



ICAO

Doc 9303

Machine Readable Travel Documents

Seventh Edition, 2015

Part 3: Specifications Common to all MRTDs



Approved by the Secretary General and published under his authority

INTERNATIONAL CIVIL AVIATION ORGANIZATION



| ICAO

Doc 9303

Machine Readable Travel Documents

Seventh Edition, 2015

Part 3: Specifications Common to all MRTDs

Approved by the Secretary General and published under his authority

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/mrtd

Doc 9303, *Machine Readable Travel Documents*
Part 3 — *Specifications Common to all MRTDs*
ISBN 978-92-9249-792-7

© ICAO 2015

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.

TABLE OF CONTENTS

	<i>Page</i>
1. SCOPE	1
2. PHYSICAL CHARACTERISTICS OF MRTDS	2
3. VISUAL INSPECTION ZONE (VIZ)	2
3.1 Languages and Characters	2
3.2 Typeface and Type Size	3
3.3 Captions/Fields	4
3.4 Convention for Writing the Name of the Holder	4
3.5 Representation of Issuing State or Organization	5
3.6 Representation of Nationality	5
3.7 Representation of Place of Birth	5
3.8 Representation of Dates	6
3.9 Displayed Identification Features of the Holder	8
4. MACHINE READABLE ZONE (MRZ)	16
4.1 Purpose of the MRZ	16
4.2 Properties of the MRZ	16
4.3 Constraints of the MRZ	17
4.4 Print Specifications	17
4.5 Machine Reading Requirements and the Effective Reading Zone	18
4.6 Convention for Writing the Name of the Holder	18
4.7 Representation of Issuing State or Organization and Nationality of Holder	20
4.8 Representation of Dates	20
4.9 Check Digits in the MRZ	20
4.10 Characteristics of the MRZ	21
4.11 Quality Specifications of the MRZ	21
5. CODES FOR NATIONALITY, PLACE OF BIRTH, LOCATION OF ISSUING STATE/AUTHORITY AND OTHER PURPOSES	22
6. TRANSLITERATIONS RECOMMENDED FOR USE BY STATES	30
7. DEVIATIONS	38
7.1 Operational Experiences	38
7.2 Deviation List Approach	38
7.3 Method	39
7.4 Publication	44
7.5 Technical Specifications	45

	<i>Page</i>
8. REFERENCES (NORMATIVE).....	49
APPENDIX A TO PART 3 EXAMPLES OF CHECK DIGIT CALCULATION (INFORMATIVE)	App A-1
APPENDIX B TO PART 3 TRANSLITERATION OF ARABIC SCRIPT IN MRTDS (INFORMATIVE)	App B-1
B.1 The Arabic Script	App B-1
B.2 The Arabic Script in the MRTD	App B-1
B.3 Recommendation for the VIZ	App B-3
B.4 Transliteration in the MRZ	App B-5
B.5 Recommendation for the MRZ.....	App B-6
B.6 Reverse transliteration of the MRZ	App B-15
B.7 Computer programs	App B-17
B.8 References (Informative)	App B-20

1. SCOPE

The Seventh Edition of Doc 9303 represents a restructuring of the ICAO specifications for Machine Readable Travel Documents. Without incorporating substantial modifications to the specifications, in this new edition Doc 9303 has been reformatted into a set of specifications for Size 1 Machine Readable Official Travel Documents (TD1), Size 2 Machine Readable Official Travel Documents (TD2), and Size 3 Machine Readable Travel Documents (TD3), as well as visas. This set of specifications consists of various separate documents in which general (applicable to all MRTDs) as well as MRTD form factor specific specifications are grouped.

This Part 3 of Doc 9303 is based on the Sixth Edition of Doc 9303, Part 1, Volume 1, *Machine Readable Passports – Passports with Machine Readable Data Stored in Optical Character Recognition Format* (2006) and the Third Edition of Doc 9303, Part 3, Volume 1, *Machine Readable Official Travel Documents – MRtds with Machine Readable Data Stored in Optical Character Recognition Format* (2008).

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 should 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 $+50^{\circ}\text{C}$ (14°F to 122°F). The MRTD should not lose its functionality after being exposed to temperatures ranging from -35°C to $+80^{\circ}\text{C}$ (-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 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 Characters”, e.g. P and ß, may also be used in the VIZ without transliteration. When mandatory data elements are in a national language that does not use the Latin alphabet, a transliteration shall also be provided.

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 should 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 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 national language that does not use Latin characters, the name of the State or other location shall appear in the national 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 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 in the national language should be accompanied by one or more translations of the name. 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.

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 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 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 is a language that does not use Latin characters, the town, city or suburb shall appear in the national 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 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 it shall be represented by its ICAO three-letter code (see 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 DDbMMbYY or DDbMMbYYYY, where b = a single blank space.

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

<i>Month</i>	<i>English</i>	<i>French</i>	<i>Spanish</i>
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 normally be shown by the last two 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 four-digit representation of the year.

Examples:

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

12bLUGb/JUILb42

where b = a single blank space, i.e. 12 LUG /JUIL 42

or

12 July 42 (using English only)

or

12JUIL42 (using French abbreviation)

or

12JUL 42 (using English or Spanish abbreviation)

or

12 07 42 (using numerical format).

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:

XXbXXbXX

XXbXXbXXXX

XXbXXXbXX where b = a single blank space.

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 the following guidelines and illustrations must be followed when taking photographs to be used as the facial image of the holder in an MRTD.

The displayed facial image shall depict a true likeness of the rightful holder of the MRTD and shall not be digitally altered or enhanced to change the subject's appearance in any way. It shall have been taken within the six months preceding the issue date of the MRTD.

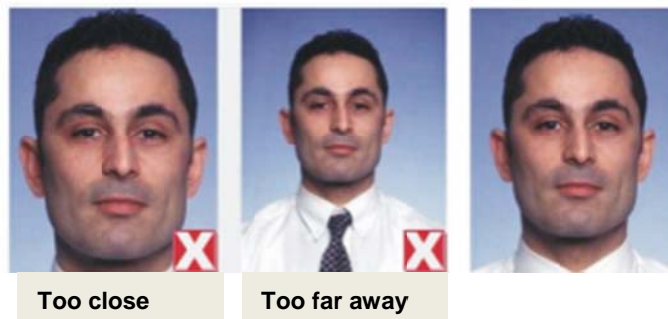
3.9.1.1 Pose

The photograph shall show a close up of the head and shoulders with the subject facing square on and looking directly at the camera with both eyes visible and with a neutral expression with the mouth closed.



The pose should be such that an imaginary horizontal line drawn between the centres of the eyes is parallel to the top and bottom edges of the rectangular image.

The facial image shall be in focus from the crown (top of the head ignoring any hair) to the chin and from the nose to the ears.



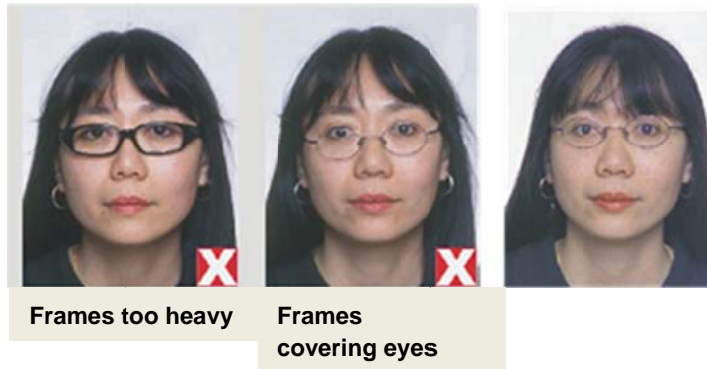
If the additional detail of one ear is required (sometimes referred to as “half-on profile”), the face shall be at such an angle to the imaginary plane as to reveal the detail of the ear while maintaining full-face frontal details on that side of the face opposite to the exposed ear.

Both edges of the face must be clearly visible. The subject shall not be looking, portrait-style, over one shoulder.

The face shall be in sharp focus and clear with no blemishes such as ink marks or creases.

The eyes must be open, and there must be no hair obscuring them.

If the subject wears glasses, the photograph must show the eyes clearly with no lights reflected in the glasses. The glasses shall not have tinted lenses. Avoid heavy frames if possible and ensure that the frames do not cover any part of the eyes. Glasses should appear only if permanently worn.



Head coverings shall not be accepted except in circumstances that the issuing State specifically approves. Such circumstances may be religious, medical or cultural. The face must be visible from the hairline to the chin and forward of the ears.



Coverings, hair, headdress or facial ornamentation which obscure the face are not permitted.



The issuing State shall use its discretion as to the extent to which facial ornaments (e.g. nose rings, studs), not obscuring the face, may appear in the portrait. A facial ornament should appear only if it is permanently worn.

A facial image of a baby should conform to the same specifications as for adults. Ideally, the baby should be in an upright position but it is acceptable to capture the facial image with the baby lying on a white or plain light-coloured blanket. Alternatively the baby may be placed in a baby seat but there shall be white or plain light-coloured background behind the head. The baby's eyes shall be open and no supporting hands visible.



There must be no other people or objects in the photograph.

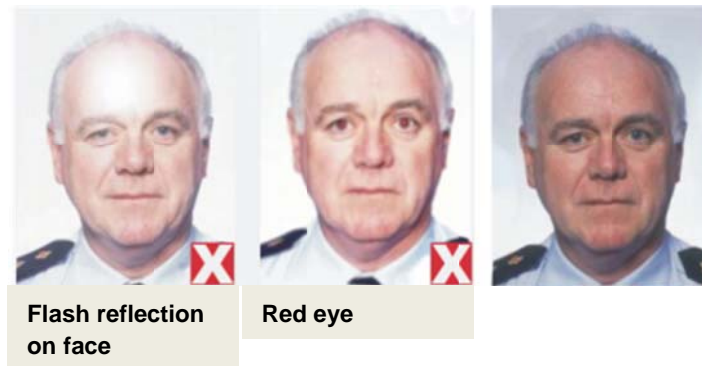


3.9.1.2 *Lighting, exposure and colour balance*

Adequate and uniform illumination shall be used to capture the facial image ensuring there are no shadows or reflections on the face or in the background.

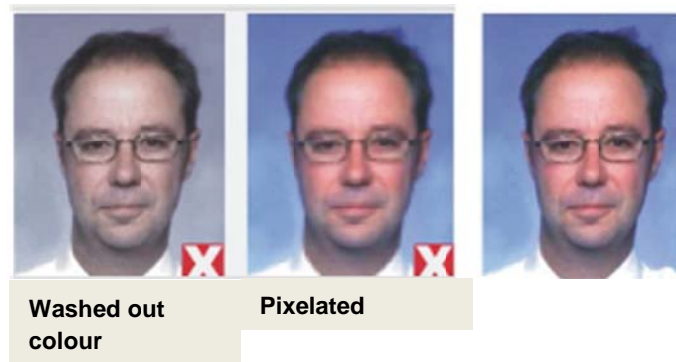


The subject's eyes must not show red eye.



The photograph must have appropriate brightness and contrast.





The displayed portrait shall be monochrome greyscale [or black and white] or a true-colour representation of the holder. Where the picture is in colour, the lighting and photographic process must be colour balanced to render skin tones faithfully.

A uniform light-coloured background shall be used to provide a contrast to the face and hair. For colour portraits, light blue, beige, light brown, pale grey or white are recommended for the background.

3.9.1.3 Placement of the portrait

The facial image 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.

3.9.1.4 Digitally printed reproduction

The portrait, whether provided in paper or digital format, must be digitally printed in the MRTD. Necessary measures shall be taken by the issuing State or organization to ensure that the displayed portrait is resistant to forgery and substitution.

Digital reproduction quality. The digital reproduction shall yield an accurate recognizable representation of the rightful holder of the document. The quality of the original captured portrait should at least be comparable to the minimum quality acceptable for photographs (resolution comparable to 6–8 line pairs per millimetre). To achieve this comparable image quality in a digital reproduction, careful attention must be given to the image capture, processing, digitization, compression and printing technology and the process used to produce the portrait, including the final preparation of the MRTD.

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

Coexistence with background security treatment(s). A digitally printed reproduction shall coexist with background security treatment(s) located within Zone V, i.e. background security printing shall not interfere with proper viewing of the displayed portrait, and vice versa.

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.

3.9.1.5 *Submission of portrait to the issuing authority*

Where the portrait is supplied to the issuing authority in the form of a print, the photograph, whether produced using conventional photographic techniques or digital techniques, should be on good or photo-quality paper and should be of the maximum specified dimensions.

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

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.

3.9.1.6 *Compliance with international standards*

The photograph shall comply with the appropriate definitions set out in [ISO/IEC 19794-5].

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.

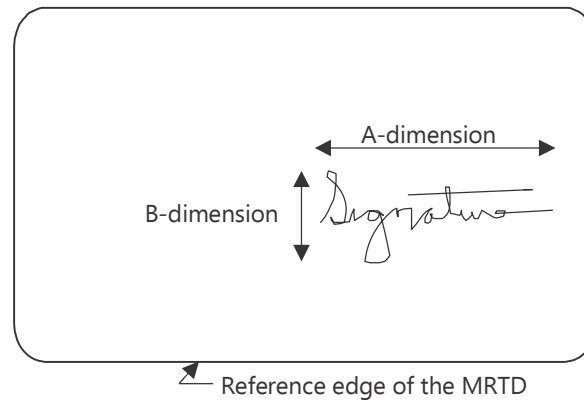


Figure 2. Orientation of the displayed signature or usual mark

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.

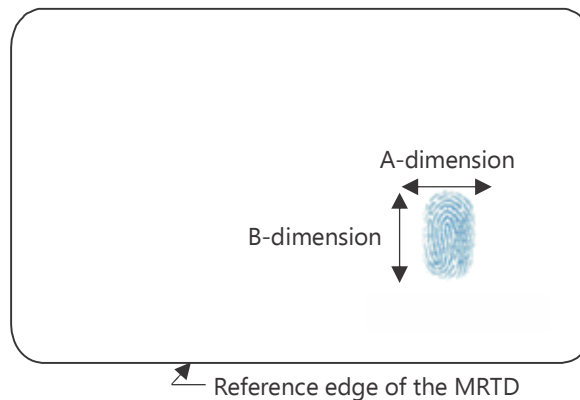


Figure 3. Orientation of the displayed single-digit fingerprint

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.

0 1 2 3 4 5 6 7 8 9
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 <

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

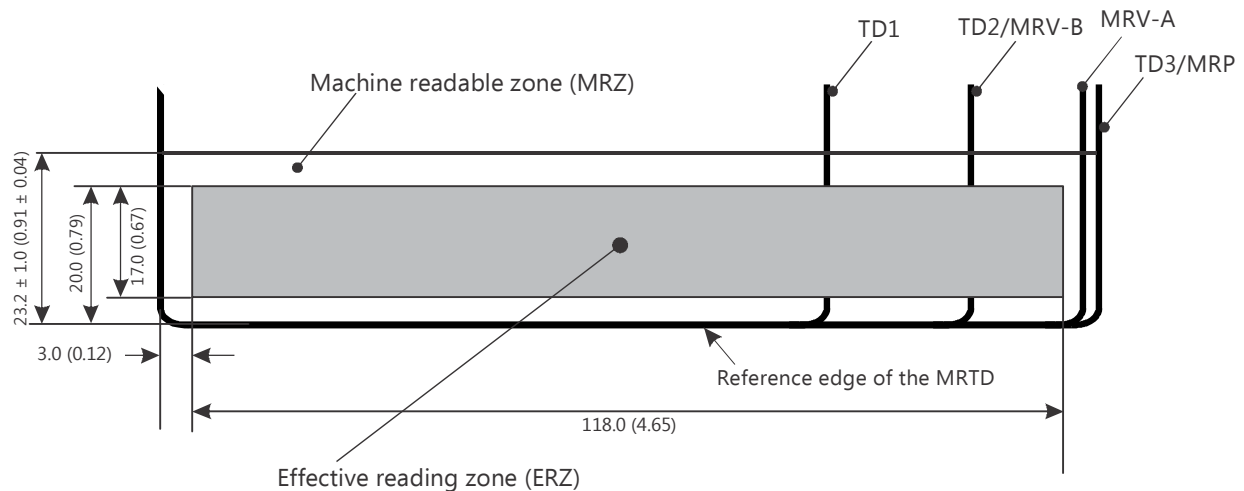
Note.— For illustrative purposes only – the characters shown are larger than actual size.

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.



Dimensions in millimetres
(inch dimensions in parentheses)

Not to scale

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<MARI A

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 listed 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 \geq 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

The following are the two- and three-letter codes for entities specified and regularly updated in [ISO 3166-1], with extensions for certain States and organizations being identified by an asterisk. The current version of the codes may be obtained from the [ISO 3166] maintenance agency - [ISO 3166/MA], ISO's focal point for country codes.

<i>Entity (short name)</i>	2-letter code	3-letter code	<i>Entity (short name)</i>	2-letter code	3-letter code
----------------------------	---------------	---------------	----------------------------	---------------	---------------

Afghanistan	AF	AFG
Åland Islands	AX	ALA
Albania	AL	ALB
Algeria	DZ	DZA
American Samoa	AS	ASM

Andorra	AD	AND
Angola	AO	AGO
Anguilla	AI	AIA
Antarctica	AQ	ATA
Antigua and Barbuda	AG	ATG

Argentina	AR	ARG
Armenia	AM	ARM
Aruba	AW	ABW
Australia	AU	AUS
Austria	AT	AUT
Azerbaijan	AZ	AZE
Bahamas	BS	BHS
Bahrain	BH	BHR
Bangladesh	BD	BGD
Barbados	BB	BRB
Belarus	BY	BLR
Belgium	BE	BEL
Belize	BZ	BLZ
Benin	BJ	BEN
Bermuda	BM	BMU
Bhutan	BT	BTN
Bolivia, Plurinational State of	BO	BOL
Bonaire, Sint Eustatius and Saba	BQ	BES
Bosnia and Herzegovina	BA	BIH
Botswana	BW	BWA
Bouvet Island	BV	BVT
Brazil	BR	BRA

British Indian Ocean Territory	IO	IOT
Brunei Darussalam	BN	BRN
Bulgaria	BG	BGR
Burkina Faso	BF	BFA
Burundi	BI	BDI
Cabo Verde	CV	CPV
Cambodia	KH	KHM
Cameroon	CM	CMR
Canada	CA	CAN
Cayman Islands	KY	CYM
Central African Republic	CF	CAF
Chad	TD	TCD
Chile	CL	CHL
China	CN	CHN
Christmas Island	CX	CXR
Cocos (Keeling) Islands	CC	CCK
Colombia	CO	COL
Comoros	KM	COM
Congo	CG	COG
Congo, Democratic Republic of the	CD	COD
Cook Islands	CK	COK
Costa Rica	CR	CRI

Côte d'Ivoire	CI	CIV
Croatia	HR	HRV
Cuba	CU	CUB
Curaçao	CW	CUW
Cyprus	CY	CYP
Czech Republic	CZ	CZE
Denmark	DK	DNK
Djibouti	DJ	DJI
Dominica	DM	DMA
Dominican Republic	DO	DOM
Ecuador	EC	ECU
Egypt	EG	EGY
El Salvador	SV	SLV
Equatorial Guinea	GQ	GNQ
Eritrea	ER	ERI
Estonia	EE	EST
Ethiopia	ET	ETH
Falkland Islands (Malvinas)	FK	FLK ¹
Faroe Islands	FO	FRO
Fiji	FJ	FJI

Finland	FI	FIN
France	FR	FRA
French Guiana	GF	GUF
French Polynesia	PF	PYF
French Southern Territories	TF	ATF
Gabon	GA	GAB
Gambia	GM	GMB
Georgia	GE	GEO
Germany	DE	D
Ghana	GH	GHA
Gibraltar	GI	GIB
Greece	GR	GRC
Greenland	GL	GRL
Grenada	GD	GRD
Guadeloupe	GP	GLP
Guam	GU	GUM
Guatemala	GT	GTM
Guernsey	GG	GGY
Guinea	GN	GIN
Guinea-Bissau	GW	GNB
Guyana	GY	GUY
Haiti	HT	HTI

¹ A dispute exists between the Governments of Argentina and the United Kingdom of Great Britain and Northern Ireland concerning sovereignty over the Falkland Islands (Malvinas).

Heard Island and McDonald Islands	HM	HMD
Holy See (Vatican City State)	VA	VAT
Honduras	HN	HND
Hong Kong Special Administrative Region of China	HK	HKG
Hungary	HU	HUN
Iceland	IS	ISL
India	IN	IND
Indonesia	ID	IDN
Iran (Islamic Republic of)	IR	IRN
Iraq	IQ	IRQ
Ireland	IE	IRL
Isle of Man	IM	IMN
Israel	IL	ISR
Italy	IT	ITA
Jamaica	JM	JAM
Japan	JP	JPN
Jersey	JE	JEY
Jordan	JO	JOR
Kazakhstan	KZ	KAZ
Kenya	KE	KEN
Kiribati	KI	KIR

Korea, Democratic People's Republic of	KP	PRK
Korea, Republic of	KR	KOR
Kuwait	KW	KWT
Kyrgyzstan	KG	KGZ
Lao People's Democratic Republic	LA	LAO
Latvia	LV	LVA
Lebanon	LB	LBN
Lesotho	LS	LSO
Liberia	LR	LBR
Libya	LY	LBY
Liechtenstein	LI	LIE
Lithuania	LT	LTU
Luxembourg	LU	LUX
Macao Special Administrative Region of China	MO	MAC
Macedonia, the former Yugoslav Republic of	MK	MKD
Madagascar	MG	MDG
Malawi	MW	MWI
Malaysia	MY	MYS
Maldives	MV	MDV
Mali	ML	MLI
Malta	MT	MLT

Marshall Islands	MH	MHL
Martinique	MQ	MTQ
Mauritania	MR	MRT
Mauritius	MU	MUS
Mayotte	YT	MYT
Mexico	MX	MEX
Micronesia (Federated States of)	FM	FSM
Moldova, Republic of	MD	MDA
Monaco	MC	MCO
Mongolia	MN	MNG
Montenegro	ME	MNE
Montserrat	MS	MSR
Morocco	MA	MAR
Mozambique	MZ	MOZ
Myanmar	MM	MMR
Namibia	NA	NAM
Nauru	NR	NRU
Nepal	NP	NPL
Netherlands	NL	NLD
Netherlands Antilles	AN	ANT
Neutral Zone	NT	NTZ
New Caledonia	NC	NCL

New Zealand	NZ	NZL
Nicaragua	NI	NIC
Niger	NE	NER
Nigeria	NG	NGA
Niue	NU	NIU
Norfolk Island	NF	NFK
Northern Mariana Islands	MP	MNP
Norway	NO	NOR
Oman	OM	OMN
Pakistan	PK	PAK
Palau	PW	PLW
Palestine, State of	PS	PSE
Panama	PA	PAN
Papua New Guinea	PG	PNG
Paraguay	PY	PRY
Peru	PE	PER
Philippines	PH	PHL
Pitcairn	PN	PCN
Poland	PL	POL
Portugal	PT	PRT
Puerto Rico	PR	PRI
Qatar	QA	QAT

Réunion	RE	REU
Romania	RO	ROU
Russian Federation	RU	RUS
Rwanda	RW	RWA
Saint Barthélemy	BL	BLM
Saint Helena, Ascension and Tristan da Cunha	SH	SHN
Saint Kitts and Nevis	KN	KNA
Saint Lucia	LC	LCA
Saint Martin (French part)	MF	MAF
Saint Pierre and Miquelon	PM	SPM
Saint Vincent and the Grenadines	VC	VCT
Samoa	WS	WSM
San Marino	SM	SMR
Sao Tome and Principe	ST	STP
Saudi Arabia	SA	SAU
Senegal	SN	SEN
Serbia	RS	SRB
Seychelles	SC	SYC
Sierra Leone	SL	SLE
Singapore	SG	SGP
Sint Maarten (Dutch part)	SX	SXM
Slovakia	SK	SVK

Slovenia	SI	SVN
Solomon Islands	SB	SLB
Somalia	SO	SOM
South Africa	ZA	ZAF
South Georgia and the South Sandwich Islands	GS	SGS
South Sudan	SS	SSD
Spain	ES	ESP
Sri Lanka	LK	LKA
Sudan	SD	SDN
Suriname	SR	SUR
Svalbard and Jan Mayen	SJ	SJM
Swaziland	SZ	SWZ
Sweden	SE	SWE
Switzerland	CH	CHE
Syrian Arab Republic	SY	SYR
Taiwan Province of China	TW	TWN
Tajikistan	TJ	TJK
Tanzania, United Republic of	TZ	TZA
Thailand	TH	THA
Timor-Leste	TL	TLS
Togo	TG	TGO
Tokelau	TK	TKL

Tonga	TO	TON
Trinidad and Tobago	TT	TTO
Tunisia	TN	TUN
Turkey	TR	TUR
Turkmenistan	TM	TKM
Turks and Caicos Islands	TC	TCA
Tuvalu	TV	TUV
Uganda	UG	UGA
Ukraine	UA	UKR
United Arab Emirates	AE	ARE
United Kingdom	GB	
— British Citizen		GBR
— British Overseas Territories Citizen		GBD*
— British National (Overseas)		GBN*
— British Overseas Citizen		GBO*
— British Protected person		GBP*

— British Subject		GBS*
United States	US	USA
United States Minor Outlying Islands	UM	UMI
Uruguay	UY	URY
Uzbekistan	UZ	UZB
Vanuatu	VU	VUT
Venezuela, Bolivarian Republic of	VE	VEN
Viet Nam	VN	VNM
Virgin Islands (British)	VG	VGB
Virgin Islands (U.S.)	VI	VIR
Wallis and Futuna	WF	WLF
Western Sahara	EH	ESH
Yemen	YE	YEM
Zambia	ZM	ZMB
Zimbabwe	ZW	ZWE

Part B — Other Codes Reserved by ISO 3166/MA

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

Part C — Codes for Use in United Nations Travel Documents

United Nations Organization or one of its officials	ZZ	UNO
United Nations specialized agency or one of its officials	ZZ	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
Common Market for Eastern and Southern Africa (COMESA)		XCO
Economic Community of West African States (ECOWAS)		XEC
International Criminal Police Organization (INTERPOL)		XPO
Sovereign Military Order of Malta or one of its emissaries		XOM

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

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

<i>Unicode</i>	<i>National character</i>	<i>Description</i>	<i>Recommended transliteration</i>
00C1	Á	A acute	A
00C0	À	A grave	A
00C2	Â	A circumflex	A
00C4	Ä	A diaeresis	AE or A
00C3	Ã	A tilde	A
0102	Ă	A breve	A
00C5	Å	A ring	AA or A
0100	Ā	A macron	A
0104	Ą	A ogonek	A
0106	Ć	C acute	C
0108	Ĉ	C circumflex	C
010C	Č	C caron	C
010A	Ċ	C dot accent	C
00C7	Ç	C cedilla	C
0110	Ð	Eth	D
010E	Ď	D caron	D
00C9	É	E acute	E
00C8	È	E grave	E
00CA	Ê	E circumflex	E
00CB	Ë	E diaeresis	E
011A	Ě	E caron	E
0116	Ė	E dot accent	E
0112	Ē	E macron	E
0118	Ę	E ogonek	E
0114	Ě	E breve	E
011C	Ĝ	G circumflex	G
011E	Ğ	G breve	G
0120	Ġ	G dot accent	G
0122	Ç	G cedilla	G
0126	Ĥ	H bar	H
0124	Ĥ	H circumflex	H

<i>Unicode</i>	<i>National character</i>	<i>Description</i>	<i>Recommended transliteration</i>
	ı	I without dot (Turkey)	I
00CD	Í	I acute	I
00CC	ì	I grave	I
00CE	î	I circumflex	I
00CF	ï	I diaeresis	I
0128	ĩ	I tilde	I
0130	í	I dot accent	I
012A	Ī	I macron	I
012E	ł	I ogonek	I
012C	ĭ	I breve	I
0134	Ĵ	J circumflex	J
0136	Ʒ	K cedilla	K
0141	Ł	L slash	L
0139	Ł	L acute	L
013D	Ľ	L caron	L
013B	ł	L cedilla	L
013F	Ł	L dot	L
0143	Ń	N acute	N
00D1	Ñ	N tilde	N or NXX
0147	Ň	N caron	N
0145	Ŋ	N cedilla	N
014B	η	Eng	N
00D8	Ø	O slash	OE
00D3	Ó	O acute	O
00D2	ò	O grave	O
00D4	ô	O circumflex	O
00D6	ö	O diaeresis	OE or O
00D5	õ	O tilde	O
0150	Ő	O double acute	O
014C	Ō	O macron	O
014E	ö	O breve	O
0154	Ŕ	R acute	R
0158	Ř	R caron	R
0156	Ŗ	R cedilla	R
015A	Ś	S acute	S
015C	Ŝ	S circumflex	S
0160	Š	S caron	S

<i>Unicode</i>	<i>National character</i>	<i>Description</i>	<i>Recommended transliteration</i>
015E	Ş	S cedilla	S
0166	Ʀ	T bar	T
0164	Ť	T caron	T
0162	Ț	T cedilla	T
00DA	Ú	U acute	U
00D9	Ù	U grave	U
00DB	Û	U circumflex	U
00DC	Ü	U diaeresis	UE or UXX or U
0168	Û	U tilde	U
016C	Ŭ	U breve	U
0170	Û	U double acute	U
016E	Ů	U ring	U
016A	Ū	U macron	U
0172	Ų	U ogonek	U
0174	Ŵ	W circumflex	W
00DD	Ÿ	Y acute	Y
0176	Ŷ	Y circumflex	Y
0178	ÿ	Y diaeresis	Y
0179	Ž	Z acute	Z
017D	Ž	Z caron	Z
017B	Ž	Z dot	Z
00FE	Þ	Thorn (Iceland)	TH
00C6	Æ	ligature AE	AE
0132	IJ	ligature IJ	IJ
0152	Œ	ligature OE	OE
00DF	ß	double s (Germany)	SS

B. Transliteration of Cyrillic Characters

<i>Unicode</i>	<i>National character</i>	<i>Recommended transliteration</i>
0410	А	A
0411	Б	B
0412	В	V
0413	Г	G (except Belorussian and Serbian = H)
0414	Д	D
0415	Е	E
0401	Ё	E (except Belorussian = IO)
0416	Ж	ZH (except Serbian = Z)
0417	З	Z
0418	И	I (except Ukrainian = Y)
0406	І	I
0419	Й	I
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
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)
0429	Ш	SH (except Serbian = S)
0429	Щ	SHCH (except Bulgarian = SHT)
042B	Ы	Y
042A	Ь	IE
042D	Э	E
042E	Ю	IU

<i>Unicode</i>	<i>National character</i>	<i>Recommended transliteration</i>
042F	Я	IA
0476	Ѹ	Y
0490	Ґ	G
045E	Ў	U
046A	Ѧ	U
	<i>f</i>	G (except in the language spoken in the former Yugoslav Republic of Macedonia = GJ)
0452	Ɔ	D
0405	Š	DZ
0408	Ј	J
041A	Ќ	K (except in the language spoken in the former Yugoslav Republic of Macedonia = KJ)
0459	Љ	LJ
045A	Њ	NJ
04BB	ћ	C
040F	Џ	DZ (except in the language spoken in the former Yugoslav Republic of Macedonia = DJ)
0454	Є	IE
0407	İ	I

C. Transliteration of Arabic Script

<i>Unicode</i>	<i>Arabic letter</i>	<i>Name</i>	<i>MRZ</i>
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 ¹
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

¹ XTA is used generally except if *teh marbuta* occurs at the end of the name component, in which case XAH is used.

<i>Unicode</i>	<i>Arabic letter</i>	<i>Name</i>	<i>MRZ</i>
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] ²
0652	◌◌◌◌◌◌◌◌	sukun	(Not encoded)
0670	◌◌◌◌◌◌◌◌◌	superscript alef	(Not encoded)
0671	آ	alef wasla	XXA
0679	ط	Tteh	XXT
067E	پ	Peh	P
067C	ٲ	teh with ring	XRT
0681	ح	hah with hamza above	XKE
0685	ح	hah with 3 dots above	XXH
0686	چ	Tchah	XC
0688	ڌ	Ddal	XXD
0689	ڊ	dal with ring	XDR
0691	ڙ	Rreh	XXR
0693	ړ	reh with ring	XRR
0696	ړ	reh with dot below and dot above	XRX

² Shadda denotes doubling: Latin character or sequence is repeated eg عباَس becomes EBBAS; فضة becomes FXDZXDZXA H.

<i>Unicode</i>	<i>Arabic letter</i>	<i>Name</i>	<i>MRZ</i>
0698	ژ	Jeh	X J
069A	ڙ	seen with dot below and dot above	XXS
069C	ڙڙڙ	seen with 3 dots below and 3 dots above	(Not encoded)
06A2	ڦ	feh with dot moved below	(Not encoded)
06A7	ڦ	qaf with dot above	(Not encoded)
06A8	ڦ	qaf with 3 dots above	(Not encoded)
06A9	ڪ	keheh	XKK
06AB	ڪ	kaf with ring	XXK
06AD	ڱ	Ng	XNG
06AF	گ	gaf	XGG
06BA	ڻ	noon ghunna	XNN
06BC	ڻ	noon with ring	XXN
06BE	ھ	heh doachashmee	XDO
06C0	ه	heh with yeh above	XYH
06C1	ه	heh goal	XXG
06C2	ه	heh goal with hamza above	XGE
06C3	ه	teh marbuta goal	XTG
06CC	ی	farsi yeh	XYA
06CD	ی	yeh with tail	XXY
06D0	ي	Yeh	Y
06D2	ے	Yeh barree	XYB
06D3	ے	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.³

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.⁴

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.

3 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.

4 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.⁵
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.

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.

⁵ 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:

<i>Issue</i>	<i>Comment</i>
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:

<i>Issue</i>	<i>Comment</i>
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:

<i>Issue</i>	<i>Comment</i>
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. 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.

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 Defect 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 Defect Lists is the PKD.

<i>Deviation Lists</i>	
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.

7.5 Technical Specifications

7.5.1 Deviation List specification

The Deviation List is implemented as a SignedData type, as specified in [RFC 3852]. All Deviation Lists MUST be produced in DER format to preserve the integrity of the signatures within them.

The range of deviations will be bounded by:

- date range (including both the issue and expiry date);
- issuer name and serial number;
- Subject Key Identifier of DSC;
- list of eMRTD numbers.

Appropriate combinations of these values will be used to accurately bind the range of MRTDs affected. When combining values, they are to be processed as joined by "AND". There is no option to process values as joined using "OR".

7.5.1.1 SignedData Type

The processing rules in [RFC 3852] apply.

- m mandatory – the field MUST be present.
 r recommended – the field SHOULD be present.
 x do not use – the field MUST NOT be populated.
 o optional – the field MAY be present.

<i>Value</i>		<i>Comments</i>
SignedData		
version	m	Value = v3
digestAlgorithms	m	
encapContentInfo	m	
eContentType	m	id-DefectList
eContent	m	The encoded contents DefectList

<i>Value</i>		<i>Comments</i>
certificates	m	States MUST include the Defect List Signer certificate and SHOULD include the CSCA certificate, which can be used to verify the signature in the signerInfos field.
crls	x	
signerInfos	m	It is RECOMMENDED that States provide only 1 signerinfo within this field.
SignerInfo	m	
version	m	The value of this field is dictated by the sid field. See [RFC 3852] Section 5.3 for rules regarding this field
sid	m	
subjectKeyIdentifier	r	It is RECOMMENDED that States support this field over issuerandSerialNumber.
digestAlgorithm	m	The algorithm identifier of the algorithm used to produce the hash value over encapsulatedContent and SignedAttrs.
signedAttrs	m	Producing States may wish to include additional attributes for inclusion in the signature, however these do not have to be processed by receiving States except to verify the signature value. signedAttrs MUST include signing time (ref. PKCS#9).
signatureAlgorithm	m	The algorithm identifier of the algorithm used to produce the signature value, and any associated parameters.
signature	m	The result of the signature generation process.
unsignedAttrs	x	

7.5.1.2 ASN.1 specification

```
DeviationList
{ iso-itu-t(2) international-organization(23) icao(136) mrttd(1) security(1)
deviationlist(7)}

DEFINITIONS IMPLICIT TAGS ::=
BEGIN

IMPORTS

-- Imports from RFC 3280 [PROFILE], Appendix A.1
AlgorithmIdentifier
  FROM PKIX1Explicit88
  { iso(1) identified-organization(3) dod(6)
    internet(1) security(5) mechanisms(5) pkix(7)
    mod(0) pkix1-explicit(18) }

-- Imports from RFC 3852
SubjectKeyIdentifier, Digest, IssuerAndSerialNumber
  FROM CryptographicMessageSyntax2004
  { iso(1) member-body(2) us(840) rsadsi(113549)
    pkcs(1) pkcs-9(9) smime(16) modules(0)
    cms-2004(24) };

DeviationListVersion ::= INTEGER {v0(0)}

DeviationList ::= SEQUENCE {
  version      DeviationListVersion,
  digestAlgorithm AlgorithmIdentifier OPTIONAL,
  deviations   SET OF Deviation
}

Deviation ::= SEQUENCE{
  documents      DeviationDocuments,
  descriptions   SET OF DeviationDescription
}

DeviationDescription ::= SEQUENCE{
  description    PrintableString OPTIONAL,
  deviationType  OBJECT IDENTIFIER,
  parameters     [0] ANY DEFINED BY deviationType OPTIONAL,
  nationalUse    [1] ANY OPTIONAL

-- The nationalUse field is for internal State use, and is not governed
-- by an ICAO specification.
}
```

```

DeviationDocuments ::= SEQUENCE {
  documentType    [0] PrintableString (SIZE(2)) OPTIONAL,
                  -- per MRZ, e.g. 'P'
  dscIdentifier   DocumentSignerIdentifier OPTIONAL,
  issuingDate     [4] IssuancePeriod OPTIONAL,
  documentNumbers [5] SET OF PrintableString OPTIONAL
}

DocumentSignerIdentifier ::= CHOICE{
  issuerAndSerialNumber [1] IssuerAndSerialNumber,
  subjectKeyIdentifier [2] SubjectKeyIdentifier,
  certificateDigest [3] Digest -- if used, digestAlgorithm must be present in
DeviationList
}

IssuancePeriod ::= SEQUENCE {
  firstIssued   GeneralizedTime,
  lastIssued    GeneralizedTime
}

-- CertField is used to define which part of a certificate is
-- affected by a coding error. Parts of the Body are identified by
-- the corresponding value of CertificateBodyField, extensions
-- by the corresponding OID identifying the extension.

CertField ::= CHOICE {
  body CertificateBodyField,
  extension OBJECT IDENTIFIER
}

CertificateBodyField ::= INTEGER {
  generic(0), version(1), serialNumber(2), signature(3), issuer(4),
  validity(5), subject(6), subjectPublicKeyInfo(7),
  issuerUniqueID(8), subjectUniqueID(9)
}

Datagroup ::= INTEGER
  {dg1(1), dg2(2), dg3(3), dg4(4), dg5(5), dg6(6),
   dg7(7), dg8(8), dg9(9), dg10(10), dg11(11),
   dg12(12), dg13(13), dg14(14), dg15(15), dg16(16),
   sod(20), com(21)}

MRZField ::= INTEGER
  {generic(0), documentCode(1), issuingState(2), personName(3),
   documentNumber(4), nationality(5), dateOfBirth(6),
   sex(7), dateOfExpiry(8), optionalData(9)}

```



```
-- Base Object Identifiers

id-icao OBJECT IDENTIFIER ::= { 2 23 136 }
id-icao-mrtd OBJECT IDENTIFIER ::= {id-icao 1}
id-icao-mrtd-security OBJECT IDENTIFIER ::= {id-icao-mrtd 1}
id-icao-DeviationList OBJECT IDENTIFIER ::= {id-icao-mrtd-security 7}
id-icao-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-DGHashWrong OBJECT IDENTIFIER ::= {id-Deviation-LDS 2}
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}

END
```

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 3166-1] ISO 3166-1:2013 Codes for the representation of names of countries and their subdivisions – Part 1:Country codes
- [ISO/IEC 7810] ISO/IEC 7810:2003, Identification cards – Physical characteristics
- [ISO/IEC 19794-5] ISO/IEC 19794-5:2005, Information technology — Biometric data interchange formats — Part 5: Face image data
- [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)		$\frac{103}{10} = 10, \text{ remainder } 3$						

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)		$\frac{125}{10} = 12, \text{ remainder } 5$							

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.

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:	0	0	0	0	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	+	
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	+	
Step 2 (sum of products)	= 392																				
Step 3 (division by modulus)	$\frac{392}{10} = 39, \text{ remainder } 2$																				

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: 3407127M9507122YT0<<<<<<<<<<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<6YT05802254M9601086<<<<<<<<

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 ب 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 diphthongs 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 national 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 name¹ 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. The “Database of Arabic Name Variants” website² gives the following sixteen variations for محمد:

- | | | |
|---------------|--------------|---------------|
| 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

Chapter 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” (ی), but to avoid confusion we have used the standard Arabic “yeh” (ي).

² See <<http://www.kanji.org/cjk/arabic/araborth.htm>>.

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:

- Deutsches 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 ³	DIN 31635	ISO 233	UN GEGN	ALA-LC	EI	
0621	ء	hamza	'	'	'	'	'	
0622	أ	alef with madda above	'ā	'ā	ā	ā	Ā	
0627	ا	alef	Ā	'				
0628	ب	beh	B	b	b	b	B	
0629	ة	teh marbuta	h,t	ṯ	h,t	h,t	a,at	
062A	ت	teh	T	t	t	t	T	
062B	ث	theh	Ṭ	ṯ	th	th	Th	
062C	ج	jeem	Ĝ	ğ	j	j	Dj	
062D	ح	hah	Ḥ	ḥ	ḥ	ḥ	ḥ	
062E	خ	khah	Ḫ	ḫ	kh	kh	Kh	
062F	د	dal	D	d	d	d	D	
0630	ذ	thal	Ḍ	ḏ	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	ḏ	ḏ	ḏ	ḏ	
0637	ط	tah	Ṭ	ṯ	ṯ	ṯ	ṯ	
0638	ظ	zah	Z	ẓ	Z	Z	Z	
0639	ع	ain	'	'	'	'	'	
063A	غ	ghain	Ĝ	ğ	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	q̣	
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	

³ The name of the character as given in Unicode and ISO/IEC 10646.

Unicode	Arabic letter	Name ³	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	ln	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		◌			
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	<u>Zh</u>
06A2 ⁴	ف	feh with dot moved below	f	f		q	
06A4	ف	veh	v			v	
06A5	ف	feh with 3 dots below	v			v	
06A7 ⁴	ق	qaf with dot above	q	q		f	
06A8 ⁴	ق	qaf with 3 dots above	v			v	
06AD	گ	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.

⁴ 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	-	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 (ي), 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 آ 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	أ	alef with madda above	XAA	B.5.5.1
0623	آ	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	أ	alef wasla	XXA	B.5.5.14

The following two letters are commonly used for foreign names:

06A4	ف	Veh	V	
06A5	ف	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 (اِ) 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 (اَ) 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 (وَ) 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 (اِ) 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 (يَ) 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 فايز (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 (-) is a graphic character and not transliterated.

B.5.5.9 *Alef maksura*

Alef maksura (ء) 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 (ِ) 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	ت	Pashto	teh with ring	XRT
0681	ح	Pashto	hah with hamza above	XKE
0685	ح	Pashto	hah with 3 dots above	XXH
0686	چ	Persian, Urdu	Tcheh	XC
0688	ڈ	Urdu	Ddal	XXD
0689	د	Pashto	dal with ring	XDR
0691	ڑ	Urdu	Rreh	XXR
0693	ر	Pashto	reh with ring	XRR
0696	ړ	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	ک	Pashto	kaf with ring	XXK
06AD	ځ		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	ۀ	Urdu	heh with yeh above	XYH
06C1	ه	Urdu	heh goal	XXG
06C2	ۀ	Urdu	heh goal with hamza above	XGE
06C3	ۀ	Urdu	teh marbuta goal	XTG
06CC	ی	Persian, Urdu	farsi yeh	XYA ⁵
06CD	ی	Pashto	yeh with tail	XXY
06D0	ی	Pashto	Yeh	Y ⁶
06D2	ے	Urdu	Yeh barree	XYB
06D3	ے	Urdu	yeh barree with hamza above	XBE

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

⁶ The character “Pashto yeh” (ی) is functionally identical to the standard “yeh” (ی).

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.

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	ء	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	پ	06A5
XGG	Gaf (Persian, Urdu)	گ	06AF
XGE	Heh goal with hamza above (Urdu)	ه	06C2
XH	Hah	ح	062D
XI	Yeh with hamza above	ي	0626
XJ	Jeh (Urdu)	ژ	0698
XKE	Hah with hamza above (Pashto)	خ	0681
XKH	Khah	ځ	062E
XKK	Keheh (Persian, Urdu)	ک	06A9
XNN	Noon ghunna (Urdu)	ن	06BA
XNG	Ng	ښ	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)	ه	06C3
XTH	Theh	ث	062B
XTT	Tah	ط	0637
XXA	Alef wasla	أ	0671
XXD	Ddal (Urdu)	ڈ	0688
XXG	Heh goal (Urdu)	ه	06C1
XXH	Hah with 3 dots above (Pashto)	څ	0685
XXK	Kaf with ring (Pashto)	ک	06AB
XXN	Noon with ring (Pashto)	ڼ	06BC
XXR	Rreh (Urdu)	ړ	0691
XXS	Seen with dot below and dot above (Pashto)	ښ	069A
XXT	Tteh (Urdu)	ټ	0679
XXY	Yeh with tail (Pashto)	ی	06CD
XYA	Farsi yeh (Persian, Urdu)	ی	06CC
XYB	Yeh barree (Urdu)	ء	06D2
XYH	Heh with yeh above (Urdu)	ه	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, 'XDX': 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 978-92-9249-792-7



9

789292

497927