point of contact. This data set was produced by the staff of the Foreign Names Committee of the U.S. Board on Geographic Names. Questions, corrections, and requests for additional information may be referred to:

Executive Secretary for Foreign Names
U.S. Board on Geographic Names
National Imagery and Mapping Agency D-56
4600 Sangamore Road
Bethesda, Maryland 20816-5003
USA

Summary
ISO/IEC 10646 and the Unicode Standard, Version 2.0

Introduction

Geographic names information, as understood in most United Nations contexts, is textual in form. This fact simplifies and, at the same time, complicates the transmission and exchange of geographic names in digital form. The process is simplified because one does not face immediately the task of defining standards for complex data structures and objects. The difficulty lay in the absence of widely implemented standards for exchange of text in writing systems beyond the basic fund of the twenty six alphabetic characters of conventional English orthography.

The International Organization for Standardization (ISO), in parallel (now in conjunction) with the Unicode Consortium, has developed International Standard ISO/IEC 10646-1:1993 as a solution for the problem of encoding text in a multilingual information processing environment. This annex provides a summary of text encoding, a brief description of the design elements behind Unicode, and sources of additional information.

Text Encoding

Text encoding in a data processing environment may be defined as the schema employed to represent a text character in binary form. A variety of standard schemes have evolved since the introduction of commercially viable computing apparatus. The basis for the majority of these is the ordered sequence of eight bits known as an octet (or byte), allowing 256 encoded characters per schema. At present, the most advanced international standard (as distinct from national or proprietary) for alphabetic text is the family of schemes under the umbrella of ISO/IEC 8859, covering many roman-alphabet orthographies as well as Arabic, Cyrillic, Greek, and Hebrew. National encoding standards, developed by national standardization authorities, have also arisen with varying degrees of international acceptance. Examples include the Russian Kod obmena informatsii - 8 (KOI-8), and the Greek national standard ELOT 928. Proprietary industry standards add to the mix, and include code pages developed and implemented by IBM, DEC, and Microsoft.

The 256-character upper bound on each set limits the utility of eight-bit encoding schemes. Applications that invoke a variety of writing systems have of necessity been designed to deal with a number of separate regional character sets, with built in or user-invoked switching mechanisms to navigate from set to set.

The limited definition of text encoding when discussing these standards is important. Encoding is the minimum set of information required to represent a text string in binary form. The graphic appearance of the resulting text string when displayed on a monitor or the printed page is not the concern of the developer of an encoding standard. The visual characteristics of the displayed text are dependent on the application that renders the code in viewable form.

Expanding the Repertoire

A more elegant solution to text encoding that would embrace a variety of writing systems began to emerge with the expansion of the Internet into common usage as a means of international communication. By the early 1990's both ISO and the newly formed commercially based Unicode Consortium had embarked on the definition of an encoding standard based on two octets per character, sometimes called *wide ASCII*. The sixteen-bit approach theoretically supports a repertoire of 2^{16} character values, or over 65,000 codes. The most current version of the standard (referred to here as Unicode 2.0) implements almost 39,000 of these.

Design Principles¹

Ten overarching principles govern the development of the Unicode standard:

Sixteen-bit characters

Unicode codes represent characters using a uniform width of 16 bits.

Full encoding

The full 16-bit code space is available to encode characters.

Characters, not glyphs

The standard encodes characters, not glyphs (i.e., the shapes of a characters when displayed).

Semantics

Characters have well-defined semantics (i.e., a definition of properties that support rendering, sorting, and other algorithms).

Plain text

The standard encodes plain text (not, e.g., rich text).

Logical order

The default for memory representation is logical order (i.e., the order in which text is typed at the keyboard).

Unification

The standard unifies duplicate characters within scripts across languages.

Dynamic composition

The standard allows for the dynamic composition of accented characters.

Equivalent sequence

Common pre-composed character-accent forms are accounted for in the standard, and can be mapped to equivalent dynamically composed code sequences.

Convertibility

Accurate convertibility (at the character level) is guaranteed between the standard and other widely accepted international and industry standards.

Sixteen-Bit Data in a Lesser-Bit Environment

Exchange of digital information requires compatibility among the sender, receiver, and transmitter. Today the limiting factor is imposed on the transmission media that tend to rely on the most elementary standards to ensure successful data exchange. Thus, Internet mail agents adhering to the Simple Mail Transfer Protocol (SMTP) are likely to transmit information in octets with seven significant bits. The exchange of 16-bit Unicode information in this environment is not trivial, requiring transformations, without loss of data, between the two environments.

Two such transformations have been developed for seven- and eight-bit environments. They are commonly referred to as Universal Character Set Transformation Formats or UTF. UTF-8 has been added to ISO/IEC 10646 through the ISO amendment process.

¹ Taken from *The Unicode Standard, Version 2.0*, The Unicode Consortium, Addison-Wesley Developers Press, Reading Massachusetts, July 1996.

Advantages of the Unicode Standard in Processing Geographic Names

The Unicode standard provides a comprehensive self-consistent framework for encoding text in a multilingual (and multiscriptual) environment. In the exchange of geographic names information there is now a practical basis for the exchange of digital toponymic information among users requiring not only Roman-script spellings (native or romanized) but the official non-Roman form as well, a common text encoding scheme being used by sender, communicator, and receiver. The following specific advantages obtain in applying Unicode to the exchange of geographic names data.

Comprehensiveness. The review conducted by this working group and documented in Appendix B has shown that the Unicode standard accounts for the majority of the writing systems (at least in their romanized form) that might be required in the global exchange of digital toponymic information.

Cost. The application of Unicode to geographic names data does not require to outlay of significant funds to develop specialized software. Users of personal computers will attain Unicode compatibility through the routine process of moving to regularly issued upgrades of commercial operating systems and applications.

Near-Term Achievability. Unicode based applications have already emerged and are commercially available. Conversion of existing systems and data bases to the Unicode standard can begin immediately.

Synchronization with Commercial Standards. The Unicode standard is driven by the market requirements of major commercial software developers. The standard will thus become the basis for industry-wide text encoding, especially in international applications.

Implementations of Unicode

As the Internet was the impetus behind the development of 16-bit text encoding, it is also where current practical applications of Unicode are most easily found. The most recent versions of the popular Microsoft Internet Explorer® and Netscape Communicator® provide comprehensive Unicode-based functions to create, host and display fully multilingual hypertext mark-up language documents.

Unicode compatible fonts are becoming available but are yet not as numerous as those compliant with the ISO 8859 family of encoding standards. Lucida Sans Unicode, shipped with Microsoft products, and the Bitstream Cyberbit font are two examples that are easily obtainable. Users should be aware however that these fonts do not currently support the entire spectrum of the Unicode standard.

A listing of commercially available software products supporting the Unicode standard may be found at www.unicode.org.