ICAO

# Doc 9303

# Machine Readable Travel Documents

Eighth Edition, 2021

## Part 3: Specifications Common to all MRTDs

Approved by and published under the authority of the Secretary General

INTERNATIONAL CIVIL AVIATION ORGANIZATION

ICAO

# Doc 9303

# Machine Readable Travel Documents

Eighth Edition, 2021

## Part 3: Specifications Common to all MRTDs

Approved by and published under the authority of the Secretary General

INTERNATIONAL CIVIL AVIATION ORGANIZATION

Published in separate English, Arabic, Chinese, French, Russian
and Spanish editions by the
INTERNATIONAL CIVIL AVIATION ORGANIZATION
999 Robert-Bourassa Boulevard, Montréal, Quebec, Canada H3C 5H7

Downloads and additional information are available at www.icao.int/Security/FAL/TRIP

**Doc 9303, *Machine Readable Travel Documents***
**Part 3 — *Specifications Common to all MRTDs***
Order No.: 9303P3
ISBN 978-92-9265-333-0 (print version)
ISBN 978-92-9275-314-6 (electronic version)

© ICAO 2021

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without prior permission in writing from the International Civil Aviation Organization.

## AMENDMENTS

Amendments are announced in the supplements to the *Products and Services Catalogue;* the Catalogue and its supplements are available on the ICAO website at www.icao.int. The space below is provided to keep a record of such amendments.

## RECORD OF AMENDMENTS AND CORRIGENDA

| AMENDMENTS | | |
|---|---|---|
| No. | Date | Entered by |
| 1 | 14/11/22 | ICAO |
| 2 | 20/3/24 | ICAO |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

| CORRIGENDA | | | |
|---|---|---|---|
| No. | Date | Language | Entered by |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of ICAO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

# TABLE OF CONTENTS

# 1. SCOPE

Part 3 defines specifications that are common to TD1, TD2 and TD3 size machine readable travel documents (MRTDs) including those necessary for global interoperability using visual inspection and machine readable (optical character recognition) means. Detailed specifications applicable to each form factor appear in Doc 9303, Parts 4 through 7.

Part 3 shall be read in conjunction with:

- Part 1 — *Introduction*;
- Part 2 — *Specifications for the Security of the Design, Manufacture and Issuance of MRTDs*;

and the relevant form factor specific part:

- Part 4 — *Specifications for Machine Readable Passports (MRPs) and other TD3 Size MRTDs*;
- Part 5 — *Specifications for TD1 Size Machine Readable Official Travel Documents (MROTDs)*;
- Part 6 — *Specifications for TD2 Size Machine Readable Official Travel Documents (MROTDs)*; and
- Part 7 — *Machine Readable Visas*.

These specifications also apply to machine readable travel documents that contain a contactless IC i.e. electronic machine readable travel documents (eMRTDs). Specifications solely for eMRTDs are contained in the following parts of Doc 9303:

- Part 9 — *Deployment of Biometric Identification and Electronic Storage of Data in MRTDs*;
- Part 10 — *Logical Data Structure (LDS) for Storage of Biometrics and other Data in the Contactless Integrated Circuit (IC)*;
- Part 11 — *Security Mechanisms for MRTDs*; and
- Part 12 — *Public Key Infrastructure for MRTDs*.

# 2. PHYSICAL CHARACTERISTICS OF MRTDS

Issuing States and organizations may choose the materials to be used in the production of their travel documents. Nevertheless, no materials shall adversely affect any other component in the MRTD, and the MRTD shall, in normal use throughout its period of validity, meet the following requirements:

- *Deformation.* The MRTD shall be of a material that bends (not creases), i.e., deformation due to normal use can be flattened by the reading device without impairing the use of the MRTD or the functioning of the reader;

- *Toxicity.* The MRTD shall present no toxic hazards in the course of normal use, as specified in [ISO/IEC 7810];

- *Resistance to chemicals.* The MRTD shall be resistant to chemical effects arising from normal handling and use, except where chemical sensitivity is added for security reasons;

- *Temperature stability.* The MRTD shall remain machine readable at operating temperatures ranging from –10°C to +50°C (14°F to 122°F). The MRTD should not lose its functionality after being exposed to temperatures ranging from –35°C to +80°C (–31°F to 176°F);

- *Humidity.* The MRTD shall be machine readable at a relative air humidity ranging from 5 per cent to 95 per cent, with a maximum wet bulb temperature of 25°C (77°F), as specified in [ISO/IEC 7810]. The MRTD should not lose its reliability after being stored at, or exposed to, a relative air humidity ranging from 0 per cent to 100 per cent (non-condensing);

- *Light.* The MRTD shall resist deterioration from exposure to light encountered during normal use, as specified in [ISO/IEC 7810].

# 3. VISUAL INSPECTION ZONE (VIZ)

The Visual Inspection Zone of an MRTD comprises the mandatory and optional data elements designed for visual inspection. The optional data elements, together with the mandatory data elements, accommodate the diverse requirements of issuing States and organizations while maintaining sufficient uniformity to ensure global interoperability for all MRTDs.

## 3.1 Languages and Characters

Latin-alphabet characters, i.e. A to Z and a to z, and Arabic numerals, i.e. 1234567890 shall be used to represent data in the VIZ. Diacritics are permitted. Latin-based national characters listed in Section 6.A “Transliteration of Multinational Latin-based Characters”, e.g. Þ and ß, may also be used in the VIZ without transliteration. When mandatory data elements are in a language that does not use the Latin alphabet, a transcription or transliteration shall also be provided.

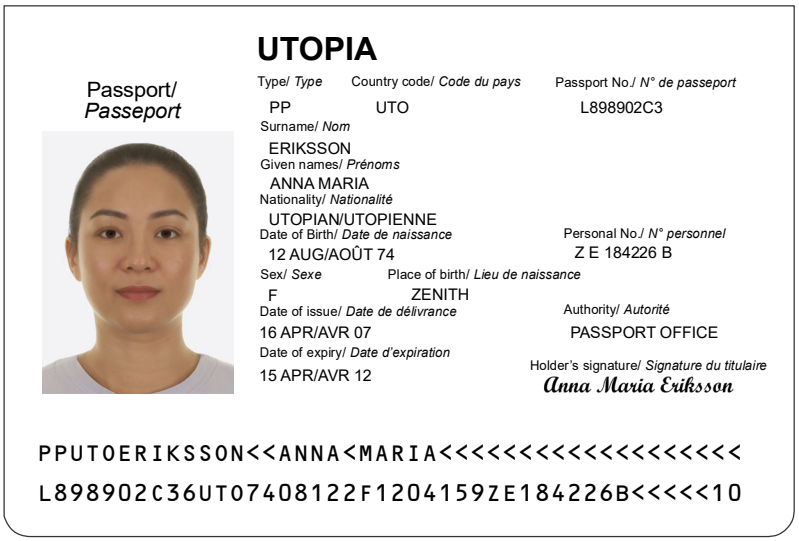


**Figure 1. Example of a VIZ and MRZ from an MRTD**

States that use other than Arabic numerals to represent numerical data in the VIZ shall provide a translation into Arabic numerals.

In the interests of facilitation, optional data elements should be entered in both the national language/working language of the issuing organization and either English, French or Spanish. Optional data in Zone VI may be entered entirely in the national script and/or language.

When including a translation, the different language shall be separated by an oblique character.

Punctuation may be included in the VIZ. For details, see Doc 9303-4.

### 3.2 Typeface and Type Size

The horizontal printing density, the typeface, the type size, the font and the vertical line spacing in the VIZ are at the discretion of the issuing State or organization. For good legibility, a type size with 10 characters per 25.4 mm (1.0 in) is recommended. A maximum of 15 characters per 25.4 mm (1.0 in) should not be exceeded. This type size has been chosen as the smallest in which information is clear and legible to a person with normal eyesight.

Use of upper-case characters is recommended. However, where a name includes a prefix, an appropriate mixture of upper- and lower-case characters may be used in the prefix (see 3.4 in this section).

Diacritical marks (accents) may be used with either lower- or upper-case characters at the option of the issuing State or organization.

### 3.3 Captions/Fields

Captions shall be used to identify all fields for mandatory data elements in the VIZ except as specified in the data element directories for each form factor in Doc 9303, Parts 4 to 7.

Captions may be in the official language of the issuing State or working language of the issuing organization. When such language uses the Latin alphabet, straight font style should be used to print the captions.

Where the official language of the issuing State or working language of the issuing organization is not English, French or Spanish, the printed caption shall be followed by an oblique character (/) and the equivalent of the caption in English, French or Spanish. An italic font style should be used for the second language.

Where the official language of the issuing State or working language of the issuing organization is English, French or Spanish, the issuing State or organization should use one of the other two languages to print the caption following the oblique (/) character. An italic font style should be used for the second language.

Captions shall be printed in a clear, linear type font in a size of 1.0 mm to 1.8 mm (0.04 in to 0.07 in).

When an optional field is not used, the caption shall not appear on the travel document.

### 3.4 Convention for Writing the Name of the Holder

The name of the holder is generally represented in two parts; the primary identifier and the secondary identifier.

The issuing State or organization shall establish which part of the name is the primary identifier. This may be the family name, the maiden name or the married name, the main name, the surname, and in some cases, the entire name where the holder's name cannot be divided into two parts. This shall be entered in the field for the primary identifier in the VIZ. It is recommended that upper-case characters be used, except in the case of a prefix, e.g. "von," "Mc" or "de la," in which case a mixture of upper- and lower-case is appropriate.

The remaining parts of the name are the secondary identifier. These may be the forenames, familiar names, given names, initials, or any other secondary names. These names shall be written in the field for the secondary identifier in the VIZ. It is recommended that upper-case characters be used throughout.

If a single field is used for the name, then the secondary identifier shall be separated from the primary identifier by a single comma (,). A comma is not needed if multiple fields are used.

Prefixes and suffixes including titles, professional and academic qualifications, honours, awards, and hereditary status, should not be included in the VIZ. However, if an issuing State or organization considers such a prefix or suffix to be legally part of the name, the prefix or suffix can appear in the VIZ. Numeric characters should not be written in the name fields of the VIZ; however, where the use of numeric characters is a legal naming convention in the issuing State, these should be represented in Roman numerals. Any prefixes, suffixes or Roman numerals shall be entered in the secondary identifier field.

National characters may be used in the VIZ. If the national characters are not Latin-based, a transcription or transliteration into Latin characters shall be provided.

### 3.5 Representation of Issuing State or Organization

Where the name of the issuing State or organization and/or the location of the issuing office or authority are in a language that does not use Latin characters, the name of the State or other location shall appear in the national language/working language of the issuing organization and also shall be either:

- transliterated into Latin characters; or
- translated into one or more languages (at least one of which must be English, French or Spanish) by which the name may be more commonly known to the international community.

The name in the different languages shall be separated by an oblique character (/) followed by at least one blank space.

Where the name of the issuing State or organization or location of the issuing office or authority is in a language that uses the Latin alphabet, but the name is more familiar to the international community in its translation into another language or languages (particularly English, French or Spanish), the name should be accompanied by one or more translations. The name in the different languages shall be separated by an oblique character (/) followed by at least one blank space.

### 3.6 Representation of Nationality

The nationality of the holder in the VIZ, in documents where this field is mandatory, shall be represented either by the three-letter code (see Section 5) or in full at the discretion of the issuing State or organization.

If the nationality is written in full and the national language of the issuing State or working language of the issuing organization is a language that does not use Latin characters, the nationality shall appear in the national/working language and also shall be either:

- transliterated into Latin characters; or
- translated into one or more languages (at least one of which must be English, French or Spanish) by which the nationality may be more commonly known to the international community.

The nationality in the different languages shall be separated by an oblique character (/) followed by at least one blank space.

Where the national language of the issuing State or working language of the issuing organization uses the Latin alphabet, but the nationality is more familiar to the international community in its translation into another language or languages (particularly English, French or Spanish), the nationality in the national/working language should be accompanied by one or more translations. The nationality in the different languages shall be separated by an oblique character (/) followed by at least one blank space.

### 3.7 Representation of Place of Birth

Inclusion of the place of birth is optional. If the place of birth is included it may be represented by the town, the city, the suburb and/or the State.

If the town, city or suburb is included and the national language of the issuing State or working language of the issuing organization is a language that does not use Latin characters, the town, city or suburb shall appear in the national/working language and also shall be either:

- transliterated into Latin characters; or
- translated into one or more languages (at least one of which must be English, French or Spanish) by which it may be more commonly known to the international community.

The town, city or suburb in the different languages shall be separated by an oblique character (/) followed by at least one blank space.

Where the national language of the issuing State or working language of the issuing organization uses the Latin alphabet, but the town, city or suburb is more familiar to the international community in its translation into another language or languages (particularly English, French or Spanish), the town, city or suburb in the national/working language should be accompanied by one or more translations. The town, city or suburb in the different languages shall be separated by an oblique character (/) followed by at least one blank space.

If the State is included, its three-letter code shall be represented as outlined in Section 5, except where no code for the State of Birth exists, in which case the name shall be written in full, and the requirements for translation and transliteration identified for town, city and suburb above apply.

*Note.— When choosing to include or omit the Place of Birth, the travel document issuing State or organization should take into consideration any current political sensitivities linked to the State or territory and whether it is a State or territory recognized by visa-issuing authorities in other countries.*

### 3.8 Representation of Dates

Dates in the VIZ of the MRTD shall be entered in accordance with the Gregorian calendar as follows:

***Day***

Days shall be shown by a two-digit number, i.e. the dates from one to nine shall be preceded by a zero. This number may be followed by a blank space before the month or may be followed immediately by the month, with no blank space.

***Month***

The month may be printed in full in the national language of the issuing State or working language of the issuing organization or abbreviated, using up to four character positions.

Where the national language of the issuing State or working language of the issuing organization is not English, French or Spanish, the month shall be followed by an oblique character (/) and the month or the abbreviation of the month up to four character positions, in one of the three languages, as shown in the table below.

Where the national language of the issuing State or working language of the issuing organization is English, French or Spanish, the issuing State or organization may also use one of the other two languages (shown in Table 1) following the oblique character (/).

The month may alternatively be printed in numerical form at the discretion of the issuing State or organization, particularly where this might facilitate the use of the MRTD by States using other than the Gregorian calendar. In this case the date would be written DDnMMnYY or DDnMMnYYYY, where n = a single blank space or a period.

**Table 1. Abbreviations of Months in English, French and Spanish**

| *Month* | *English* | *French* | *Spanish* |
|---|---|---|---|
| JANUARY | JAN | JAN | ENE |
| FEBRUARY | FEB | FÉV | FEB |
| MARCH | MAR | MARS | MAR |
| APRIL | APR | AVR | ABR |
| MAY | MAY | MAI | MAYO |
| JUNE | JUN | JUIN | JUN |
| JULY | JUL | JUIL | JUL |
| AUGUST | AUG | AOÛT | AGO |
| SEPTEMBER | SEP | SEPT | SEPT |
| OCTOBER | OCT | OCT | OCT |
| NOVEMBER | NOV | NOV | NOV |
| DECEMBER | DEC | DÉC | DIC |

***Year***

The year will be shown by the last two or four digits and may be preceded by a blank space, or it may follow the month immediately with no blank space. Both formats are acceptable.

When the month is represented numerically, the issuing State or organization may use the two- or four-digit representation of the year, and separate the month and year by a blank space or a period.

*Note.— States are encouraged to use the four digit representation of the year for all date formats*.

Examples:

12 July 1942 on an MRTD data page issued in Italian with French translation of the month could appear as:

12nLUGn/JUILn1942

where n = a single blank space, i.e. 12 LUG/JUIL 1942

or

12nLUGn/JUILn42
where n = a single blank space, i.e. 12 LUG/JUIL 42

or

12 July 1942 or 12 July 42 (using English only)

or

12JUIL1942 or 12JUIL42 (using French abbreviation)

or

12JUL 1942 or 12JUL 42 (using English or Spanish abbreviation)

or

12 07 42 or 12.07.42 (using numerical format).

or

12 07 1942 or 12.07.1942 (using numerical format with four-digit year).

*Unknown date of birth*. Where a date of birth is completely unknown, that data element shall appear in the date format used for dates of birth by the issuing State or organization but with Xs representing unknown elements (numbers and/or letters) of the date.

Examples:
XXnXXnXX
XXnXXnXXXX
XXnXXXnXX where n = a single blank space or a period (if numerical format is used).

If only part of the date of birth is unknown, only that part (day, month, year) of the date shall be represented by Xs as per the date format used by the issuing State or organization.

### 3.9 Displayed Identification Features of the Holder

Doc 9303 identifies mandatory and optional identification feature(s) of the holder which must be displayed within the VIZ, i.e. facial image, signature or usual mark and/or single-digit fingerprint for each type of MRTD as well as the position, dimensions and scaling for the identification features.

### *3.9.1 Displayed facial image*

To ensure compatibility with facial recognition systems, portrait capturing shall comply with relevant specifications outlined in [ISO/IEC 39794-5].

The displayed facial image, whether provided in paper or digital formal, shall:

- be digitally printed in the MRTD;
- depict a true likeness of the rightful holder of the MRTD; and
- not be digitally altered or enhanced to change the subject's appearance in any way.

Necessary measures shall be taken by the issuing State or organization to ensure that the displayed portrait is resistant to forgery and substitution.

#### *3.9.1.1 Image Printing for Portrait Submission*

The physical portrait shall yield an accurate recognizable representation of the subject. The quality of the original captured image should at least be comparable to the minimum quality acceptable for paper photographs (resolution comparable to 6 – 8 line pairs per millimetre). To achieve this comparable image quality in a digital reproduction, careful attention shall be given to the image capture, processing, digitization, compression and printing technology and the process used to produce the portrait. The printing process shall maintain the width to height ratio of the original image.

*Note.— Many issuing States use a printing/re-scanning procedure for document application. This approach is acceptable; however, caution should be taken to ensure quality according to the guidelines and requirements indicated below and in [ISO/IEC 39794-5]. If a new design of the application process is considered, digital submission should be taken into consideration as the preferred technology whenever possible.*

*Print resolution.* The printing process should produce a smooth image that is capable of accurately rendering fine contrasted facial details, such as wrinkles and moles. All flesh tones from both light- and dark-complexioned subjects should be printed accurately and limited hot spots or shadow drop-outs apparent. Smooth facial details should be rendered without noticeable posterization or contouring.

*Saturation and colour*. With the exception of glare or glints caused by small areas of possible specular (mirror-like) reflection, only a small portion of the printed image should be saturated in white or black. Excluding the background area, using luminosity, the number of fully saturated 0 value pixels shall be less than 0.1%, and the number of fully saturated 255 value pixels shall be less than 0.1%.

No portion of the background or the subject's garments should be printed fully white and details should be apparent in dark shadow regions.

Printed photos shall be colour images having balanced colour channels. It may be assumed that the capture device (digital camera or scanner) is correctly white balanced.

*Paper properties and portrait size.* The photograph shall be on photo-quality paper. Examples of such paper are the following (other technologies with similar properties are also acceptable):

- Instant photographic standard gloss,
- Dye sublimation photographic semi-gloss,
- Silver halide photographic semi-gloss, or

- Drylab photographic inkjet bases standard gloss.

The photograph paper shall have a low roughness, non-structured surface (no pearl or silkscreen effect). Submitted portraits should have a minimum width of 35 mm. The inter eye distance (IED) should be at least 10 mm.

Newly designed application processes still relying on printed portrait submission should consider using larger photo sizes, such as, e.g., 7 cm by 10 cm. Larger photos reduce the risk of quality losses in the process chain. However, a switch to larger photos will have process implications to be considered.

*Moiré or visible dot patterns*. Digitization of printed photos may introduce artefacts, such as moiré, and certain printing processes may exacerbate the generation of such artefacts. The printing process employed should allow accurate face recognition when its prints are scanned with a document scanner at a spatial sampling rate of 120 pixels per centimetre (300 pixels per inch) in each axis.

If a printed photo has been produced through a periodic half-toning process, scanning the photo will almost invariably introduce moiré patterns. Thus, those printers, such as inkjet and laser printers, which inherently employ half-toning to simulate continuous tones, should use non-periodic (or dithered) half-toning methods. Furthermore, the printing process should not produce dot patterns visible to the unaided eye.

*Note.— It is often useful to provide a transparent template to a person responsible for photo quality evaluation. The template would display the limits of head size and rotation (roll) and, when superimposed on the photo, could assist in the determination of whether a printed photo is compliant to the requirements. Samples of such tools can be found in [ISO/IEC 39794-5].*

*3.9.1.2 Scanning of Submitted Portraits*

Submitted portraits shall comply with the relevant specifications outlined in section 3.9.1.1 and in [ISO/IEC 39794-5].

*Properties of the submitted portrait.* Submitted portraits should be 45.0 mm x 35.0 mm (1.77 in x 1.38 in) in dimension. This will provide adequate resolution for scaling to required size for use on the MRTD while having adequate resolution for facial recognition purposes.

Multiple scan/print steps shall not be used in an application process. If the portrait has been printed for submission and is subsequently scanned, all remaining production steps shall be digital.

A submitted portrait shall have been captured within the last six months before application, as outlined in [ISO/IEC 39794-5]. Portraits with a capture time dating back more than three months should not be accepted. Issuers should consider the use of the metadata encoded with the digital image to assure that the photograph is recent.

If printed portraits are submitted, evidence on the capturing date should be requested. This may be the printed manufacturing date on the back side of the photo, or a dated invoice of the photographer. The complete card should be provided if the portrait is part of a photo card (e.g., a 10x15 print containing 2x2 images).

The submitted portrait shall be clean, not bent, not scratched, not folded and not damaged. There shall be no ink marks or creases on the printed portrait.

Where the portrait is supplied to the issuing authority in digital form, the requirements specified by the issuing authority must be adhered to.

*Pixel count and Modulation Transfer Function (MTF)*. The final scanned images shall have a pixel count as specified in [ISO/IEC 39794-5]. MTF20 should occur at 4,7 cy/mm or higher for scanners. The scanner's MTF should be the same in both axes. Image enhancement processing using either built-in hardware or software-based image sharpening generally should not be used to boost the MTF.

Example:

The optical properties of the image can be maintained if the digital camera original image MTF20 should occur at approximately 80% or higher of the Nyquist frequency when using the MTF test method according to [ISO 12233]. The size of a freckle/mole that should be detectable in face photos is 2 to 3 mm. Rulers make good fiducial markers to make measurements on the image.

The MTF analysis should be done using the appropriate target from ISO 12233. Informative examples can be found in [ISO/IEC 39794-5].

Example:

A typical printed image with 10 mm IED should be scanned at a sampling rate of at least 300 ppi.

The MTF will be limited by the size of the paper photo and the resolution (fineness of detail) therein. To obtain higher resolution from scanned images, the issuer should consider increasing the size requirement for printed portraits.

Particular care shall be taken in the acquisition process in order to avoid any kind of image dimensional stretching in any direction.

The width to height ratio of the final image is defined by the application process of the issuer, a typical value is 7:9. Necessary modifications shall be made by cropping and shall not be made by stretching.

*Colour, sharpness, and saturation.* The scanned portrait shall have the same colour as the submitted one. The human eye shall not be able to detect differences between the portrait and scanned result when viewed on a colour corrected display device and under daylight conditions. The portrait shall have appropriate brightness and contrast that show skin tones naturally.

The number of quantization levels should be at least 256 levels per colour, with three colours per pixel. The scanned image shall comply with the colour requirements outlined in [ISO/IEC 39794-5].

Since red-green-blue (RGB) colour space and its derivatives are inherently device-dependent, the scanner's output shall be converted to one of the well-defined, device-independent colour spaces as outlined in [ISO/IEC 39794-5].

Saturation occurs when significant numbers of pixels have values that are at the limits of quantization, i.e., at the levels of 0 or 255, if quantization of eight bits per colour is employed. Acceptable scanned face images should not have a significant number of pixels in saturation in the facial region.

The scanned portrait shall be centred, clear and in sharp focus with no shadows. It shall not have visible compression artefacts.

*3.9.1.3 Image Printing for MRTD production*

The portrait printed on the data page shall be derived from the same digital image source as the image stored electronically in the MRTD. However, due to the influence of printing technologies as well as to the application of several security features to the portrait and to the data page, the image may not be exactly the same. Examples for possible deviations are the printer resolution, removed background in the printed portrait, image enhancements, dithering of grayscale content, or guilloches occurring in the print.

*Note.— The implementation of the portrait on or into the MRTD should be done considering the properties of the different materials and technologies in use. It is possible that the printing technology itself introduces specific features into the printed portrait.*

The digital reproduction shall yield an accurate recognizable representation of the subject. To achieve such image quality in a document data page, careful attention shall be given to the processing, compression and printing technology and the process used to produce the portrait. Printed portraits have specific features that depend on categories of printing technologies.

The primary printed image on the MRTD may be either greyscale or colour.

Any face printing process should produce a smooth image that is capable of accurately rendering fine facial details, such as contrasted wrinkles, contrasted moles, and contrasted scars, as small as two millimetres in diameter on the face positioned anywhere in the printed image area. Such details shall be detectable when viewed with the naked eye at a distance of 0.3 m.

All flesh tones from both light- and dark-complexioned subjects should be printed accurately and no hot spots or shadow drop-out should be apparent. Smooth facial details should be rendered without posterization or contouring.

*Size*. The portrait dimensions should meet the specifications outlined in [ISO/IEC 39794-5]. Necessary modifications shall be made by cropping and shall not be made by stretching. In cases where the background has been removed from the image, the correct width or height of the printed image may be impossible to determine. In such cases, the height-to-width ratio is considered to be maintained if the ratio between *IED* and eye-to-mouth distance (*EM)* of the printed image is the same as of the portrait.

*Tonal range*. The tonal range of the printed image shall not interfere with facial details important for human identification when making a comparison of the printed image to the document holder.

*Moiré or visible dot patterns*. Moiré or dot patterns in the printed image should be minimized. Any such patterns in the printed image shall not interfere with facial details important for human identification when making a comparison of the printed image to the document holder.

*Portrait placement in an MRTD and coexistence with security printing*. The printed portrait shall be centred within Zone V, with the crown (top of the head ignoring any hair) nearest the top edge of the MRTD. The crown-to-chin portion of the facial image shall be 70 to 80 per cent of the longest dimension defined for Zone V, maintaining the aspect ratio between the crown-to-chin and ear-to-ear details of the face of the holder. The 70 to 80 per cent requirement may mean cropping the picture so that not all the hair is visible.

If present, a digitally printed reproduction shall coexist with background security treatment(s) located within Zone V, i.e., the background security printing shall not interfere with proper viewing of the displayed portrait, and vice versa, yet still offer protection to the displayed portrait.

*Coexistence with final preparation treatment(s) of the MRTD*. A displayed portrait shall coexist with final preparation treatment(s), i.e. final preparation treatment(s) shall not interfere with proper viewing of the displayed portrait, and vice versa.

*Border*. A border or frame shall not be used to outline a digitally printed reproduction.

*3.9.1.4 Compliance with international standards*

The photograph shall comply with the appropriate definitions set out in [ISO/IEC 39794-5].

### *3.9.2 Displayed signature or usual mark*

A displayed signature or usual mark, the acceptability of which is at the issuing State or organization's discretion, appears in Zone IV. A displayed signature or usual mark shall be an original created on the MRTD, a digitally printed reproduction of an original or, where permitted by specifications defined in Doc 9303 Parts 4 to 7 specific to the preparation of the different types of MRTDs, on a substrate that can be securely affixed to the MRTD. Necessary measures shall be taken by the issuing State or organization to ensure that the displayed signature or usual mark is resistant to forgery and substitution. The displayed signature or usual mark shall meet the following requirements.

*Orientation*. The displayed signature or usual mark shall be displayed with its A-dimension parallel to the reference (longer) edge of the MRTD as defined in Figure 2.

*Size*. The displayed signature or usual mark shall be of such dimensions that it is discernible by the human eye (i.e. reduced in size by no more than 50 per cent), and the aspect ratio (A-dimension to B-dimension) of the original signature or usual mark is maintained.

*Scaling for reproduction using digital printing*. In the event the displayed signature or usual mark is scaled up or scaled down, the aspect ratio (A-dimension to B-dimension) of the original signature or usual mark shall be maintained.

*Cropping for reproduction using digital printing*. The issuing State or organization should take steps to eliminate or minimize cropping.

*Colour*. The displayed signature or usual mark shall be displayed in a colour that affords a definite contrast to the background.

*Borders*. Borders or frames shall not be permitted or used to outline the displayed signature or usual mark.

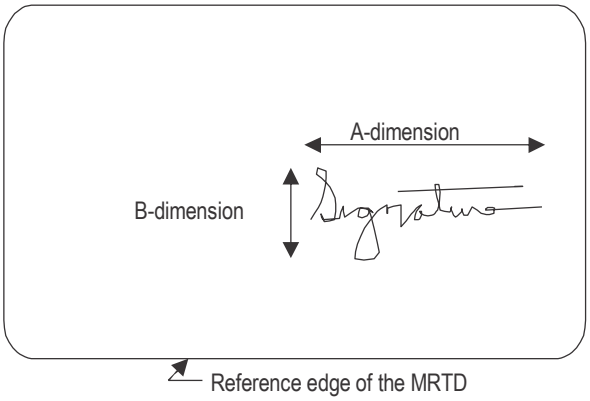


**Figure 2. Orientation of the displayed signature or usual mark**

### *3.9.3 Displayed single-digit fingerprint*

A displayed single-digit fingerprint, if required by the issuing State or organization, shall be either an original created on the MRTD substrate by the holder or, more probably, a digitally printed reproduction of an original. Necessary measures shall be taken by the issuing State or organization to ensure that the single-digit fingerprint is resistant to forgery and substitution. The single-digit fingerprint shall meet the following requirements.

*Orientation*. The A-dimension (width) of the displayed single-digit fingerprint shall be parallel to the reference edge of the MRTD as defined in Figure 3. The top of the finger shall be that portion of the single-digit fingerprint furthest away from the reference edge of the MRTD. (See Doc 9303-6, Figure 10 and Figure 12.)

*Size*. The displayed single-digit fingerprint shall be a one-to-one replication (A-dimension versus B-dimension) of the original print.

*Scaling for reproduction using digital printing*. Scaling of a single-digit fingerprint shall not be permitted.

*Cropping for reproduction using digital printing*. The issuing State or organization should take steps to eliminate or minimize cropping.

*Colour*. The displayed single-digit fingerprint shall be displayed in a colour that affords a definite contrast to the background.

*Borders*. Borders or frames shall not be permitted or used to outline the displayed single-digit fingerprint.

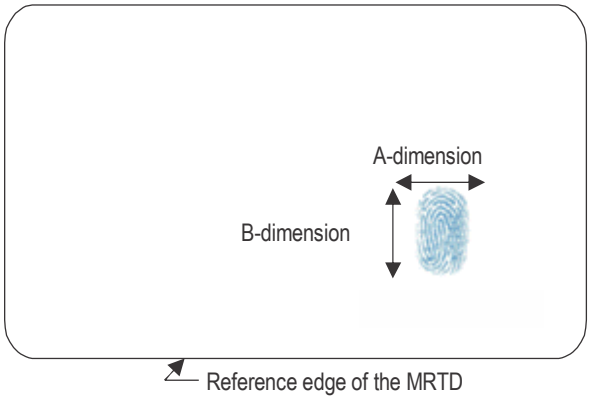


**Figure 3. Orientation of the displayed single-digit fingerprint**

# 4. MACHINE READABLE ZONE (MRZ)

## 4.1 Purpose of the MRZ

MRTDs produced in accordance with Doc 9303 incorporate an MRZ to facilitate inspection of travel documents and reduce the time taken up in the travel process by administrative procedures. In addition, the MRZ provides verification of the information in the VIZ and may be used to provide search characters for a database inquiry. As well, it may be used to capture data for registration of arrival and departure or simply to point to an existing record in a database.

The MRZ provides a set of essential data elements in a format, standardized for each type of MRTD that can be used by all receiving States regardless of their national script or customs.

The data in the MRZ are formatted in such a way as to be readable by machines with standard capability worldwide. It must be stressed that the MRZ is reserved for data intended for international use in conformance with international standards for MRTDs. The MRZ is a different representation of the data than is found in the VIZ.

## 4.2 Properties of the MRZ

The data in the MRZ must be visually readable as well as machine readable. Data presentation must conform to a common standard such that all machine readers configured in conformance with Doc 9303 can recognize each character and communicate in a standard protocol (e.g. ASCII) that is compatible with the technology infrastructure and the processing requirements defined by the receiving State.

To meet these requirements, OCR-B typeface is the specified medium for storage of data in the MRZ. The MRZ as defined herein is recognized as the machine reading technology essential for global interchange and is therefore mandatory in all types of MRTDs.

## 4.3 Constraints of the MRZ

The only characters allowed in the MRZ are a common set of characters (Figure 4) which can be used by all States. National characters generally appear only in the computer-processing systems of the States in which they apply and are not available globally. They shall not, therefore, appear in the MRZ.

Diacritical marks are not permitted in the MRZ. Even though they may be useful to distinguish names, the use of diacritical marks in the MRZ would confuse machine-reading equipment, resulting in less accurate database searches and slower clearance of travellers.

The number of character positions available for data in the MRZ is limited and varies according to the type of MRTD. The length of the data elements inserted in the MRZ must conform to the size of the respective fields as specified in the MRZ data element directory in the applicable Part 4 to 7 of Doc 9303.

In some instances, names in the MRZ may not appear in the same form as in the VIZ. In the VIZ, non-Latin and national characters may be used to represent more accurately the data in the script of the issuing State or organization. Such characters are not permitted in the MRZ.

### 4.4 Print Specifications

Machine readable data shall be printed in OCR-B type font, size 1, constant stroke width characters, at a fixed width spacing of 2.54 mm (0.1 in), i.e. horizontal printing density of 10 characters per 25.4 mm (1.0 in). Printed characters are restricted to those defined in Figure 4.


**Figure 4. Subset of OCR-B Characters from [ISO 1073-2] for use in machine readable travel documents**

*Note.— For illustrative purposes only – the characters shown are larger than actual size.*

### 4.5 Machine Reading Requirements and the Effective Reading Zone

*Effective reading zone*. A fixed-dimensional reading area (effective reading zone (ERZ) of 17.0 mm × 118.0 mm (0.67 in × 4.65 in)), sized to accommodate the largest MRTD, is defined to allow use of a single machine reader for all sizes of MRTDs. The location of the ERZ is as defined in Figure 5. The provision of the ERZ is not intended to allow additional tolerance for the printing positions defined in Parts 4, 5, 6 and 7 specific to the preparation of the different types of MRTDs. The ERZ is intended to allow for variances due to the manual placement of machine readable visas (MRVs) and the fanning effect of the pages that takes place when reading an interior page of an MRP. It also allows for the reading of MRTDs with either two or three lines of machine readable data.

To combat the threat to travel document security posed by, for example, photocopiers, security features are permitted in the MRZ, and any such security feature shall not interfere with accurate reading of the OCR characters at the B900 range, as defined in [ISO 1831]. While OCR characters must be visible, as specified in 4.2, to ensure that all MRTDs, including those with security features in the MRZ, can be successfully read, the OCR characters in the MRZ shall be machine readable at least in the near infrared portion of the spectrum (i.e. the B900 band defined in [ISO 1831]).

*Note.— The dimensions of the effective reading zone (ERZ) illustrated are based on a standardized ERZ for all machine readable travel documents to allow use of a single machine reader.*

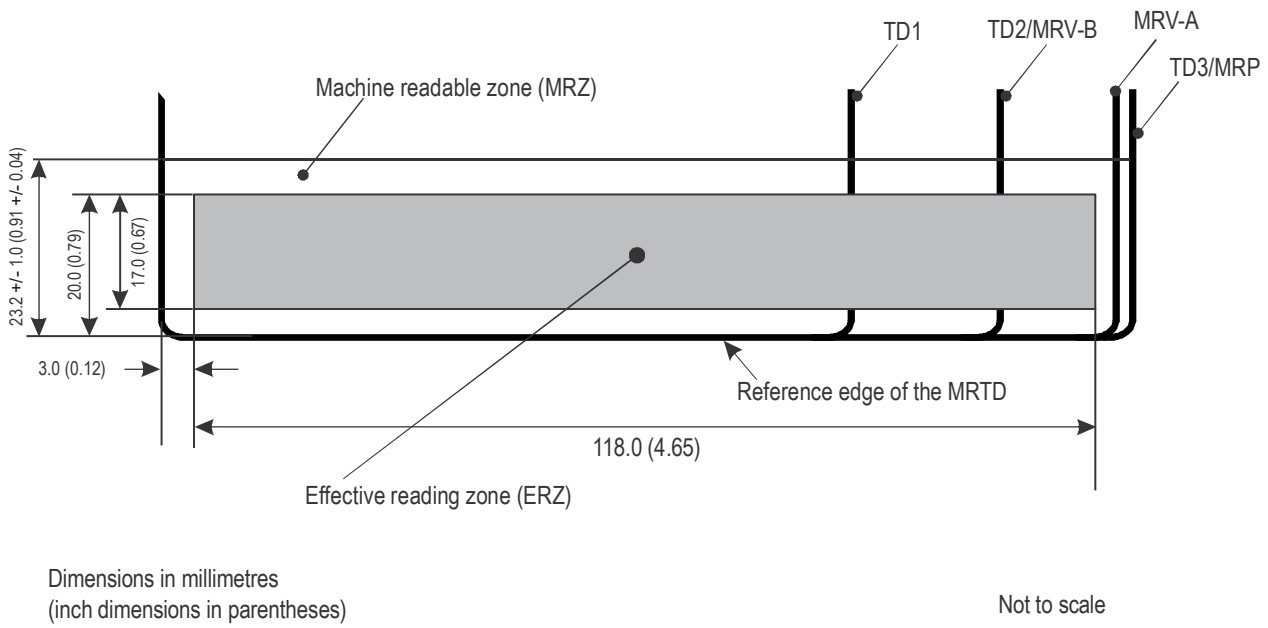


**Figure 5. Schematic diagram of the MRTD effective reading zone**

### 4.6 Convention for Writing the Name of the Holder

To achieve global interoperability, the primary and secondary identifiers in the MRZ shall be printed using upper-case OCR-B characters, illustrated in Figure 4, without diacritical marks, and conform to the number of character positions available. As such, names in the MRZ are represented differently from those in the VIZ. The issuing State or organization shall transliterate national characters using only the allowed OCR-B characters and/or truncate, as specified in the form factor specific Parts 4 to 7 of Doc 9303. Transliteration tables for the most commonly used Latin, Cyrillic and Arabic families of languages are provided in Section 6.

The primary identifier, using the Latin character transliteration (if applicable), shall be written in the MRZ as specified in the form factor specific Parts 4 to 7 of Doc 9303. The primary identifier shall be followed by two filler characters (<<). The secondary identifier, using the Latin character transliteration (if applicable), shall be written starting in the character position immediately following the two filler characters.

If the primary or secondary identifiers have more than one name component, each component shall be separated by a single filler character (<).

Filler characters (<) should be inserted immediately following the final secondary identifier (or following the primary identifier in the case of a name having only a primary identifier) through to the last character position in the machine readable line.

The number of character positions in the name field is limited and differs for the different types of MRTDs. If the primary and secondary identifiers, written in the relevant machine readable line using the above procedure, exceed the available character positions, then truncation shall be carried out using the procedure set out in the form factor specific Parts 4 to 7 of Doc 9303. In all other cases, the name shall not be truncated.

Examples of truncation of names are contained in the form factor specific Parts 4 to 7 of Doc 9303.

Prefixes and suffixes, including titles, professional and academic qualifications, honours, awards, and hereditary status (such as Dr., Sir, Jr., Sr., II and III) shall not be included in the MRZ except where the issuing State considers these to be legally part of the name. In such cases, prefixes or suffixes shall be represented as components of the secondary identifier(s).

Numeric characters shall not be used in the name fields of the MRZ.

Punctuation characters are not allowed in the MRZ. Where these appear as part of a name, they should be treated as follows:

*Apostrophe:*

This shall be omitted; name components separated by the apostrophe shall be combined, and no filler character shall be inserted in its place in the MRZ.

*Example* VIZ: D'ARTAGNAN
MRZ: `DARTAGNAN`

*Hyphen:*

Where a hyphen appears between two name components, it shall be represented in the MRZ by a single filler character (<). (i.e. hyphenated names shall be represented as separate components).

*Example* VIZ: MARIE-ELISE
MRZ: `MARIE<ELISE`

*Comma:*

Where a comma is used in the VIZ to separate the primary and secondary identifiers, the comma shall be omitted in the MRZ, and the primary and secondary identifiers shall be separated in the MRZ by two filler characters (<<).

*Example* VIZ: ERIKSSON, ANNA MARIA
MRZ: `ERIKSSON<<ANNA<MARIA`

Otherwise, where a comma is used in the VIZ to separate two name components, it shall be represented in the MRZ as a single filler character (<).

*Example* VIZ: ANNA, MARIA
MRZ: `ANNA<MARIA`

*Other punctuation characters:*

All other punctuation characters shall be omitted from the MRZ (i.e. no filler character shall be inserted in their place in the MRZ).

### 4.7 Representation of Issuing State or Organization and Nationality of Holder

The three-letter codes referenced in Section 5 shall be used to complete the fields for the issuing State or organization and the nationality of the holder in the MRZ.

### 4.8 Representation of Dates

Dates in the MRZ of the MRTD shall be shown as a six-digit string consisting of the last two digits for the year (YY) immediately followed by two digits for the number of the month (MM) and by two digits for the day (DD). The structure is as follows: YYMMDD.

Following this format, 12 July 1942 will be shown as: 420712.

If all or part of the date of birth is unknown, the relevant character positions shall be completed with filler characters (<).

### 4.9 Check Digits in the MRZ

A check digit consists of a single digit computed from the other digits in a series. Check digits in the MRZ are calculated on specified numerical data elements in the MRZ. The check digits permit readers to verify that data in the MRZ is correctly interpreted.

A special check digit calculation has been adopted for use in MRTDs. The check digits shall be calculated on modulus 10 with a continuously repetitive weighting of 731 731 ..., as follows.

*Step 1*. Going from left to right, multiply each digit of the pertinent numerical data element by the weighting figure appearing in the corresponding sequential position.

*Step 2*. Add the products of each multiplication.

*Step 3*. Divide the sum by 10 (the modulus).

*Step 4*. The remainder shall be the check digit.

For data elements in which the number does not occupy all available character positions, the symbol < shall be used to complete vacant positions and shall be given the value of zero for the purpose of calculating the check digit.

When the check digit calculation is applied to data elements containing alphabetic characters, the characters A to Z shall have the values 10 to 35 consecutively, as follows:

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 |

Data elements included in the check digit calculation and check digit location for each document type are contained in the form factor specific Parts 4 to 7 of Doc 9303. Examples of check digit calculation are found in Informative Appendix A to this Part.

### 4.10 Characteristics of the MRZ

Except as otherwise specified herein, the MRTD shall conform with [ISO 1831] concerning the following matters:

- optical properties of the substrate to be used;
- optical and dimensional properties of the image patterns forming OCR characters; and
- basic requirements related to the position of OCR characters on the substrate.

Machine readable data shall be arranged from left to right in fixed-length fields in two lines (upper and lower) except for TD1 size travel documents where there are three lines (upper, middle and lower). The data are presented in the order specified in the data structure tables in the form factor specific Parts 4 to 7 of Doc 9303 and located on the document as shown in those parts. Data shall be entered in each field, beginning with the left-hand character position.

Where the entered data do not occupy all the character positions specified for the relevant field, the symbol < shall be used to fill the unoccupied positions.

### 4.11 Quality Specifications of the MRZ

In general, the print quality shall conform to [ISO 1831] Range X, except as otherwise provided herein. Except where otherwise noted, all quality specifications set forth hereunder shall conform to the requirements of Section 2 of this Part and shall apply to the MRTD after final preparation and, in the case of visas, after placement in the passport or other travel document.

*Substrate quality*. [ISO 1831], 4.3 through 4.3.2, shall be used for reference only.

*Substrate opacity*. The substrate used, measured before and after final preparation (and for visas, prior to placement in the passport or other travel document), shall be within the definition of at least medium opacity as specified in [ISO 1831], 4.4.1 and 4.4.3.

*Substrate gloss*. The level of gloss is not specified.

*Fluorescence*. The reflectance of the substrate in the visible spectrum shall exhibit no visibly detectable fluorescence when irradiated by ultraviolet light, except where this is a predictable fluorescence for security reasons.

*Alternative substrates*. The aforementioned quality specifications should be followed irrespective of the substrate material.

*Spectral band.* The OCR print shall be legible visually and shall be black (B425 through B680 as defined in [ISO 1831]). The OCR print shall also absorb in the B900 band as defined in [ISO 1831] (i.e. near infrared). This property must test successfully when the characters are machine-read through any protective material that may have been applied to the surface of the document.

*Print contrast signal (PCS)*. After final preparation, the minimum print contrast signal (PCS/min), when measured as specified in [ISO 1831], shall be as follows: PCS/ min ≥ 0.6 at the B900 spectral band.

*Character stroke width*. The stroke width after final preparation shall be as specified for Range X in [ISO 1831], 5.3.1.

*Contrast variation ratio (CVR)*. After final preparation, the CVR should be as is shown for Range X in [ISO 1831], i.e. CVR < 1.50.

*Spots and extraneous marks*. [ISO 1831], 5.4.4.6 and 5.4.5.12 shall apply at the reading surface (see also B.6 of Appendix B and C.5.10 of Appendix C to [ISO 1831]).

*Voids*. The value of "d" as defined in [ISO 1831], 5.4.5.9 shall be equal to 0.4 at the reading surface.

*Line separation*. Refer to the form factor specific Parts 4 to 7 of Doc 9303.

*Line spacing*. Refer to the form factor specific Parts 4 to 7 of Doc 9303.

*Skew of the MRZ lines*. The effect of the actual skew of the MRZ lines and the actual skew of the MRZ characters shall not exceed 3 degrees measured from the reference edge nor shall the skew of MRZ or character misalignment result in the MRZ lines or any part thereof appearing outside the printing zone as defined in the form factor specific Parts 4 to 7 of Doc 9303.

## 5. CODES FOR NATIONALITY, PLACE OF BIRTH, LOCATION OF ISSUING STATE/AUTHORITY AND OTHER PURPOSES

### Part A — Letter Codes

Two- and three-letter codes shall be obtained from the [ISO 3166] maintenance agency - [ISO 3166/MA], ISO's focal point for country codes. These codes are regularly updated in [ISO 3166-1] and are publically available (https://www.iso.org/iso-3166-country-codes.html).

Codes not included in [ISO 3166-1], such as extensions for other States and organizations, or other exceptions, are outlined in the following table:

| *Entity (short name)* | 2-letter code | 3-letter code | *Entity (short name)* | 2-letter code | 3-letter code |
|---|---|---|---|---|---|
| British Overseas Territories Citizen | | GBD | British Protected Person | | GBP |
| British National (Overseas) | | GBN | Germany | DE | D |
| British Overseas Citizen | | GBO | Kosovo[1] | KS | RKS |
| British Subject | | GBS | | | |

### Part B — Other Codes Reserved by ISO 3166/MA

| | | |
|---|---|---|
| European Union (EU) | EU | EUE |

1. The KS and RKS codes are operationally in use, although not reflected in [ISO 3166-1].

**Part C — Codes for Use in United Nations Travel Documents**

| | | |
|---|---|---|
| United Nations Organization or one of its officials | UN | UNO |
| United Nations specialized agency or one of its officials | UN | UNA |
| Resident of Kosovo to whom a travel document has been issued by the United Nations Interim Administration Mission in Kosovo (UNMIK) | | UNK |

**Part D — Codes for Other Issuing Authorities**

| | | |
|---|---|---|
| African Development Bank (ADB) | | XBA |
| African Export-Import Bank (AFREXIM bank) | | XIM |
| Caribbean Community or one of its emissaries (CARICOM) | | XCC |
| Council of Europe | | XCE |
| Common Market for Eastern and Southern Africa (COMESA) | | XCO |
| Economic Community of West African States (ECOWAS) | | XEC |
| International Criminal Police Organization (INTERPOL) | | XPO |
| Organization of Eastern Caribbean States (OECS) | | XES |
| Parliamentary Assembly of the Mediterranean (PAM) | | XMP |
| Sovereign Military Order of Malta or one of its emissaries | | XOM |
| Southern African Development Community | | XDC |

**Part E — Codes for Persons Without a Defined Nationality**

| | | |
|---|---|---|
| Stateless person, as defined in Article 1 of the 1954 Convention Relating to the Status of Stateless Persons | | XXA |
| Refugee, as defined in Article 1 of the 1951 Convention Relating to the Status of Refugees as amended by the 1967 Protocol | | XXB |
| Refugee, other than as defined under the code XXB above | | XXC |
| Person of unspecified nationality, for whom issuing State does not consider it necessary to specify any of the codes XXA, XXB or XXC above, whatever that person's status may be. This category may include a person who is neither stateless nor a refugee but who is of unknown nationality and legally residing in the State of issue. | | XXX |

**Part F — Codes Deprecated in [ISO 3166] (referenced for backward compatibility)**

| | | |
|---|---|---|
| Netherlands Antilles | AN | ANT |
| Neutral Zone | NT | NTZ |

**Part G — Codes Used in Specimen Documents**

In order to establish a standardized way to identify specimen documents, it is recommended to set the nationality of the document holder to “Utopia” for sample documents.

| | | |
|---|---|---|
| Utopia | UT | UTO |

**Part H — Codes Used by ICAO**

The following code, not reflected in [ISO 3166], will be utilized by ICAO only when digitally signing a master list.

| | | |
|---|---|---|
| International Civil Aviation Organization (ICAO) | IA | IAO |

# 6. TRANSLITERATIONS RECOMMENDED FOR USE BY STATES

The following tables contain the most commonly used national characters of the Latin, Cyrillic and Arabic families of languages.

### A. Transliteration of Multinational Latin-based Characters

| *Unicode* | *National character* | *Description* | *Recommended transliteration* |
|---|---|---|---|
| 00C0 | À | A grave | A |
| 00C1 | Á | A acute | A |
| 00C2 | Â | A circumflex | A |
| 00C3 | Ã | A tilde | A |
| 00C4 | Ä | A diaeresis | AE or A |
| 00C5 | Å | A ring above | AA or A |
| 00C6 | Æ | ligature AE | AE |
| 00C7 | Ç | C cedilla | C |
| 00C8 | È | E grave | E |
| 00C9 | É | E acute | E |
| 00CA | Ê | E circumflex | E |
| 00CB | Ë | E diaeresis | E |
| 00CC | Ì | I grave | I |
| 00CD | Í | I acute | I |
| 00CE | Î | I circumflex | I |
| 00CF | Ï | I diaeresis | I |
| 00D0 | Đ | Eth | D |
| 00D1 | Ñ | N tilde | N or NXX |
| 00D2 | Ò | O grave | O |
| 00D3 | Ó | O acute | O |
| 00D4 | Ô | O circumflex | O |
| 00D5 | Õ | O tilde | O |
| 00D6 | Ö | O diaeresis | OE or O |
| 00D8 | Ø | O stroke | OE |
| 00D9 | Ù | U grave | U |
| 00DA | Ú | U acute | U |
| 00DB | Û | U circumflex | U |
| 00DC | Ü | U diaeresis | UE or UXX or U |
| 00DD | Ý | Y acute | Y |
| 00DE | Þ | Thorn (Iceland) | TH |

| *Unicode* | *National character* | *Description* | *Recommended transliteration* |
|---|---|---|---|
| 0100 | Ā | A macron | A |
| 0102 | Ă | A breve | A |
| 0104 | Ą | A ogonek | A |
| 0106 | Ć | C acute | C |
| 0108 | Ĉ | C circumflex | C |
| 010A | Ċ | C dot above | C |
| 010C | Č | C caron | C |
| 010E | Ď | D caron | D |
| 0110 | Đ | D stroke | D |
| 0112 | Ē | E macron | E |
| 0114 | Ĕ | E breve | E |
| 0116 | Ė | E dot above | E |
| 0118 | Ę | E ogonek | E |
| 011A | Ě | E caron | E |
| 011C | Ĝ | G circumflex | G |
| 011E | Ğ | G breve | G |
| 0120 | Ġ | G dot above | G |
| 0122 | Ģ | G cedilla | G |
| 0124 | Ĥ | H circumflex | H |
| 0126 | Ħ | H stroke | H |
| 0128 | Ĩ | I tilde | I |
| 012A | Ī | I macron | I |
| 012C | Ĭ | I breve | I |
| 012E | Į | I ogonek | I |
| 0130 | İ | I dot above | I |
| 0131 | I | I without dot (Turkey) | I |
| 0132 | IJ | ligature IJ | IJ |
| 0134 | Ĵ | J circumflex | J |
| 0136 | Ķ | K cedilla | K |
| 0139 | Ĺ | L acute | L |
| 013B | Ļ | L cedilla | L |
| 013D | Ľ | L caron | L |
| 013F | Ŀ | L middle dot | L |
| 0141 | Ł | L stroke | L |
| 0143 | Ń | N acute | N |
| 0145 | Ņ | N cedilla | N |
| 0147 | Ň | N caron | N |

| *Unicode* | *National character* | *Description* | *Recommended transliteration* |
|---|---|---|---|
| 014A | Ŋ | Eng | N |
| 014C | Ō | O macron | O |
| 014E | Ŏ | O breve | O |
| 0150 | Ő | O double acute | O |
| 0152 | Œ | ligature OE | OE |
| 0154 | Ŕ | R acute | R |
| 0156 | Ŗ | R cedilla | R |
| 0158 | Ř | R caron | R |
| 015A | Ś | S acute | S |
| 015C | Ŝ | S circumflex | S |
| 015E | Ş | S cedilla | S |
| 0160 | Š | S caron | S |
| 0162 | Ţ | T cedilla | T |
| 0164 | Ť | T caron | T |
| 0166 | Ŧ | T stroke | T |
| 0168 | Ũ | U tilde | U |
| 016A | Ū | U macron | U |
| 016C | Ŭ | U breve | U |
| 016E | Ů | U ring above | U |
| 0170 | Ű | U double acute | U |
| 0172 | Ų | U ogonek | U |
| 0174 | Ŵ | W circumflex | W |
| 0176 | Ŷ | Y circumflex | Y |
| 0178 | Ÿ | Y diaeresis | Y |
| 0179 | Ź | Z acute | Z |
| 017B | Ż | Z dot above | Z |
| 017D | Ž | Z caron | Z |
| 1E9E | ß | double s (Germany) | SS |

### B. Transliteration of Cyrillic Characters

| *Unicode* | *National character* | *Recommended transliteration* |
|---|---|---|
| 0401 | Ё | E (except Belorussian = IO) |
| 0402 | Ћ | D |
| 0404 | Є | IE (except if Ukrainian first character, then =YE) |
| 0405 | S | DZ |
| 0406 | I | I |
| 0407 | Ї | I (except if Ukrainian first character, then =YI) |
| 0408 | J | J |
| 0409 | Љ | LJ |
| 040A | Њ | NJ |
| 040C | Ќ | K (except in the language spoken in the former Yugoslav Republic of Macedonia = KJ) |
| 040E | ў | U |
| 040F | Џ | DZ (except in the language spoken in the former Yugoslav Republic of Macedonia = DJ) |
| 0410 | А | A |
| 0411 | Б | B |
| 0412 | В | V |
| 0413 | Г | G (except Belorussian, Serbian, and Ukrainian = H) |
| 0414 | Д | D |
| 0415 | Е | E |
| 0416 | Ж | ZH (except Serbian = Z) |
| 0417 | З | Z |
| 0418 | И | I (except Ukrainian = Y) |
| 0419 | Й | I (except if Ukrainian first character, then =Y) |
| 041A | К | K |
| 041B | Л | L |
| 041C | М | M |
| 041D | Н | N |
| 041E | О | O |
| 041F | П | P |
| 0420 | Р | R |
| 0421 | С | S |
| 0422 | Т | T |
| 0423 | У | U |
| 0424 | Ф | F |

| *Unicode* | *National character* | *Recommended transliteration* |
|---|---|---|
| 0425 | Х | KH (except Serbian and in the language spoken in the former Yugoslav Republic of Macedonia = H) |
| 0426 | Ц | TS (except Serbian and in the language spoken in the former Yugoslav Republic of Macedonia = C) |
| 0427 | Ч | CH (except Serbian = C) |
| 0428 | Ш | SH (except Serbian = S) |
| 0429 | Щ | SHCH (except Bulgarian = SHT) |
| 042A | Ъ | IE |
| 042B | Ы | Y |
| 042D | Э | E |
| 042E | Ю | IU (except if Ukrainian first character, then =YU) |
| 042F | Я | IA (except if Ukrainian first character, then =YA) |
| 046A | Ѫ | U |
| 0474 | Ѵ | Y |
| 0490 | Ґ | G |
| 0492 | Ғ | G (except in the language spoken in the former Yugoslav Republic of Macedonia = GJ) |
| 04BA | Һ | C |

### C. Transliteration of Arabic Script

| *Unicode* | *Arabic letter* | *Name* | *MRZ* |
|---|---|---|---|
| 0621 | ء | hamza | X E |
| 0622 | آ | alef with madda above | X A A |
| 0623 | أ | alef with hamza above | X A E |
| 0624 | ؤ | waw with hamza above | U |
| 0625 | إ | alef with hamza below | I |
| 0626 | ئ | yeh with hamza above | X I |
| 0627 | ا | alef | A |
| 0628 | ب | beh | B |

| *Unicode* | *Arabic letter* | *Name* | *MRZ* |
|---|---|---|---|
| 0629 | ة | teh marbuta | XTA/XAH[2] |
| 062A | ت | teh | T |
| 062B | ث | theh | XTH |
| 062C | ج | jeem | J |
| 062D | ح | hah | XH |
| 062E | خ | khah | XKH |
| 062F | د | dal | D |
| 0630 | ذ | thal | XDH |
| 0631 | ر | reh | R |
| 0632 | ز | zain | Z |
| 0633 | س | seen | S |
| 0634 | ش | sheen | XSH |
| 0635 | ص | sad | XSS |
| 0636 | ض | dad | XDZ |
| 0637 | ط | tah | XTT |
| 0638 | ظ | zah | XZZ |
| 0639 | ع | ain | E |
| 063A | غ | ghain | G |
| 0640 | - | tatwheel | (Not encoded) |
| 0641 | ف | feh | F |
| 0642 | ق | qaf | Q |
| 0643 | ك | kaf | K |
| 0644 | ل | lam | L |
| 0645 | م | meem | M |
| 0646 | ن | noon | N |

2. XTA is used generally, except if *teh marbuta* occurs at the end of the name component, in which case XAH is used.

| *Unicode* | *Arabic letter* | *Name* | *MRZ* |
|---|---|---|---|
| 0647 | ه | heh | H |
| 0648 | و | waw | W |
| 0649 | ى | alef maksura | XAY |
| 064A | ي | yeh | Y |
| 064B | ً | fathatan | (Not encoded) |
| 064C | ٌ | dammatan | (Not encoded) |
| 064D | ٍ | kasratan | (Not encoded) |
| 064E | َ | fatha | (Not encoded) |
| 064F | ُ | damma | (Not encoded) |
| 0650 | ِ | kasra | (Not encoded) |
| 0651 | ّ | shadda | [DOUBLE][3] |
| 0652 | ْ | sukun | (Not encoded) |
| 0670 | ٰ | superscript alef | (Not encoded) |
| 0671 | ٱ | alef wasla | XXA |
| 0679 | ٹ | tteh | XXT |
| 067C | ټ | teh with ring | XRT |
| 067E | پ | peh | P |
| 0681 | ځ | hah with hamza above | XKE |
| 0685 | څ | hah with 3 dots above | XXH |
| 0686 | چ | tcheh | XC |
| 0688 | ڈ | ddal | XXD |
| 0689 | ډ | dal with ring | XDR |
| 0691 | ڑ | rreh | XXR |
| 0693 | ړ | reh with ring | XRR |
| 0696 | ږ | reh with dot below and dot above | XRX |

3. Shadda denotes doubling: Latin character or sequence is repeated e.g.عبّاس becomes EBBAS; فضّة becomes FXDZXDZXAH.

| *Unicode* | *Arabic letter* | *Name* | *MRZ* |
|---|---|---|---|
| 0698 | ژ | jeh | XJ |
| 069A | ښ | seen with dot below and dot above | XXS |
| 069C | ڜ | seen with 3 dots below and 3 dots above | (Not encoded) |
| 06A2 | ڢ | feh with dot moved below | (Not encoded) |
| 06A7 | ڧ | qaf with dot above | (Not encoded) |
| 06A8 | ڨ | qaf with 3 dots above | (Not encoded) |
| 06A9 | ک | keheh | XKK |
| 06AB | ڭ | kaf with ring | XXK |
| 06AD | ڭ | ng | XNG |
| 06AF | گ | gaf | XGG |
| 06BA | ں | noon ghunna | XNN |
| 06BC | ڼ | noon with ring | XXN |
| 06BE | ھ | heh doachashmee | XDO |
| 06C0 | ۀ | heh with yeh above | XYH |
| 06C1 | ہ | heh goal | XXG |
| 06C2 | ۂ | heh goal with hamza above | XGE |
| 06C3 | ۃ | teh marbuta goal | XTG |
| 06CC | ی | farsi yeh | XYA |
| 06CD | ۍ | yeh with tail | XXY |
| 06D0 | ې | yeh | Y |
| 06D2 | ے | yeh barree | XYB |
| 06D3 | ۓ | yeh barree with hamza above | XBE |

# 7. DEVIATIONS

As States worldwide continue to adopt MRTDs, the increased complexity and the rise in deviations have led to a need for reporting deviations from standards or the normal practice of a State through a standardized mechanism. Deviations are defined as MRTDs that contain elements that do not precisely conform to the ICAO specifications and the governing ISO and RFC standards. Deviations are generally observed within Country Signing Certificate Authorities (CSCA) or Document Signer Certificates (DSCs). Nonetheless, States have also indicated issues related to the LDS and MRZ fields within their MRTDs. The purpose of this section is to detail the mechanism by which issuing States can publish their deviations.

While travel documents may contain deviations, they may still be usable in border management systems. For documents that are otherwise valid, they may remain in use for several years. Consequently, relying parties should identify their own processes for handling any published deviations.

## 7.1 Operational Experiences

For a long time the only method for managing deviations was through the general advice given by issuing States via diplomatic means. This section includes deviations affecting large numbers of MRTDs that might be reported so as to assist borders in making a determination on whether travel documents are valid, forged or the product of a substitution. Some examples of operational errors include MRZ, LDS and PKI deviations.

While the MRZ has been in use for many years some recent examples of known MRZ errors are:

- MRZ date of birth does not match VIZ page date of birth.
- MRZ citizenship incorrectly reports the country of birth rather than citizenship.

In most cases travel documents with a non-conforming MRZ will be recalled by the issuing State. Since there is a gap between issuance and the subsequent reissuance, travellers may be forced to use their deviating MRTD. During this time, a published deviation may alleviate potential problems for travellers.[4]

For LDS and PKI deviations, some could go undetected for long periods of time, as many States are not yet performing Passive and Active Authentication as specified by Doc 9303. However, issuing States are strongly encouraged to publish deviations in order assist the global community in the technical adoption of MRTDs.[5]

## 7.2 Deviation List Approach

The approach described in this section aims to provide a standardized means for issuing States to publish and distribute a Travel Document Deviation List. It is based on principles established during the development of the CSCA Master List (see Doc 9303-12), in that a signed Deviation List for each State's non-conformities will be provided via the ICAO PKD or the issuing authority through a website or a LDAP-server. The PKD is used to support the dissemination of information relevant to the management of deviations.

4. Non conformities that affect single documents or small numbers of eMRTDs will not be addressed by this section, it is up to the issuing State to recall and re-issue individual documents.
5. For any instance where there has been a security issue related to a PKI certificate, the proper response is revocation as described in Doc 9303-12. Further guidance is outside the scope of this section.

Deviations are categorized into four specific areas:

- Keys and Certificates;
- Logical Data Structure (LDS);
- Machine Readable Zone (MRZ);
- Chip.

For each of these categories deviations will be described to one level only, for example:

| | |
|---|---|
| Category: | LDS |
| Error | DG2 |

Additional information will be provided via an operational parameter as made available by each State and/or a free text field in the reporting framework allowing the notifying State to add any descriptive text required. The notifying State can include links to additional information within the free text field. For certificate errors, the issuer will have the option to issue a new certificate, but this will not be mandatory.

The decision to advise relying parties of a non-conformity remains solely with the issuing State. In deciding whether to create a Deviation List, States should take into consideration that as traveller self-processing border solutions become more common, failure to communicate information relevant to non-conforming travel documents may cause delays and inconvenience for travellers, which will reflect poorly on both the issuing State and the border process as a whole.

Deviation Lists provide a means of reporting deviations affecting thousands of travel documents rather than a few or a few hundred. It is appropriate for States to manage small numbers of non-conforming travel documents directly.

## 7.3 Method

### *7.3.1 Deviation elements*

The elements that make up an MRTD range from paper to RFID chips, with each element protected in some way by security features that can be defined and thus tested by inspection systems during the life of the travel document. Security features employed on the physical travel document are both overt and covert. This section considers only deviation elements within the MRZ, LDS and PKI.

The MRZ is a fixed-dimensional area located on the MRTD data page, containing mandatory and optional data formatted for machine reading using OCR methods. Doc 9303 provides the specifications for the MRZ, including:

- purpose;
- constraints;
- transliteration; and
- data structure of the MRZ lines.

The conformity of the MRZ is routinely tested by inspection systems via data comparison with the corresponding VIZ page data and recalculation of the MRZ check digits.

The authenticity and integrity of data stored on MRTD RFID chip is protected by Passive Authentication. This security mechanism is based on digital signatures and Public Key Infrastructure (PKI).

The structure of the MRTD LDS is defined by Doc 9303-10. While there are no specific tests to establish conformity, the data stored within the LDS is in part a subset of data available from the MRZ or VIZ page of the MRTD. Consequently, the same tests apply for the digital MRZ and VIZ data as would be applied to the MRZ and VIZ page. Authenticity of the LDS is provided through the correct application of Passive Authentication by inspection systems, while Active Authentication is performed by the chip. A brief description is below:

**Passive Authentication (PA)** is based on digital signatures and consists of the following PKI components:

1. **Country Signing CA (CSCA):** Every State establishes a CSCA as its national trust point in the context of eMRTDs. The CSCA issues public key certificates for one or more (national) Document Signers. In addition each CSCA issues Certificate Revocation Lists (CRLs) of all revoked certificates.[6]

2. **Document Signers (DS):** A Document Signer digitally signs data to be stored on MRTDs; this signature is stored in the Document Security Object for each document.

**Active Authentication (AA):** Where AA is implemented, each chip contains its own AA Key Pair. The private Key is stored in the chip's secure memory with the Public Key stored at LDS Data Group 15.

### *7.3.2 Issuing Deviation Lists*

Deviation Lists MUST NOT be issued directly by a CSCA, instead the CSCA SHALL authorize a Deviation List Signer (see Doc 9303-12) to compile, sign and publish Deviation Lists. For Deviation List specifications, see Doc 9303-12.

The procedures to be performed for issuing a Deviation List SHOULD be reflected in the published certification policies of the issuing CSCA.

### *7.3.3 Receiving a Deviation List*

Every Receiving State defines its own policies under which it accepts a Deviation List and how deviations are handled during the inspection of documents. Those policies are, in general, private information.

The Receiving State will at its sole discretion choose to allow MRTDs with a deviation to be utilized.

6. Since CRLs are a security reporting mechanism and are constantly reissued, no defects reporting is necessary for them and they are therefore outside the scope of this Part.

### *7.3.4 Categories of Deviations*

#### *7.3.4.1 Keys and certificates*

Certificate and key deviations are restricted to the following:

| *Issue* | *Comment* |
|---|---|
| Certificate | Described to the Field or Extension |
| Keys | Described to the Field or Extension |
| AA | Described to the error/problem only |

*Note.— Where a reporting State decides to issue a new certificate, the certificate MUST NOT be included in the Deviation List, but could be pointed to via the free text field.*

#### *7.3.4.2 Logical Data Structure (LDS)*

LDS deviations are restricted to the following:

| *Issue* | *Comment* |
|---|---|
| EF.Com | Described to the encoding error |
| DG's | Described to the Data Group |
| EF.sod | Described to the issue (e.g. DSC) |

#### *7.3.4.3 Machine Readable Zone (MRZ)*

MRZ deviations are restricted to the following:

| *Issue* | *Comment* |
|---|---|
| Match to VIZ | Described to the field |
| Check Digits | Described to the responsible check digit |
| Wrong Information encoded | Described to the MRZ field |

### *7.3.5 Deviation type definitions*

Categories of deviations and corresponding parameters may be extended over time and will be maintained in Doc 9303.

Each deviation is described by a deviationDescription element. The deviation is identified by an Object Identifier deviationType and may be further detailed by parameters. The field description MAY contain further information, such as how the nature of the deviation cannot be adequately described by the governing deviationType.

| DeviationType | Parameters | Description |
|---|---|---|
| **Certificate/Key Deviation** | | |
| `id-Deviation-CertOrKey` | None | A generic certificate or key related deviation not covered by the more detailed deviations below. |
| `id-Deviation-CertOrKey-DSSignature` | None | The signature of the Document Signer Certificate is wrong. |
| `id-Deviation-CertOrKey-DSEncoding` `CertField` | `CertField` | The Document Signer Certificate contains a coding error. |
| `id-Deviation-CertOrKey-CSCAEncoding` | `CertField` | The Country Signing CA Certificate contains a coding error. |
| `id-Deviation-CertOrKey-AAKeyCompromised` | None | The key for Active Authentication may be compromised and should not be relied upon. |
| | | |
| **LDS Deviation** | | |
| `id-Deviation-LDS` | None | A generic LDS related deviation not covered by the more detailed deviations below. |
| `id-Deviation-LDS-DGMalformed` | `Datagroup` | The TLV encoding of the given datagroup is corrupted. |
| `id-Deviation-LDS-DGHashWrong` | `Datagroup` | The hash value of the given datagroup in the EF.SOD is wrong. |
| `id-Deviation-LDS-SODSignatureWrong` | None | The signature contained in EF.SOD is wrong. |
| `id-Deviation-LDS-COMinconsistent` | None | EF.COM and EF.SOD are inconsistent. |
| | | |

| DeviationType | Parameters | Description |
|---|---|---|
| **MRZ Deviation** | | |
| `id-Deviation-MRZ` | None | A generic MRZ related deviation not covered by the more detailed deviation below. |
| `id-Deviation-MRZ-WrongData` | MRZField | The given field of the MRZ contains wrong data (e.g. inconsistent with VIZ), but the derived BAC key is usable to open the chip.<br>If the derived BAC key is not usable, additionally id-Deviation-Chip SHALL be included in the Deviation List. |
| `id-Deviation-MRZ-WrongCheckDigit` | MRZField | The check digit to given field of the MRZ is calculated wrong. |
| | | |
| **Chip Deviation** | | |
| `id-Deviation-Chip` | None | The Chip is not usable, e.g. wrong BAC key, broken antenna or other physical defect. |
| | | |

ICAO Object Identifiers are specified in 9303-10, 9303-11, and 9303-12. A list of the Deviation Object Identifiers follows:

```
-- Deviation List Base Object identifiers
id-icao-mrtd-security-DeviationList OBJECT IDENTIFIER ::= {id-icao-mrtd-
security 7}

id-icao-mrtd-security-DeviationListSigningKey OBJECT IDENTIFIER ::= {id-icao-
mrtd-security 8}

-- Deviation Object Identifiers and Parameter Definitions
id-Deviation-CertOrKey OBJECT IDENTIFIER ::= {id-icao-DeviationList 1}

id-Deviation-CertOrKey-DSSignature OBJECT IDENTIFIER ::= {id-Deviation-
CertOrKey 1}

id-Deviation-CertOrKey-DSEncoding OBJECT IDENTIFIER ::= {id-Deviation-
CertOrKey 2}

id-Deviation-CertOrKey-CSCAEncoding OBJECT IDENTIFIER ::= {id-Deviation-
CertOrKey 3}

id-Deviation-CertOrKey-AAKeyCompromised OBJECT IDENTIFIER ::= {id-
Deviation-CertOrKey 4}

id-Deviation-LDS OBJECT IDENTIFIER ::= {id-icao-DeviationList 2}

id-Deviation-LDS-DGMalformed OBJECT IDENTIFIER ::= {id-Deviation-LDS 1}
```

```
id-Deviation-LDS-SODSignatureWrong OBJECT IDENTIFIER ::= {id-Deviation-LDS
3}

id-Deviation-LDS-COMInconsistent OBJECT IDENTIFIER ::= {id-Deviation-LDS 4}

id-Deviation-MRZ OBJECT IDENTIFIER ::= {id-icao-DeviationList 3}

id-Deviation-MRZ-WrongData OBJECT IDENTIFIER ::= {id-Deviation-MRZ 1}

id-Deviation-MRZ-WrongCheckDigit OBJECT IDENTIFIER ::= {id-Deviation-MRZ 2}

id-Deviation-Chip OBJECT IDENTIFIER ::= {id-icao-DeviationList 4}

id-Deviation-NationalUse OBJECT IDENTIFIER ::= {id-icao-DeviationList 5}
```

#### *7.3.6 Identification of deviant documents*

Documents affected by a deviation MAY be identified by several different means:

- by the Document Signer Certificate used to sign these documents; the Document Signer can be either identified by:
  - the Distinguished Name of the Issuer in combination with the Serial Number of the certificate (`issuerAndSerialNumber`),
  - the `subjectKeyIdentifier` uniquely identifying the Document Signer, or
  - the hash of the Document Signer certificate (`certificateHash`); the hash function to be used is the same as used in the signature of the Deviation List.
- by a range of issuing dates (`startIssuingDate, endIssuingDate`)
- by a list of document numbers (`listOfDocNumbers`).

Each method has advantages and disadvantages for the issuer of a Deviation List as well for the receiver of a Deviation List. These include:

- Identification by Document Signer allows recognition of a deviation by the inspection systems only after Passive Authentication was performed. Additionally, identification by Document Signer might be too coarse to accurately identify only defect documents, i.e. the deviation affects only part of the documents signed by a given Document Signer.
- The Issuing Date is not part of the machine readable zone, and also in general not available in the electronic LDS. Therefore this is not suitable for automated processing. Additionally, depending on the Issuing State, the Issuing Date might not be the actual date of passport personalization, but the application date, and therefore not accurate enough to identify only affected documents.
- A list of document numbers is difficult to compile if document numbers are not issued sequentially. A list of document numbers grows quite quickly to unmanageable size if many documents are affected by a defect.

It is RECOMMENDED to give as much identifying information on affected documents as possible. If several methods for identification are given, the conditions MUST be met simultaneously to identify a document. It is at the discretion of the Relying State to decide which means of identification given in a Deviation List entry are used to identify affected documents.

## 7.4 Publication

Deviation Lists can be published via the ICAO PKD and/or the issuing authority through a website or LDAP server. The primary distribution point for Deviation Lists is the PKD.

| | *Deviation Lists* |
|---|---|
| Primary Distribution | PKD |
| Secondary Distribution | Website/LDAP |

### *7.4.1 Publication by the issuing State*

Deviation Lists can be published via a website or an LDAP-server of the issuing authority.

### *7.4.2 Publication on the PKD*

The PKD operates as a central repository for Deviation Lists.

The procedure for publishing a Deviation List is as follows:

1. Deviation Lists are sent to the write PKD, as part of the usual certificate upload process as defined in the PKD Interface Specification and PKD Procedures Manual.

2. The ICAO PKD office validates the signatures of uploaded Deviation Lists as specified in the PKD Procedures Manual.

3. Valid Deviation Lists are moved to the read PKD.

4. The distributing State will determine if its Deviation List will be publicly available, or restricted to PKD member States.

### *7.4.3 Relying parties*

To be able to verify a Deviation List, a relying party needs to have received the corresponding CSCA certificate of the issuing State by out-of-band communications. It is up to the Relying Party to decide how to handle MRTDs with a corresponding entry in the issuing State's Deviation List.

## 8. REFERENCES (NORMATIVE)


[ISO 1073-2] ISO 1073-2:1976, Alphanumeric character sets for optical recognition – Part 2: Character set OCR-B – Shapes and dimensions of the printed image

[ISO 1831] ISO 1831:1980, Printing specifications for optical character recognition

[ISO 1664-2] ISO 11664-2:2007(E)/CIE S014-2/E: 2006, CIE Standard Illuminants for Colorimetry

[ISO 12233] ISO 12233: Photography – Electronic still picture imaging – Resolution and spatial frequency responses

[ISO 3166-1] ISO 3166-1:2013 Codes for the representation of names of countries and their subdivisions – Part 1:Country codes

[ISO 3166/MA] ISO 3166 Maintenance Agency https://www.iso.org/iso/home/standards/country_codes.htm

[ISO/IEC 7810] ISO/IEC 7810:2003, Identification cards – Physical characteristics

[ISO/IEC 39794-5] ISO/IEC 39794-5:2019, Extensible biometric data interchange formats – Part 5: Face image data

[ISO/IEC 7501] ISO/IEC 7501 multipart standard: Machine Readable Travel Documents

[ISO/IEC 10918-1] ISO/IEC 10918-1:1994, Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines

[ISO/IEC 15444-1] ISO/IEC 15444-1:2004, Information technology – JPEG 2000 image coding system: Core coding system

[ISO/IEC 15948] ISO/IEC 15948:2004, Information technology – Computer graphics and image processing – Portable Network Graphics (PNG): Functional specification

[ISO/IEC 14496-2] ISO/IEC 14496-2 Information technology – Coding of audio-visual objects Part 2: Visual [MPEG4]

[IEC 61966-2-1] IEC 61966-2-1: Multimedia systems and equipment – Colour measurement and management – Part 2-1: Colour management – Default RGB colour space – sRGB

[IEC 61966-8] IEC 61966-8:2001, Multimedia systems and equipment – Colour measurement and management – Part 8: Multimedia colour scanners

[TR-03121-3] BSI: Technical Guideline TR-03121-3: Biometrics for public sector applications, Part 3: Application Profiles and Function Modules, Volume 1: Verification scenarios for ePassport and Identity Card, Version 3.0.1. 2013

[RFC 3852] Cryptographic Message Syntax – July 2004

[RFC 5280] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, W. Polk, “Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile“, May 2008

— — — — — — — —

# APPENDIX A TO PART 3 — EXAMPLES OF CHECK DIGIT CALCULATION (INFORMATIVE)

*Example 1 — Application of check digit to date field*

Using 27 July 1952 as an example, with the date in numeric form, the calculation will be:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Date: | 5 | 2 | 0 | 7 | 2 | 7 |
| | Weighting: | 7 | 3 | 1 | 7 | 3 | 1 |
| Step 1 (multiplication) | Products: | 35 | 6 | 0 | 49 | 6 | 7 |

Step 2 (sum of products) 35 + 6 + 0 + 49 + 6 + 7 = 103

Step 3 (division by modulus) $\frac{103}{10}$ = 10, remainder 3

Step 4. Check digit is the remainder, 3. The date and its check digit shall consequently be written as `5207273`.

*Example 2 — Application of check digit to document number field*

Using the number AB2134 as an example for coding a 9-character, fixed-length field (e.g. passport number), the calculation will be:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | A | B | 2 | 1 | 3 | 4 | < | < | < |
| Assigned numeric values: | 10 | 11 | 2 | 1 | 3 | 4 | 0 | 0 | 0 |
| Weighting: | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 |
| Step 1 (multiplication) Products: | 70 | 33 | 2 | 7 | 9 | 4 | 0 | 0 | 0 |

Step 2 (sum of products) 70 + 33 + 2 + 7 + 9 + 4 + 0 + 0 + 0 = 125

Step 3 (division by modulus) $\frac{125}{10}$ = 12, remainder 5

Step 4. Check digit is the remainder, 5. The number and its check digit shall consequently be written as `AB2134<<<5`.

*Examples of the calculation of composite check digits.*

The calculation method for composite check digits is the same for all MRTDs. However, the location and number of the digits to be included in the calculation are different between the different types of documents. For completeness, examples of each are included here.

*Example 3 — Composite check digit calculation for TD3 documents*

Using the lower line of MRZ data from a TD3 data page that follows, as an example for coding the composite check digit, the calculation will be:

Character positions 1-43: Example with no alpha-numeric characters in “optional data” field.

HA672242<6YT05802254M9601086<<<<<<<<<<<<<<0

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | H | A | 6 | 7 | 2 | 2 | 4 | 2 | < | 6 |
| Assigned numeric values: | 17 | 10 | 6 | 7 | 2 | 2 | 4 | 2 | 0 | 6 |
| Weighting: | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 |
| Step 1 (multiplication) Products: | 119 | 30 | 6 | 49 | 6 | 2 | 28 | 6 | 0 | 42 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | 5 | 8 | 0 | 2 | 2 | 5 | 4 | 9 | 6 | 0 |
| Assigned numeric values: | 5 | 8 | 0 | 2 | 2 | 5 | 4 | 9 | 6 | 0 |
| Weighting: | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 |
| Step 1 (multiplication) Products: | 15 | 8 | 0 | 6 | 2 | 35 | 12 | 9 | 42 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | 1 | 0 | 8 | 6 | < | < | < | < | < | < |
| Assigned numeric values: | 1 | 0 | 8 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 |
| Step 1 (multiplication) Products: | 1 | 0 | 24 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | < | < | < | < | < | < | < | < | 0 |
| Assigned numeric values: | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 |
| Step 1 (multiplication) Products: | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Step 2 (sum of products) | 119 | + | 30 | + | 6 | + | 49 | + | 6 | + | 2 | + | 28 | + | 6 | + | 0 | + | 42 |
| Step 2 (sum of products) | 15 | + | 8 | + | 0 | + | 6 | + | 2 | + | 35 | + | 12 | + | 9 | + | 42 | + | 0 |
| Step 2 (sum of products) | 1 | + | 0 | + | 24 | + | 6 | + | 0 | + | 0 | + | 0 | + | 0 | + | 0 | + | 0 |
| Step 2 (sum of products) | 0 | + | 0 | + | 0 | + | 0 | + | 0 | + | 0 | + | 0 | + | 0 | + | 0 | | |
| Step 2 (sum of products) | 448 | | | | | | | | | | | | | | | | | | |
| Step 3 (division by modulus) | $\frac{448}{10}$ = 44, remainder 8 | | | | | | | | | | | | | | | | | | |

Step 4. Check digit is the remainder, 8. The lower line of MRZ data together with its composite check digit may consequently be written as follows:

HA672242<6YT05802254M9601086<<<<<<<<<<<<<<08

*Example 4 — Composite check digit calculation for TD1 documents*

Using the upper and middle lines of MRZ data of a TD1 that follow as an example for coding the composite check digit, the calculation will be:

Upper machine readable line (character positions 1–30): I<YTOD231458907<<<<<<<<<<<<<<<

Middle machine readable line (character positions 1–29): 3407127M9507122YTO<<<<<<<<<<<

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | D | 2 | 3 | 1 | 4 | 5 | 8 | 9 | 0 | 7 |
| Assigned numeric values: | 13 | 2 | 3 | 1 | 4 | 5 | 8 | 9 | 0 | 7 |
| Weighting: | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 |
| Step 1 (multiplication) Products: | 91 | 6 | 3 | 7 | 12 | 5 | 56 | 27 | 0 | 49 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | < | < | < | < | < | < | < | < | < | < |
| Assigned numeric values: | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 |
| Step 1 (multiplication) Products: | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| Sample data element: | < | < | < | < | < |
| Assigned numeric values: | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 1 | 7 | 3 | 1 | 7 |
| Step 1 (multiplication) Products: | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | 3 | 4 | 0 | 7 | 1 | 2 | 7 | 9 | 5 | 0 |
| Assigned numeric values: | 3 | 4 | 0 | 7 | 1 | 2 | 7 | 9 | 5 | 0 |
| Weighting: | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 |
| Step 1 (multiplication) Products: | 9 | 4 | 0 | 21 | 1 | 14 | 21 | 9 | 35 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | 7 | 1 | 2 | 2 | < | < | < | < | < | < |
| Assigned numeric values: | 7 | 1 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 |
| Step 1 (multiplication) Products: | 7 | 7 | 6 | 2 | 0 | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Sample data element: | < | < | < | < | < |
| Assigned numeric values: | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 7 | 3 | 1 | 7 | 3 |
| Step 1 (multiplication) Products: | 0 | 0 | 0 | 0 | 0 |

| | |
|---|---|
| Step 2 (sum of products) | 91 + 6 + 3 + 7 + 12 + 5 + 56 + 27 + 0 + 49 + |
| Step 2 (sum of products) | 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + |
| Step 2 (sum of products) | 0 + 0 + 0 + 0 + 0 + 9 + 4 + 0 + 21 + 1 + |
| Step 2 (sum of products) | 14 + 21 + 9 + 35 + 0 + 7 + 7 + 6 + 2 + 0 + |
| Step 2 (sum of products) | 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 |
| Step 2 (sum of products) | = 392 |
| Step 3 (division by modulus) | 392 = 39, remainder 2<br>10 |

Step 4. Check digit is the remainder, 2. The middle line of MRZ data together with its composite check digit may consequently be written as follows: `3407127M9507122YTO<<<<<<<<<<<2`.

*Example 5 — Composite check digit calculation for TD2 documents.*

Using the lower line of MRZ data that follows as an example for coding the composite check digit, the calculation will be:
Lower machine readable line (character positions 1–35):

`HA672242<6YTO5802254M9601086<<<<<<<`

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | H | A | 6 | 7 | 2 | 2 | 4 | 2 | < | 6 |
| Assigned numeric values: | 17 | 10 | 6 | 7 | 2 | 2 | 4 | 2 | 0 | 6 |
| Weighting: | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 |
| Step 1 (multiplication) Products: | 119 | 30 | 6 | 49 | 6 | 2 | 28 | 6 | 0 | 42 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | 5 | 8 | 0 | 2 | 2 | 5 | 4 | 9 | 6 | 0 |
| Assigned numeric values: | 5 | 8 | 0 | 2 | 2 | 5 | 4 | 9 | 6 | 0 |
| Weighting: | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 |
| Step 1 (multiplication) Products: | 15 | 8 | 0 | 6 | 2 | 35 | 12 | 9 | 42 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sample data element: | 1 | 0 | 8 | 6 | < | < | < | < | < | < |
| Assigned numeric values: | 1 | 0 | 8 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| Weighting: | 1 | 7 | 3 | 1 | 7 | 3 | 1 | 7 | 3 | 1 |
| Step 1 (multiplication) Products: | 1 | 0 | 24 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |

| | |
|---|---|
| Sample data element: | < |
| Assigned numeric values: | 0 |
| Weighting: | 7 |
| Step 1 (multiplication) Products: | 0 |

Step 2 (sum of products) 119 + 30 + 6 + 49 + 6 + 2 + 28 + 6 + 0 + 42 +

Step 2 (sum of products) 15 + 8 + 0 + 6 + 2 + 35 + 12 + 9 + 42 + 0 +

Step 2 (sum of products) 1 + 0 + 24 + 6 + 0 + 0 + 0 + 0 + 0 + 0 +

Step 2 (sum of products) 0

Step 2 (sum of products) = 448

Step 3 (division by modulus) $\frac{448}{10}$ = 44, remainder 8

Step 4. Check digit is the remainder, 8. The lower line of MRZ data together with its composite check digit may consequently be written as follows:

HA672242<6YT05802254M9601086<<<<<<<8.

— — — — — — — —

# APPENDIX B TO PART 3 — TRANSLITERATION OF ARABIC SCRIPT IN MRTDS (INFORMATIVE)

## B.1 The Arabic Script

The Arabic script is used by the Arabic language, the official language of about 24 countries from Morocco to Oman. The Arabic script is also used by other languages, notably Farsi in Iran; Pashto and Dari in Afghanistan; Urdu in Pakistan; and many others, including Kurdish, Assyrian, Hausa and Uighur. In the past it was used for the languages of Central Asia, for example, Tajik and Uzbek.

The Arabic script is cursive, and a letter will often change its shape depending upon whether it is standing alone (isolated); at the start of a word (initial); in the body of a word (medial); or at the end (final). For example, the letter ب (beh) changes its shape to بـ at the beginning of the word بكر (Bakr) — note that Arabic reads from right to left, so the first letter is at the right hand side. We are not concerned here with these different letter shapes (glyphs), only the basic letter code — represented by the isolated shape.

Arabic and the other languages using the Arabic script are usually written using consonants alone. Thus the name محمد (Mohammed) as written consists of just four consonants, which may be approximated in Latin as "Mhmd". The vowels are added at the discretion of the translator to achieve a phonetic equivalent. Arabic can also be "vocalized" if the vowel marks ("harakat") are added to modify the pronunciation. However, the harakat are normally omitted.

The standard Arabic script consists of 32 consonants, 18 vowels and dipthongs and three other signs. In addition there are over 100 national characters in the Arabic script when used with non-Arabic languages, although some of these are obsolete and no longer in use.

## B.2 The Arabic Script in the MRTD

### *B.2.1 VIZ*

The VIZ has a mandatory field for the name (refer to specifications for each form factor in Doc 9303, Parts 4 through 7). Doc 9303-3, paragraph 3.1, states:

> *"When mandatory data elements are in a language that does not use the Latin alphabet, a transliteration shall also be provided."*

Thus if the name is written in the Arabic script, a Latin representation shall be included. While Doc 9303 refers to this representation as a "transliteration", it is commonly a phonetic equivalent and should be more correctly termed a "transcription".

For example:

the name[1] in Arabic script: **ابو بكر محمد بن زكريا الرازي**

and a transcription into Latin characters: **Abū Bakr Mohammed ibn Zakarīa al-Rāzi**

Firstly note that Doc 9303-3, paragraph 3.2, allows the use of diacritical marks (e.g. the **ā** in **al-Rāzi**) in the VIZ at the option of the issuing State.

Secondly, note that this particular transcription into Latin characters is only one of many possibilities. For example, the following variations for **محمد** are applied variants:

| | | |
|---|---|---|
| 1. Muhammad | 2. Moohammad | 3. Moohamad |
| 4. Mohammad | 5. Mohamad | 6. Muhamad |
| 7. Muhamad | 8. Mohamed | 9. Mohammed |
| 10. Mohemmed | 11. Mohemmed | 12. Muhemmed |
| 13. Muhamed | 14. Muhammed | 15. Moohammed |
| 16. Mouhammed | | |

In some countries it is common to replace the final “d” with “t”, so this leads to a total of 32 variations for **محمد**.

The transcription scheme used depends upon the language and regional accent of the Arabic script source (non-Arabic languages such as Farsi, Pashto and Urdu also use the Arabic script); the language of the Latin script speaker; and the transcription scheme used.

### *B.2.2 MRZ*

Section 4 of this part of Doc 9303 describes the MRZ.

The MRZ provides a set of essential data elements in a format standardized for each type of MRTD that can be used by all receiving States regardless of their national script or customs. The data in the MRZ are formatted in such a way as to be readable by machines with standard capability worldwide and, as a consequence, the MRZ is a different representation of the data than is found in the VIZ. National characters generally appear only in the computer-processing systems of the States in which they apply and are not available globally. They shall not, therefore, appear in the MRZ.

The Name Field of the MRZ consists, in the case of the MRP, of 39 character positions, and only the OCR-B subset of A-Z and < may be used. Thus Arabic characters shall not be used in the MRZ, and “equivalent” OCR-B characters must be used to represent them.

The conversion of the name in the Arabic script to the Latin characters of the MRZ, constrained by the use of only the OCR-B characters (A-Z and <), is problematical. In addition, the uncertainty introduced if a phonetic-based transcription is allowed means that database searches can become useless.

For example, from the same example used above:

the name in Arabic script: **ابو بكر محمد بن زكريا الرازي**

1. Abū Bakr al-Rāzi was a great Persian scientist and doctor of about 1 100 years ago. In Persian (Farsi), his name is usually spelt with a final Persian “yeh” (ى), but to avoid confusion we have used the standard Arabic “yeh” (ي).

and one **transcription** into Latin characters for the MRZ:

ABU<BAKR<MOHAMMED<IBN<ZAKARIA<AL<RAZI

However the MRZ is likely to be one of at least 32 variants based on the name “Mohammed” alone. “Zakaria” may be written “Zakariya”; “ibn” as “bin”; and “al” as “el”. Just these variations lead to 256 alternatives.

To draw the contrast, a **transliteration** of the above name محمد, for example, applying the Buckwalter table (see below) to the four Arabic characters, would be “mHmd”. In this case, each Arabic character maps into a single Latin character. No allowance is made for phonetics.

The complete Buckwalter transliteration of the name above is:

**Abw<bAkr<mHmd<bn<zkryAY<AlrAzY**

Unfortunately, the Buckwalter table uses lower-case (a-z) and special characters (‘,|,>,$,<,},*,_,~) so is not suitable for use in the MRZ (see http://www.qamus.org/transliteration.htm).

## B.3 Recommendation for the VIZ

### *B.3.1 Transcription in the VIZ*

As stated above, Doc 9303-3, paragraph 3.1, mandates the inclusion of a “transliteration” in the VIZ when a national script other than Latin is used. Related Doc 9303-3, paragraph 3.4, refers specifically to the requirement for names.

There is confusion about the terms “transliteration” and “transcription”. A “transliteration” is a strictly one-to-one representation of the non-Latin script. A “transcription” is a more loose representation, often based on phonetics (how the name “sounds” when spoken). Of course, often sounds made in one language do not have equivalents in another, and it depends on the target language, for example, “ch”, “sh” and “th” are pronounced differently in English and French and German. Compare the English transcription “Omar Khayyam” with the German transcription “Omar Chajjam” for the name of the mathematician and poet عمر خيام.

There are many “transcription” schemes:

- Deutches Institut für Normung: DIN 31635 (1982)
- Deutsche Morgenländische Gesellschaft (1936)
- International Standards Organisation: ISO/R 233 (1961), ISO 233 (1984)[3], ISO 233-2 (1993)
- British Standards Institute: BS 4280 (1968)
- United Nations Group of Experts on Geographical Names (UNGEGN): UN (1972) [4]
- Qalam (1985)
- American Library Association – Library of Congress: ALA-LC (1997) [1]
- The Encyclopedia of Islam, new edition: EI (1960) [2]

Some countries maintain their citizens’ names in birth or citizen registers in both Arabic and Latin script, where the Latin version is an approved transcription of the Arabic version. These countries may wish to continue to enter the approved Latin transcription in the VIZ.

**Recommendation**

**Doc 9303-3, in paragraphs 3.1 and 3.4 as stated above, makes it mandatory to provide a Latin character equivalent in the VIZ, so it is at the discretion of the issuing State as to whether this is a phonetic transcription, or a copy of the MRZ transliteration (as described below).**

### *B.3.2 Transcription schemes*

Some of the transcription schemes are presented below:

| *Unicode* | *Arabic letter* | *Name*[2] | *DIN 31635* | *ISO 233* | *UN GEGN* | *ALA-LC* | *EI* |
|---|---|---|---|---|---|---|---|
| 0621 | ء | hamza | ' | ' | ' | ' | ' |
| 0622 | آ | alef with madda above | 'ā | 'â | ā | ā | Ā |
| 0627 | ا | alef | Ā | ' | | | |
| 0628 | ب | beh | B | b | b | b | B |
| 0629 | ة | teh marbuta | h,t | ṫ | h,t | h,t | a,at |
| 062A | ت | teh | T | t | t | t | T̲ |
| 062B | ث | theh | T̲ | ṯ | th | th | Th |
| 062C | ج | jeem | Ğ | ğ | j | j | Dj |
| 062D | ح | hah | ḥ | ḥ | ḫ | ḥ | ḥ |
| 062E | خ | khah | ḫ | h̲ | kh | kh | Kh |
| 062F | د | dal | D | d | d | d | D |
| 0630 | ذ | thal | D̲ | d̲ | dh | dh | Dh |
| 0631 | ر | reh | R | r | r | r | R |
| 0632 | ز | zain | Z | z | z | z | Z |
| 0633 | س | seen | S | s | s | s | S |
| 0634 | ش | sheen | Š | š | sh | sh | Sh |
| 0635 | ص | sad | ṣ | ṣ | ṣ | ṣ | ṣ |
| 0636 | ض | dad | ḍ | ḍ | ḍ | ḍ | ḍ |
| 0637 | ط | tah | ṭ | ṭ | ṭ | ṭ | ṭ |
| 0638 | ظ | zah | ẓ | ẓ | ẓ | ẓ | ẓ |
| 0639 | ع | ain | ' | ' | ' | ' | ' |
| 063A | غ | ghain | Ġ | ġ | gh | gh | Gh |
| 0640 | ـ | tatwheel | [graphic filler, not transcribed] | | | | |
| 0641 | ف | feh | F | f | f | f | F |
| 0642 | ق | qaf | Q | q | q | q | ḳ |
| 0643 | ك | kaf | K | k | k | k | K |
| 0644 | ل | lam | L | l | l | l | L |
| 0645 | م | meem | M | m | m | m | M |
| 0646 | ن | noon | N | n | n | n | N |
| 0647 | ه | heh | H | h | h | h | H |
| 0648 | و | waw | W | w | w | w | W |

2. The name of the character as given in Unicode and ISO/IEC 10646.

| *Unicode* | *Arabic letter* | *Name*[2] | *DIN 31635* | *ISO 233* | *UN GEGN* | *ALA-LC* | *EI* |
|---|---|---|---|---|---|---|---|
| 0649 | ى | alef maksura | Ā | ỳ | y | y | Ā |
| 064A | ي | yeh | Y | y | y | y | Y |
| 064B | ً | fathatan | An | á' | a | an | |
| 064C | ٌ | dammatan | Un | ú | u | un | |
| 064D | ٍ | kasratan | In | í | i | in | |
| 064E | َ | fatha | A | a | a | a | A |
| 064F | ُ | damma | u | u | u | u | U |
| 0650 | ِ | kasra | i | i | i | i | I |
| 0651 | ّ | shadda | [double] | ¯ | [double] | [double] | [double] |
| 0652 | ْ | sukun | | º | | | |
| 0670 | ٰ | superscript alef | ā | ā | ā | ā | Ā |

Other national characters are:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 067E | پ | peh | p | | | p | P |
| 0686 | چ | tcheh | č | | | ch,zh | Č |
| 0698 | ژ | jeh | ž | | | zh | Zh |
| 06A2[4] | ڢ | feh with dot moved below | f | f | | q | |
| 06A4 | ڤ | veh | v | | | v | |
| 06A5 | ڥ | feh with 3 dots below | v | | | v | |
| 06A7[4] | ڧ | qaf with dot above | q | q | | f | |
| 06A8[3] | ڨ | qaf with 3 dots above | v | | | v | |
| 06AD | ڭ | ng | G | | | g | G |
| 06AF | گ | gaf | G | | | g | G |

## B.4 Transliteration in the MRZ

### *B.4.1 Transliteration of European languages in the MRZ*

It is worth considering the situation of the national characters of European languages. Doc 9303-3, Section 6 “Transliterations Recommended for use by States” includes a table: *Transliteration of Multinational Latin-based Characters*.

Most of the national characters have their diacritical marks omitted for inclusion in the MRZ. There are a group of nine characters that are treated specially, for example, the character “Ñ” can be transliterated into the MRZ as “NXX”, thus preserving its uniqueness and importance for database searches.

3. Obsolete characters

For example:

the name in a European national script: **Térèsa CAÑON**

and the transliteration into the MRZ: CANXXON<<TERESA

While the MRZ representation appears unaesthetic (and may lead to complaints), the purpose is for machine reading, thus enabling the original name to be recovered for database searches and the like. Thus the MRZ results in the name being recognized as **CAÑON** as distinct from **CANON.**

### *B.4.2 Use of UNICODE*

Internally, computers use encoding schemes to represent the characters of different languages. A common encoding scheme is UNICODE, which is nearly equivalent to the ISO/IEC standard 10646 (UNICODE character indices are used in the tables below).

Representations of all the characters of the Arabic script can be found in UNICODE. The UNICODE character indices are usually given as a four-digit hexadecimal number (hexadecimal is base 16, and uses the numerals 0-9 and letters A-F to represent the 16 possible numbers). All Arabic characters are located in row 06 which forms the first two digits of the numbers (i.e. 06XX).

*For example:*

**ابو بكر محمد بن زكريا الرازي**

can be encoded in UNICODE as:

| | |
|---|---|
| **ابو** | Alef (ا) - Beh (ب) - Waw (و) => 0627 + 0628 + 0648 |
| **بكر** | Beh (ب) – Kaf (ك) - Reh (ر) => 0628 + 0643 + 0631 |
| **محمد** | Meem (م) – Hah (ح) – Meem (م) – Dal (د) => 0645 + 062D + 0645 + 062F |
| **بن** | Beh (ب) – Noon (ن) => 0628 + 0646 |
| **زكريا** | Zain (ز) – Kaf (ك) – Reh (ر) – Yeh (ي) – Alef (ا) => 0632 + 0643 + 0631 + 064A + 0627 |
| **الرازي** | Alef (ا) – Lam (ل) – Reh (ر) – Alef (ا) – Zain (ز) - Yeh (ي) => 0627 + 0644 + 0631 + 0627 + 0632 + 064A |

## B.5 Recommendation for the MRZ

### *B.5.1 Factors affecting transliteration in the MRZ*

Doc 9303-3, paragraph 4.1 states, "... the MRZ provides verification of the information in the VIZ and may be used to provide search characters for a database inquiry." Paragraph 4.1 also states that "The data in the MRZ are formatted in such a way as to be readable by machines with standard capability worldwide", and "The MRZ is a different representation of the data than is found in the VIZ." However, in paragraph 4.2 it is stated that "the data in the MRZ must be visually readable as well as machine readable."

The aim here is to transliterate the Arabic name into equivalent Latin characters in the MRZ such that there is only one possible representation for the name. This is necessary to avoid ambiguity and make database and alert list searching as accurate as possible for reliable identification. At the same time, the MRZ must be as far as possible a recognizable representation of the name as displayed in the VIZ so that it is visually readable for the purposes of advanced passenger processing and similar uses.

### *B.5.2 Existing transliteration schemes*

There are several transliteration schemes in use: Standard Arabic Technical Transliteration System (SATTS), Buckwalter and ASMO 449. These are presented below:

| *Unicode* | *Arabic letter* | *Name* | *SATTS* | *Buckwalter* | *ASMO 449* |
|---|---|---|---|---|---|
| 0621 | ء | hamza | E | ' | A |
| 0622 | آ | alef with madda above | (missing) | \| | B |
| 0623 | أ | alef with hamza above | (missing) | > | C |
| 0624 | ؤ | waw with hamza above | (missing) | & | D |
| 0625 | إ | alef with hamza below | (missing) | < | E |
| 0626 | ئ | yeh with hamza above | (missing) | } | F |
| 0627 | ا | alef | A | A | G |
| 0628 | ب | beh | B | b | H |
| 0629 | ة | teh marbuta | ? | p | I |
| 062A | ت | teh | T | t | J |
| 062B | ث | theh | C | v | K |
| 062C | ج | jeem | J | j | L |
| 062D | ح | hah | H | H | M |
| 062E | خ | khah | O | x | N |
| 062F | د | dal | D | d | O |
| 0630 | ذ | thal | Z | * | P |
| 0631 | ر | reh | R | r | Q |
| 0632 | ز | zain | ; | z | R |
| 0633 | س | seen | S | s | S |
| 0634 | ش | sheen | : | $ | T |
| 0635 | ص | sad | X | S | U |
| 0636 | ض | dad | V | D | V |
| 0637 | ط | tah | U | T | W |
| 0638 | ظ | zah | Y | Z | X |
| 0639 | ع | ain | " | E | Y |
| 063A | غ | ghain | G | g | Z |
| 0640 | ـ | tatwheel | (missing) | _ | 0x60 |
| 0641 | ف | feh | F | f | A |
| 0642 | ق | qaf | Q | q | B |
| 0643 | ك | kaf | K | k | C |
| 0644 | ل | lam | L | l | D |
| 0645 | م | meem | M | m | E |
| 0646 | ن | noon | N | n | F |
| 0647 | ه | heh | ? | h | G |
| 0648 | و | waw | W | w | H |
| 0649 | ى | alef maksura | (missing) | Y | I |
| 064A | ي | yeh | I | y | J |

| Unicode | Arabic letter | Name | SATTS | Buckwalter | ASMO 449 |
|---|---|---|---|---|---|
| 064B | ً | fathatan | (missing) | F | K |
| 064C | ٌ | dammatan | (missing) | N | L |
| 064D | ٍ | kasratan | (missing) | K | M |
| 064E | َ | fatha | (missing) | a | N |
| 064F | ُ | damma | (missing) | u | O |
| 0650 | ِ | kasra | (missing) | i | P |
| 0651 | ّ | shadda | (missing) | ~ | Q |
| 0652 | ْ | sukun | (missing) | o | R |
| 0670 | ٰ | superscript alef | (missing) | ` | (missing) |

As can be seen from inspection of the tables, these schemes use Latin characters outside of the range A-Z, so are fundamentally unsuitable for use in the MRZ.

The ASMO 449 scheme has an arbitrary allocation of Latin characters, whereas Buckwalter approximates some of the phonetic equivalents.

SATTS does not distinguish between heh (ه) and teh marbuta (ة), or between final yeh (ي) and alif maksura (ى), and it cannot transliterate an alif madda ( آ ).

### *B.5.3 Other considerations*

The recommended transliteration scheme cannot be put forward without considering the environment in which the MRTD operates. In particular, the name in the MRZ should be as close as possible in appearance and form as the name derived from other sources. The Passenger Name Record (PNR) used by airlines and forwarded to immigration authorities in Advanced Passenger Information (API) schemes is one example. While the transliteration in the MRZ will almost always not be exactly the same as the transcription in the VIZ (and other phonetic derivatives such as the PNR), the scheme recommended here attempts to make the names in the two zones recognizably similar.

For this purpose the character ‘X’ is used as an “escape” character in the same sense as in the Transliteration of Multinational Latin-based Characters table, except only one ‘X’ is used, and it is used before the character it modifies rather than after (e.g. “XTH” versus “NXX”). One or two characters follow each ‘X’ to represent one Arabic letter. This use of ‘X’ is possible as ‘X’ does not exist in the existing transcription and transliteration schemes for Arabic.

[The difference in the usage of ‘X’ in Arabic and Latin-based transliteration is unlikely to cause confusion. For the proper application of reverse transliteration, the original script must be defined, preferably based on the country of issue.]

In some transliteration entries, a second ‘X’ is used after the initial ‘X’: for example, alef with madda above آ is “XAA”, alef wasla ٱ is “XXA”. This technique is used primarily to avoid introducing other characters which would make the MRZ less readable by humans.

The intention is that human operators viewing the raw MRZ data from existing systems will be instructed to ignore any ‘X’ characters. The resulting name should resemble that from other sources. The raw MRZ data will also be lacking vowels that would normally be included in the VIZ transcription and in other sources such as the PNR. However if human operators are instructed that the vowels are missing then the MRZ data should be regarded as a fair representation of the transcribed phonetic version.

The transliteration will also not encompass the assimilation (sandhi) of the article before the “sun letters” as this is essentially a phonetic feature, and hence the spelling may not match the phonetic transcription of the VIZ (for example, “AL-RAZI” may be “AR-RAZI” in the VIZ).

The “shadda” (symbol to denote doubling of letters) results in the denoted character being repeated in the MRZ (doubled). Search algorithms should take into account that the “shadda” may not always be present.

### *B.5.4 Recommended transliteration scheme for Standard Arabic*

Using the Buckwalter transliteration table as a base, and taking into account the common phonetic equivalents listed in the transcription schemes (paragraph B.3.2), a recommended transliteration scheme that uses only the Latin characters A-Z can be formulated. As there is a precedent of using ‘X’ for variations (paragraph B.5.3), the character ‘X’ is used as an “escape” character to denote that the one or two characters that follow the ‘X’ represent a single Arabic letter.

| *Unicode* | *Arabic letter* | *Name* | *MRZ* | *Comments* |
|---|---|---|---|---|
| 0621 | ء | hamza | XE | |
| 0622 | آ | alef with madda above | XAA | B.5.5.1 |
| 0623 | أ | alef with hamza above | XAE | B.5.5.2 |
| 0624 | ؤ | waw with hamza above | U | B.5.5.3 |
| 0625 | إ | alef with hamza below | I | B.5.5.4 |
| 0626 | ئ | yeh with hamza above | XI | B.5.5.5 |
| 0627 | ا | alef | A | |
| 0628 | ب | beh | B | |
| 0629 | ة | teh marbuta | XTA/XAH | B.5.5.6 |
| 062A | ت | teh | T | |
| 062B | ث | theh | XTH | |
| 062C | ج | jeem | J | |
| 062D | ح | hah | XH | B.5.5.7 |
| 062E | خ | khah | XKH | |
| 062F | د | dal | D | |
| 0630 | ذ | thal | XDH | |
| 0631 | ر | reh | R | |
| 0632 | ز | zain | Z | |
| 0633 | س | seen | S | |
| 0634 | ش | sheen | XSH | |
| 0635 | ص | sad | XSS | |
| 0636 | ض | dad | XDZ | |
| 0637 | ط | tah | XTT | |
| 0638 | ظ | zah | XZZ | |
| 0639 | ع | ain | E | |
| 063A | غ | ghain | G | |
| 0640 | ـ | tatwheel | (note 1) | B.5.5.8 |
| 0641 | ف | feh | F | |
| 0642 | ق | qaf | Q | |
| 0643 | ك | kaf | K | |
| 0644 | ل | lam | L | |
| 0645 | م | meem | M | |
| 0646 | ن | noon | N | |
| 0647 | ه | heh | H | B.5.5.7 |
| 0648 | و | waw | W | |
| 0649 | ى | alef maksura | XAY | B.5.5.9 |
| 064A | ي | yeh | Y | |

| *Unicode* | *Arabic letter* | *Name* | *MRZ* | *Comments* |
|---|---|---|---|---|
| 064B | ً | fathatan | (note 1) | B.5.5.10 |
| 064C | ٌ | dammatan | (note 1) | B.5.5.10 |
| 064D | ٍ | kasratan | (note 1) | B.5.5.10 |
| 064E | َ | fatha | (note 1) | B.5.5.10 |
| 064F | ُ | damma | (note 1) | B.5.5.10 |
| 0650 | ِ | kasra | (note 1) | B.5.5.10 |
| 0651 | ّ | shadda | (doubling) | B.5.5.11 |
| 0652 | ْ | sukun | (note 1) | B.5.5.12 |
| 0670 | ٰ | superscript alef | (note 1) | B.5.5.13 |
| 0671 | ٱ | alef wasla | XXA | B.5.5.14 |

The following two letters are commonly used for foreign names:

| | | | | |
|---|---|---|---|---|
| 06A4 | ڤ | veh | V | |
| 06A5 | ڥ | feh with 3 dots below | XF | |

*Note 1.— Not encoded.*

### *B.5.5 Comments on Transliteration Table*

#### *B.5.5.1 Alef with madda above*

*Alef with madda above* ( آ ) is not represented in the ALA-LC Romanisation Tables [1]. However, both Interpol [5] and Dr Hoogland [6] recommend the transliteration XAA.

#### *B.5.5.2 Alef with hamza above*

*Alef with hamza above* ( أ ) is not represented in the ALA-LC Romanisation Tables [1]. However, Interpol [5] recommends the transliteration XAE.

#### *B.5.5.3 Waw with hamza above*

*Waw with hamza above* ( ؤ ) is not represented in the ALA-LC Romanisation Tables [1]. U is used here as *waw with hamza above* is commonly transcribed by “U”.

#### *B.5.5.4 Alef with hamza below*

*Alef with hamza below* ( إ ) is not represented in the ALA-LC Romanisation Tables [1]. The transliteration used here is I as that Latin letter is otherwise unused, and *alef with hamza below* often commences names such as إبراهيم **(Ibrahim)** where the *alef with hamza below* is commonly transcribed by “I”.

#### *B.5.5.5 Yeh with hamza above*

*Yeh with hamza above* ( ئ ) is not represented in the ALA-LC Romanisation Tables [1]. The transliteration used here is XI as *yeh with hamza above* is used in names such as فائز **(Faiz)** where the *yeh with hamza above* is commonly transcribed by “I”.

*B.5.5.6 Teh marbuta*

*Teh marbuta* ( ة ) is represented in the ALA-LC Romanisation Tables [1] as H or T or TAN, depending upon the context. Dr Hoogland [6] recommends XTA. The transliteration here of *teh marbuta* has two alternatives: XTA is used generally except if *teh marbuta* occurs at the end of the name component, in which case XAH is used. This is because feminine names often use *teh marbuta* to modify a masculine name, e.g. فاطمة **(Fatimah)**. Search algorithms should take these two possibilities into account.

*B.5.5.7 Hah and heh*

The transliterations for *hah* ( ح ) and *heh* (ه ) have been swapped at the advice of Interpol [5]. *Hah* is now XH and *heh* is H.

*B.5.5.8 Tatwheel*

*Tatwheel* ( ـ ) is a graphic character and not transliterated.

*B.5.5.9 Alef maksura*

*Alef maksura* (ى ) is now transliterated as XAY at the recommendation of Dr Hoogland [6]. Other characters are transliterated as XY_, thus the former XY is incompatible.

*B.5.5.10 Short vowels fatha, damma, kasra, fathatan, dammatan and kasratan*

The optional short vowels (haracat) are not generally used in names and are not transliterated.

*B.5.5.11 Shadda*

*Shadda* (ّ ) denotes a doubling of the consonant below it, so this is transliterated by doubling the appropriate character. Search algorithms should note that *shaddah* is optional and sometimes a doubling of the character will be present and sometimes not.

Note the special case of الله (Allah).

*B.5.5.12 Sukun*

*Sukun* (ْ ) denotes the absence of a vowel, is optional, and is not transliterated.

*B.5.5.13 Superscript alef*

*Superscript alef* ( ٰ ) (“vowel-dagger-alef”) is not transliterated.

*B.5.5.14 Alef wasla*

*Alef wasla* (ٱ ) is now transliterated as XXA at the recommendation of Interpol [5]. Other characters are transliterated XA_, thus the former XA is incompatible. Dr Hoogland [6] also recommends XXA.

### *B.5.6 Recommended transliteration scheme for other languages*

Persian is spoken in Iran (Farsi), Afghanistan (Dari), Tajikistan and Uzbekistan.
Pashto is spoken in Afghanistan and western Pakistan.
Urdu is spoken in Pakistan and India.

| *Unicode* | *Arabic letter* | *Language* | *Name* | *MRZ* |
|---|---|---|---|---|
| 0679 | ٹ | Urdu | tteh | XXT |
| 067E | پ | Persian, Urdu | peh | P |
| 067C | ټ | Pashto | teh with ring | XRT |
| 0681 | ځ | Pashto | hah with hamza above | XKE |
| 0685 | څ | Pashto | hah with 3 dots above | XXH |
| 0686 | چ | Persian, Urdu | tcheh | XC |
| 0688 | ڈ | Urdu | ddal | XXD |
| 0689 | ډ | Pashto | dal with ring | XDR |
| 0691 | ڑ | Urdu | rreh | XXR |
| 0693 | ړ | Pashto | reh with ring | XRR |
| 0696 | ږ | Pashto | reh with dot below and dot above | XRX |
| 0698 | ژ | Persian, Urdu | jeh | XJ |
| 069A | ښ | Pashto | seen with dot below and dot above | XXS |
| 06A9 | ک | Persian, Urdu | keheh | XKK |
| 06AB | ګ | Pashto | kaf with ring | XXK |
| 06AD | ڭ | | ng | XNG |
| 06AF | گ | Persian, Urdu | gaf | XGG |
| 06BA | ں | Urdu | noon ghunna | XNN |
| 06BC | ڼ | Pashto | noon with ring | XXN |
| 06BE | ھ | Urdu | heh doachashmee | XDO |
| 06C0 | ہ | Urdu | heh with yeh above | XYH |
| 06C1 | ہ | Urdu | heh goal | XXG |
| 06C2 | ۂ | Urdu | heh goal with hamza above | XGE |
| 06C3 | ۃ | Urdu | teh marbuta goal | XTG |
| 06CC | ی | Persian, Urdu | farsi yeh | XYA[4] |
| 06CD | ۍ | Pashto | yeh with tail | XXY |
| 06D0 | ې | Pashto | yeh | Y[5] |
| 06D2 | ے | Urdu | yeh barree | XYB |
| 06D3 | ۓ | Urdu | yeh barree with hamza above | XBE |

4. The letter “farsi yeh” (ی) is functionally identical to the standard “yeh” (ي) but in the isolated and final forms is graphically identical to the standard “alef maksura” (ى), so could be transliterated as ‘Y’ or “XAY”. Database matching algorithms should take this into account.
5. The character “Pashto yeh” (ې) is functionally identical to the standard “yeh” (ي).

### *B.5.7 Example of transliteration for Standard Arabic*

The example above,

**ابو بكر محمد بن زكريا الرازي**

can be encoded in the MRZ as:

| | |
|---|---|
| **ابو** | Alef (ا) - Beh (ب) - Waw (و) => ABW |
| **بكر** | Beh (ب) – Kaf (ك) - Reh (ر) => BKR |
| **محمد** | Meem (م) – Hah (ح) – Meem (م) – Dal (د) => MXHMD |
| **بن** | Beh (ب) – Noon (ن) => BN |
| **زكريا** | Zain (ز) – Kaf (ك) – Reh (ر) – Yeh (ي) – Alef (ا) => ZKRYA |
| **الرازي** | Alef (ا) – Lam (ل) – Reh (ر) – Alef (ا) – Zain (ز) - Yeh (ي) => ALRAZY |

i.e. ABW<BKR<MXHMD<BN<ZKRYA<ALRAZY

The advantages of this transliteration are:

1. The name in the Arabic script is always transliterated to the same Latin representation. This means that database matches are more likely to result;

2. The process is reversible — the name in the Arabic script can be recovered.

To recover the name in the Arabic script:

| | |
|---|---|
| ABW | A=Alef (ا) - B=Beh (ب) - W=Waw (و) => ابو |
| BKR | B=Beh (ب) – K=Kaf (ك) - R=Reh (ر) => بكر |
| MXHMD | M=Meem (م) – XH=Hah (ح) – M=Meem (م) – D=Dal (د) => محمد |
| BN | B=Beh (ب) – N=Noon (ن) => بن |
| ZKRYA | Z=Zain (ز) – K=Kaf (ك) – R=Reh (ر) – Y=Yeh (ي) – A=Alef (ا) => زكريا |
| ALRAZY | A=Alef (ا) – L=Lam (ل) – R=Reh (ر) – A=Alef (ا) – Z=Zain (ز) - Y=Yeh (ي) => الرازي |

The rationale for omitting the harakat and other diacritical marks is that they are optional and mostly not used. Therefore they should be treated the same way as the diacritical marks on European national characters (e.g. é, è, ç ) which are used for pronunciation purposes.

As well, the optional inclusion of the harakat would be detrimental for accurate database matches.

### *B.5.8 Recommended transliteration scheme for Moroccan, Tunisian and Maghrib Arabic*

Moroccan, Tunisian and Maghrib Arabic add four letters to the standard Arabic script:

| *Unicode* | *Arabic letter* | *Name* | *MRZ* |
|---|---|---|---|
| 069C | ﭺ | seen with 3 dots below and 3 dots above (Moroccan) | (note 1) |
| 06A2 | ﭒ | feh with dot moved below (Maghrib) | (note 1) |
| 06A7 | ﭪ | qaf with dot above (Maghrib) | (note 1) |
| 06A8 | ﭫ | qaf with 3 dots above (Tunisian) | (note 1) |

*Note 1.— These characters are obsolete and not transliterated (at the recommendation of Dr Hoogland [6]).*

### *B.5.9 Further examples*

Arabic: هاري الشماع
VIZ: **Hari Al-Schamma**
MRZ: HARY<ALXSHMAE<<<<<<<<<<<<<<<<<<<<<<<<

Arabic: سمير بادمكدوذيل
VIZ: **Samir Badmakduthal**
MRZ: SMYR<BADMKDWXDHYL<<<<<<<<<<<<<<<<<<<<<

Arabic: جمال عبد الناصر
VIZ: **Gamal Abdel Nasser**
MRZ: JMAL<EBD<ALNAXSSR<<<<<<<<<<<<<<<<<<<<<<

Arabic: العباس عبد الله بن محمد السفاح
VIZ: **al-'Abbās 'Abdu'llāh ibn Muhammad as-Saffāh**
MRZ: ALEBAS<EBD<ALLXH<BN<MXHMD<ALSFAXH<<<<<<

Arabic: عبدالله محمد بن عمر بن الحسين فخر الدين الرازي
VIZ: **Abdullah Muhammad ibn Umar ibn al-Husayn Fakhr al-Din al-Razi**
MRZ[6]: EBD<ALLXH<MXHMD<BN<EMR<BN<ALXHSYN<FXKHR

Arabic: عبدالعزيز بن متعب
VIZ: **Abdul Aziz bin Mithab**
MRZ: EBD<ALEZYZ<BN<MTEB<<<<<<<<<<<<<<<<<<<<

Arabic: إسماعيل عزّ الدين
VIZ: **Isma'il Izz-ud-din**
MRZ: ISMAEYL<EZZ<ALDYN<<<<<<<<<<<<<<<<<<<<<

Arabic: جميلة نعيمة
VIZ: **Jamillah Na'ima**
MRZ: JMYLXAH<NEYMXAH<<<<<<<<<<<<<<<<<<<<<<

6. Truncated as specified in the form factor specific Parts 4 to 7 of Doc 9303.

### *B.5.10 Order of names in the MRZ*

Doc 9303-3, paragraphs 4.6 and Parts 4-7, specify how primary and secondary identifiers shall be printed. This Appendix does not attempt to define primary and secondary identifiers in Arabic names. It is for the issuing authority to make that determination. But as an example:

the name in Arabic script: **ابو بكر محمد بن زكريا الرازي**

1) if the component BN<ZKRYA<ALRAZY is considered the primary identifier, then the MRZ is:

`BN<ZKRYA<ALRAZY<<ABW<BKR<MXHMD<<<<<<<<<`

2) if the component ALRAZY is considered the primary identifier, then the MRZ is:

`ALRAZY<<ABW<BKR<MXHMD<BN<ZKRYA<<<<<<<<<`

## B.6 Reverse Transliteration of the MRZ

### *B.6.1 Table for Reverse Transliteration of the MRZ*

Using the table hereunder, the Latin characters in the MRZ can be mapped back into the original Arabic script. Note that 'X' is an "escape" character and the following one or two Latin characters must be used to obtain the corresponding Arabic letter.

| *MRZ* | *Name of Arabic letter* | *Arabic letter* | *Unicode* |
|---|---|---|---|
| A | alef | ا | 0627 |
| B | beh | ب | 0628 |
| D | dal | د | 062F |
| E | ain | ع | 0639 |
| F | feh | ف | 0641 |
| G | ghain | غ | 063A |
| H | heh | ه | 0647 |
| I | alef with hamza below | إ | 0625 |
| J | jeem | ج | 062C |
| K | kaf | ك | 0643 |
| L | lam | ل | 0644 |
| M | meem | م | 0645 |
| N | noon | ن | 0646 |
| P | peh (Persian, Urdu) | پ | 067E |
| Q | qaf | ق | 0642 |
| R | reh | ر | 0631 |
| S | seen | س | 0633 |
| T | teh | ت | 062A |
| U | waw with hamza above | ؤ | 0624 |
| V | veh | ڤ | 06A4 |
| W | waw | و | 0648 |
| Y | yeh or yeh (Pashto) | ي / ې | 064A/06D0 |
| Z | zain | ز | 0632 |

| MRZ | Name of Arabic letter | Arabic letter | Unicode |
|---|---|---|---|
| XAA | alef with madda above | آ | 0622 |
| XAE | alef with hamza above | أ | 0623 |
| XAH | teh marbuta (see also xta) | ة | 0629 |
| XAY | alef maksura | ى | 0649 |
| XBE | yeh barree with hamza above | ۓ | 06D3 |
| XC | tcheh (Persian, Urdu) | چ | 0686 |
| XDH | thal | ذ | 0630 |
| XDO | heh doachashmee | ھ | 06BE |
| XDR | dal with ring (Pashto) | ډ | 0689 |
| XDZ | dad | ض | 0636 |
| XE | hamza | ء | 0621 |
| XF | feh with 3 dots below | ڟ | 06A5 |
| XGG | gaf (Persian, Urdu) | گ | 06AF |
| XGE | heh goal with hamza above (Urdu) | ۂ | 06C2 |
| XH | hah | ح | 062D |
| XI | yeh with hamza above | ئ | 0626 |
| XJ | jeh (Urdu) | ژ | 0698 |
| XKE | hah with hamza above (Pashto) | ځ | 0681 |
| XKH | khah | خ | 062E |
| XKK | keheh (Persian, Urdu) | ک | 06A9 |
| XNN | noon ghunna (Urdu) | ں | 06BA |
| XNG | ng | ڭ | 06AD |
| XRR | reh with ring (Pashto) | ړ | 0693 |
| XRT | teh with ring | ټ | 067C |
| XRX | reh with dot below and dot above (Pashto) | ږ | 0696 |
| XSH | sheen | ش | 0634 |
| XSS | sad | ص | 0635 |
| XTA | teh marbuta (see also XAH) | ة | 0629 |
| XTG | teh marbuta goal (Urdu) | ۃ | 06C3 |
| XTH | theh | ث | 062B |
| XTT | tah | ط | 0637 |
| XXA | alef wasla | ٱ | 0671 |
| XXD | ddal (Urdu) | ڈ | 0688 |
| XXG | heh goal (Urdu) | ہ | 06C1 |
| XXH | hah with 3 dots above (Pashto) | څ | 0685 |
| XXK | kaf with ring (Pashto) | ګ | 06AB |
| XXN | noon with ring (Pashto) | ڼ | 06BC |
| XXR | rreh (Urdu) | ڑ | 0691 |
| XXS | seen with dot below and dot above (Pashto) | ښ | 069A |
| XXT | tteh (Urdu) | ٹ | 0679 |
| XXY | yeh with tail (Pashto) | ې | 06CD |
| XYA | farsi yeh (Persian, Urdu) | ی | 06CC |
| XYB | yeh barree (Urdu) | ے | 06D2 |
| XYH | heh with yeh above (Urdu) | ۀ | 06C0 |
| XZZ | zah | ظ | 0638 |

## B.7 Computer Programs

### *B.7.1 Arabic to MRZ*

This program written in Python is offered as an example of converting Arabic characters (in Unicode) to the MRZ format.

The Arabic characters are contained in a file “Arabic source.txt” and the corresponding MRZ data is written to a file “MRZ output.txt”.

```
*******************************************************************************

#    # -*- coding: iso-8859-15 -*-

import unicodedata
import encodings.utf_8_sig
import codecs

# TRANSLITERATE
def Arabic_to_MRZ(unicode_string):
   transform = {0x20: '<', 0x21: 'XE', 0x22: 'XAA', 0x23: 'XAE', 0x24: 'U',
               0x25: 'I', 0x26: 'XI', 0x27: 'A', 0x28: 'B', 0x29: 'XAH',
               0x2A: 'T', 0x2B: 'XTH', 0x2C: 'J', 0x2D: 'XH', 0x2E: 'XKH',
               0x2F: 'D', 0x30: 'XDH', 0x31: 'R', 0x32: 'Z', 0x33: 'S', 0x34: 'XSH',
               0x35: 'XSS', 0x36: 'XDZ', 0x37: 'XTT', 0x38: 'XZZ', 0x39: 'E',
               0x3A: 'G', 0x41: 'F', 0x42: 'Q', 0x43: 'K', 0x44: 'L',
               0x45: 'M', 0x46: 'N', 0x47: 'H', 0x48: 'W', 0x49: 'XAY',
               0x4A: 'Y', 0x71: 'XXA', 0x79: 'XXT', 0x7E: 'P', 0x7C: 'XRT',
               0x81: 'XKE', 0x85: 'XXH', 0x86: 'XC', 0x88: 'XXD', 0x89: 'XDR',
               0x91: 'XXR', 0x93: 'XRR', 0x96: 'XRX', 0x98: 'XJ', 0x9A: 'XXS',
               0xA4: 'XV', 0xA5: 'XF', 0xA9: 'XKK', 0xAB: 'XXK', 0xAD: 'XNG',
               0xAF: 'XGG', 0xBA: 'XNN', 0xBC: 'XXN', 0xBE: 'XDO', 0xC0: 'XYH',
               0xC1: 'XXG', 0xC2: 'XGE', 0xC3: 'XTG',
               0xCC: 'XYA', 0xCD: 'XXY', 0xD0: 'Y', 0xD2: 'XYB', 0xD3: 'XBE'}
   name_in = unicode_string
   name_out = ""
   for c in name_in:
# check for shadda (double)
      if ord(c) == 0x51:
         name_out = name_out + char
      else:
         if ord(c) in transform:
            char = transform[ord(c)]
            name_out = name_out + char
   print name_out
   return name_out

#
#   MAIN - Arabic to MRZ
#

# open input and output files
```

```
fin = encodings.utf_8_sig.codecs.open('Arabic source.txt', 'r') #b', 'utf-8-sig', 'ignore', 1)
fout = open('MRZ output.txt', 'w')

# loop through the input file

try:
    for arabic_name in fin:
        MRZ_name = Arabic_to_MRZ(arabic_name)
        fout.write(MRZ_name)
        fout.write('\n')
finally:
    fin.close()
fout.flush()
fout.close()
```

**********

### *B.7.2 MRZ to Arabic*

This program written in Python is offered as an example of converting MRZ characters to Arabic characters (in Unicode).

The MRZ characters are contained in a file “MRZ source.txt” and the corresponding Arabic data is written to a file “Arabic output.txt”.

**********

```
#    # -*- coding: iso-8859-15 -*-

import unicodedata
import encodings.utf_8_sig
import codecs

# TRANSLITERATE
def MRZ_to_Arabic(ascii_string):
    transform = { '<': 0x20, 'XE': 0x21, 'XAA':0x22, 'XAE': 0x23, 'U': 0x24,
                  'I': 0x25, 'XI': 0x26, 'A': 0x27, 'B': 0x28, 'XAH': 0x29,
                  'T': 0x2A, 'XTH': 0x2B, 'J': 0x2C, 'XH': 0x2D, 'XKH': 0x2E,
                  'D': 0x2F, 'XDH': 0x30, 'R': 0x31, 'Z': 0x32, 'S': 0x33, 'XSH': 0x34,
                  'XSS': 0x35, 'XDZ': 0x36, 'XTT': 0x37, 'XZZ': 0x38, 'E': 0x39,
                  'G': 0x3A, 'F': 0x41, 'Q': 0x42, 'K': 0x43, 'L': 0x44, 'M': 0x45,
                  'N': 0x46, 'H': 0x47, 'W': 0x48, 'XAY': 0x49, 'Y': 0x4A, 'XXA': 0x71,
                  'XXT': 0x79, 'P': 0x7E, 'XRT': 0x7C, 'XKE': 0x81, 'XXH': 0x85,
                  'XC': 0x86, 'XXD': 0x88, 'XDR': 0x89, 'XXR': 0x91, 'XRR': 0x93,
                  'XRX': 0x96, 'XJ': 0x98, 'XXS': 0x9A, 'XV': 0xA4, 'XF': 0xA5,
                  'XKK': 0xA9, 'XXK': 0xAB, 'XNG': 0xAD, 'XGG': 0xAF,
                  'XNN': 0xBA, 'XXN': 0xBC, 'XDO': 0xBE, 'XYH': 0xC0,
                  'XXG': 0xC1, 'XGE': 0xC2, ‘XTA’: 0x29, 'XTG': 0xC3, 'XYA': 0xCC,
                  'XXY': 0xCD, 'I': 0xD0, 'XYB': 0xD2, 'XBE': 0xD3}
    name_in = ascii_string
```

```
    name_out = ""
    # if this character is not X, does it appear by itself in the table?
    search_string = ''
    last_string = ''
    iloop = 0
    while iloop < len(name_in):
        search_string = search_string + name_in[iloop]
        if search_string in transform:
            if search_string <> last_string:
                name_out = name_out + chr((transform[search_string]))
            #insert shadda if double found
            else:
                name_out = name_out + chr(0x51)
            if search_string <> '<':
                name_out = name_out + chr(0x06)
            else:
                name_out = name_out + chr(0x00)
            #remember last string
            if search_string <> '<':
                last_string = search_string
            else:
                last_string = ''
            #clear the search string once found
            search_string = ''
        iloop = iloop + 1
    print name_out
    return name_out

#
#   MAIN - MRZ to Arabic
#

# open input and output files

fin = open('MRZ source.txt', 'r')
fout = open('Arabic output.txt', 'wb') #b', 'utf-8-sig', 'strict', 1)
fout.write(encodings.utf_8_sig.codecs.BOM)


# loop through the input file

try:
    for MRZ_name in fin:
        Arabic_name = MRZ_to_Arabic(MRZ_name)
        Arabic_name = Arabic_name + chr(0x0D) + chr(0x00) + chr(0x0A) + chr(0x00)
        fout.write(Arabic_name)
finally:
    fin.close()
fout.flush()
fout.close()

*****************************************************************************
```

### B.8 References (Informative)

[1] *ALA-LC Romanization Tables: Transliteration Schemes for Non-Roman Scripts.* Randal K. Berry (ed.). Library of Congress, 1997.

[2] *The Encyclopedia of Islam*. New Edition. Leiden, 1960.

[3] *ISO 233:1984. Documentation - Transliteration of Arabic characters into Latin characters*. International Organization for Standardization, 1984-12-15.

[4] *United Nations Romanization Systems for Geographical Names. Report on Their Current Status*. Compiled by the UNGEGN Working Group on Romanization Systems. Version 2.1. June 2002.

[5] *IPSG comments to the document: Transliteration of Arabic Fonts in Machine Readable Travel Documents - Technical Report - Version 2.3 dated 15 Feb 2008*. Interpol, Lyon, 17 March 2008.

[6] Private correspondence, Dr. Jan Hoogland, Department of Arabic, University of Nijmegen, the Netherlands, 23 March 2008.

[7] *Comments on the Translation of Arabic Fonts in Machine Readable Travel Documents TECHNICAL REPORT AMA 13052008*, Mr. Abdalla M. Askar, Emirates Identity Authority.

**— END —**

ISBN 978-92-9275-314-6