
**Informatika in dokumentacija - Transliteracija arabskih znakov v latinične znake -
3. del: Perzijski jezik - Transliteracija**

Information and documentation -- Transliteration of Arabic characters into Latin
characters - Part 3: Persian language — Transliteration

iTeh STANDARD PREVIEW

Information et documentation -- Translittération des caractères arabes en caractères
latins - Partie 3: Persan — Translittération

[oSIST ISO/DIS 233-3:2021](https://standards.iteh.ai/catalog/standards/sist/233-3-2021/233-3-2021)

Ta slovenski standard je istoveten z: ISO/DIS 233-3

<https://standards.iteh.ai/catalog/standards/sist/233-3-2021/233-3-2021>

ICS:

01.140.10 Pisanje in prečkovanje Writing and transliteration

oSIST ISO/DIS 233-3:2021

en

iTeh STANDARD PREVIEW **(standards.iteh.ai)**

[oSIST ISO/DIS 233-3:2021](https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021)

<https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021>

DRAFT INTERNATIONAL STANDARD

ISO/DIS 233-3

ISO/TC 46

Secretariat: AFNOR

Voting begins on:
2021-09-14Voting terminates on:
2021-12-07

Information and documentation — Transliteration of Arabic characters into Latin characters —

Part 3: Persian language — Transliteration

*Information et documentation — Translittération des caractères arabes en caractères latins —
Partie 3: Persan — Translittération*

ICS: 01.140.10

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[oSIST ISO/DIS 233-3:2021](https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021)

<https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021>

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

This document is circulated as received from the committee secretariat.



Reference number
ISO/DIS 233-3:2021(E)

© ISO 2021

iTeh STANDARD PREVIEW (standards.iteh.ai)

[oSIST ISO/DIS 233-3:2021](https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021)

<https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021>



COPYRIGHT PROTECTED DOCUMENT

© ISO 2021

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

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

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Transliteration tables	2
4.1 Table 1 — Consonants.....	2
4.2 Vowels.....	4
4.2.1 Table 2a — Vowels – for fully reversible system.....	4
4.2.2 Table 2b — Vowels – modified transliteration options.....	4
4.3 Arabic elements in the Persian language.....	5
4.3.1 Table 3 — Conventional signs.....	5
4.4 Table 4 — Different forms of <i>hamzeh</i>	6
4.5 Table 5 — Punctuation marks.....	7
4.6 Table 6 — Persian numerals.....	7
Annex A (normative) Different positional forms of characters	8
Annex B (normative) General principles	9
Bibliography	12

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[oSIST ISO/DIS 233-3:2021](https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021)

<https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021>

ISO/DIS 233-3:2021(E)

Foreword

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

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

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

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

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

This document was prepared by Technical Committee ISO/TC 46 *Information and documentation*.

This second edition cancels and replaces the first edition (ISO 233-3:1999), which has been technically revised.

The main changes compared to the previous edition are as follows:

- Added missing diacritical signs to the tables and corrected some errors elsewhere in the text. Added distinction in transliteration between ١ and ١̇. Changed transliteration of خ (table 1, row 9) and ض (table 1, row 18), from ḳ to x and ẓ to ẓ respectively. Changed transliteration of tanvin (table 3, row 2) from " to ã.
- Incorporated options for 3 “levels” of transliteration: strict, i.e. fully reversible (with and without vowels); and a modified but not fully reversible system, which, for example, distinguishes in transliteration when the characters ٣ and ٣ function as long vowels or as consonants, and, in the case of ٣, other functions of the character.
- Added hexadecimal character codes (ISO/IEC 10646 or Unicode) to all tables containing Persian characters and transliterations and therefore omitted [Annexes B](#) and D, Annex C thus becoming [Annex B](#).
- Added Terms and definitions section.

A list of all parts in the ISO 233 series can be found on the ISO website.

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

Introduction

This part of ISO 233 is one of a series of International Standards, dealing with the conversion of systems of writing. The aim of this part of ISO 233 and others in the series is to provide a means for international communication of written messages in a form which permits the automatic transmission and reconstitution of these, by men or machines. The system of conversion, in this case, must be univocal and entirely reversible to allow for retransliteration.

This means that consideration to phonetic and aesthetic matters or to certain national customs is not a priority: all these considerations are, indeed, ignored by the machine performing the function.

This part of ISO 233 may be used by anyone who has a clear understanding of the system and is certain that it can be applied without ambiguity. The result obtained will not give a correct pronunciation of the original text in a person's own language, but it will serve as a means of finding automatically the original graphism and thus allow anyone who has knowledge of the original language to pronounce it correctly. Similarly, one can only pronounce correctly a text written in, for example, English or Polish, if one has a knowledge of English or Polish.

The existence in Perso-Arabic script of short vowels and other diacritical marks, which are pronounced but often not written somewhat complicates this reading of the text but as those with knowledge of the language can read and mentally fill in the missing signs/sounds when reading the original script, so they can with the transliterated version, e.g. the word رَدِب, which consists of 3 consonants and 2 short vowels (transliteration: pedar) when written without short vowels would be رَدِب (transliteration: pdr).

To address the issue of short vowels and other signs that are unwritten, and the fact that some characters perform more than one function (e.g. characters that can function as either a vowel or a consonant), ISO 233-3 incorporates three levels: strict and fully reversible, univocal, with short vowels only transliterated if written in the source text; strict and fully reversible, univocal, with short vowels included for clarity, regardless of their presence or absence in the source text; and a modified version of the system that while not fully reversible includes the short vowels and takes account of the different functions performed by some characters (see 4.1 and 4.2 for further details).

The adoption of this part of ISO 233 for international communication leaves every country free to adopt for its own use a national standard which may be different, on condition that it is compatible with this part of ISO 233. The system proposed herein should make this possible and be acceptable to international use if the graphisms it creates are such that they may be converted automatically into the graphisms used in any strict national systems.

The adoption of national standards compatible with this part of ISO 233 will permit the representation, in an international publication, of the morphemes of each language according to the customs of the country where it is spoken. It will be possible to simplify this representation in order to take into account the number of the character sets available on different kinds of machines.

iTeh STANDARD PREVIEW **(standards.iteh.ai)**

[oSIST ISO/DIS 233-3:2021](https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021)

<https://standards.iteh.ai/catalog/standards/sist/288e79a1-a918-4868-ba9b-3e2076602a04/osist-iso-dis-233-3-2021>

Information and documentation — Transliteration of Arabic characters into Latin characters —

Part 3: Persian language — Transliteration

1 Scope

This part of ISO 233 establishes a system for the transliteration into Latin characters of the Arabic characters (often called Perso-Arabic script) used to write the Persian language. This simplification of the stringent rules established by ISO 233:1984 is especially intended to facilitate the processing of bibliographic information (e.g. catalogues, indices, citations, etc.).

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of ISO 233. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this part of ISO 233 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

oSIST ISO/DIS 233-3:2021

ISO 233, *Documentation — Transliteration of Arabic characters into Latin characters*

<https://standards.iso.org/iso/233-3:2021>

ISO 233-2, *Information and documentation — Transliteration of Arabic characters into Latin characters — Part 2: Arabic language — Simplified transliteration*

ISO/IEC 10646, *Information technology — Universal coded character set (UCS)*

3 Terms and definitions

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

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1

character

element of an alphabetical or other type of writing system that graphically represents a phoneme, a syllable, a word or even a prosodical characteristic of a given language; it is used either alone (for example, a letter, a syllabic sign, an ideographical character, a digit, a punctuation mark) or in combination (such as an accent or a diacritical mark)

Note 1 to entry: A letter having an accent or a diacritical mark, for example â, è, ö, is therefore a character in the same way as a basic letter.

3.2

vowel

a speech sound produced by unobstructed flow of air through the mouth

ISO/DIS 233-3:2021(E)

3.3

consonant

a speech sound produced by complete or partial closure of the vocal tract

3.4

transliteration

process which consists of representing the *characters* (3.1) of an alphabetical or syllabic system of writing by the characters of a conversion alphabet

3.5

retransliteration

process whereby the *characters* (3.1) of a conversion alphabet are transformed back into those of the converted writing system

3.6

transcription

process whereby the sounds of a given language are noted by the system of signs of a conversion language

3.7

romanization

conversion of non-Latin writing systems to the Latin alphabet

4 Transliteration tables

“Hex” in the following tables refers to hexadecimal character codes in the standard ISO/IEC 10646 (Unicode).

4.1 Table 1 — Consonants¹⁾

The shaded rows, numbered 1b, 30c etc., provide for option three described above, a modified but not fully-reversible transliteration. They reference the use of consonant characters as vowels (see also 4.2 vowels). Where full reversibility is required only the main numbered rows should be used and the shaded rows ignored.

No.	Persian character	Hex	Transliteration into Latin character	Hex
1 ¹⁾	ا	0627	â	00E2
1b	ا	0627	<i>Omit in word-initial position (see 4.2)</i>	-
1c	ا	0627	<i>ā in word-medial or final position</i>	0101
2	ب	0628	b	0062
3	پ	067E	p	0070
4	ت	062A	t	0074
5	ث	062B	s	0073+0331
6	ج	062C	j	006A
7	چ	0686	č	010D
8	ح	062D	h	1E25
9	خ	062E	x	0078
10	د	062F	d	0064
11	ذ	0630	z	1E95
12	ر	0631	r	0072

1) For the transliteration of hamzeh see 4.4.

No.	Persian character	Hex	Transliteration into Latin character	Hex
13	ز	0632	z	007A
14	ژ	0698	ž	017E
15	س	0633	s	0073
16	ش	0634	š	0161
17	ص	0635	ş	1E63
18	ض	0636	ẓ	017C+0324
19	ط	0637	ţ	1E6D
20	ظ	0638	ẓ	1E93
21	ع	0639	‘	02BB (2018) ²
22	غ	063A	ġ	0121
23	ف	0641	f	0066
24	ق	0642	q	0071
25	ک	06A9	k	006B
26	گ	06AF	g	0067
27	ل	0644	l	006C
28	م	0645	m	006D
29	ن	0646	n	006E
30	و	0648	v ³	0076
30b	و	0648	ū when functioning as a long vowel	016B
30c	و	064F+0648	ow	006F+0077
30d	وا	0627+0648	ū when occurring in word-initial position	016B
31	ه	0647	h	0068
32	ی	06CC	Y ⁴	0079
32b	ی	06CC	ī when functioning as a long vowel	012B
32c	ی	0650+06CC	ey	0065+0079
32d	ی	06CC	-ye when functioning as ez āfeh (see note 8)	0079+0065
32e	یا	0627+06CC	ī when occurring in word-initial position	012B
33	ی	06CC (0649)	á when representing alif maqsurah (in words of Arabic origin) *	00E1

¹ For transliteration of *alef* with *maddeh* see 4.2; and with *hamzeh* see 4.4

² 02BB is recommended. Some transliterations use 2018 so this may be encountered

³ In the modified transliteration, this character should be transliterated as ‘v’ when functioning as a consonant.

⁴ In the modified transliteration, this character should be transliterated as ‘y’ when functioning as a consonant.

* The *alif maqsurah* is a feature of loan words and names of Arabic origin. It should have a different Unicode encoding (0649) but in appearance it looks the same as the Persian *ye* (06CC) and the latter is likely to occur in actual usage.

Different position forms of Persian characters (initial, medial, final and separate) are shown in [Annex A](#).