INTERNATIONAL CIVIL AVIATION ORGANIZATION

CIVIL AVIATION AUTHORITY
الهيئة العامة للطيران المدني

ICAO Regional Seminar on MRTDs, Biometrics and Border Security

دودندوة الايكاو الإقليمية عن وثائق السفر المقروءة آليا، والأدوات البيومترية وأمن الحدود

الدوحة، قطر

31 October - 2 November 2011, Doha, Qatar

٣١ أكتوبر إلى ٢ نوفمبر ٢٠١١

Arabic Transliteration Rules in Document 9303

Mike ELLIS

ISO, Australia

V3.4.1 10/2011

Main points:

1. Identification: name in ARABIC script is only reliable basis

2. Countries that use the Arabic script should have the benefits of the MRZ

The Arabic name must appear in Latin characters in VIZ

Passport / جواز سفر   Type   Country Code   Passport No.
P   ASU   A2222222

Name   Mahmut Abdul Rahiim   محمـود عبـدالرحيم   الاسم

Date of Birth   21/08/71   ١٣٩١/٠٧/٠١   تاريخ الميلاد

ID No.   1000146819   Sex   M   نكر الجنس   السجل المدني

Date of Issue   29/01/2000   ١٤٢٠/١٠/٢٢   تاريخ الإصدار

Date of Expiry   تاريخ الانتهاء   Issuing Authority   مكان الإصدار
04/12/04   ١٤٢٥/١٠/٢٢   مركز المعلومات

P<ASUMAHMUT<ABDUL<RAHIIM<<<<<<<<<<<<<<<<<<<<
A2222222<0ASU7

8.3 *Languages and characters*.
When the mandatory elements of Zones I, II and III are in a national language that does not use the Latin alphabet, a transliteration shall also be provided.

Status quo: name in VIZ copied to MRZ

Many phonetic transcriptions of same Arabic name

Many phonetic transcriptions

Many phonetic transcriptions

Many phonetic transcriptions

Many phonetic transcriptions

MRZ same as VIZ   **Much variation in (Latin) name**

IDENTIFICATION: the Arabic name is unique

So the solution is to base the MRZ on the Arabic Name.

But the MRZ can only contain OCR-B A-Z and <

Passport / جواز سفر	Type P	Country Code ASU	Passport No. A2222222

محمود عبدالرحيم

Name Mahmut Abdul Rahim الاسم

Date of Birth 21/08/71 تاريخ الميلاد ١٣٩١/٠٧/٠١

ID No. 1000146819 الجنس ذكر Sex M السجل المدني

Date of Issue 29/01/2000 تاريخ الإصدار ١٤٢٠/١٠/٢٢

Date of Expiry تاريخ الانتهاء Issuing Authority مكان الإصدار
04/12/04 ١٤٢٥/١٠/٢٢ مركز المعلومات

محمود عبدالرحيم

P<ASUMAHMUT<ABDUL<RAHIIM<<<<<<<<<<<<
A2222222<0ASU7108211M0412040I000146819<<<<33

9.4.1 Names in the MRZ are represented differently from those in the VIZ. National characters must be transliterated using only the allowed OCR character set *[A..Z]*.

Solution: Transliteration of Arabic name into Latin

Transliteration table based on closest match + 'escape'

Technical Report – Appendix 1

Unicode	Arabic letter	Name	MRZ
0621	ء	hamza	XE
0622	آ	alef with madda above	XAA
0623	أ	alef with hamza above	XAE
0624	ؤ	waw with hamza above	XW
0625	إ	alef with hamza below	I
0626	ئ	yeh with hamza above	XI
0627	ا	alef	A
0628	ب	beh	B
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

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

Technical Report – Appendix 1

Unicode	Arabic letter	Name	MRZ
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
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
0648	و	waw	W

Technical Report – Appendix 1

Unicode	Arabic letter	Name	MRZ
0649	ى	alef maksura	XAY
064A	ي	yeh	Y
0651	ّ	shadda	[DOUBLE][1]
0671	أ	alef wasla	XXA
0679	ٹ	Tteh	XXT
067E	پ	Peh	P
067C	ټ	teh with ring	XRT
0681	ځ	hah with hamza above	XKE
0685	څ	hah with 3 dots above	XXH
0686	چ	Tcheh	XC
0688	ڈ	Ddal	XXD
0689	ډ	dal with ring	XDR
0691	ڑ	Rreh	XXR
0693	ړ	reh with ring	XRR
0696	ږ	reh with dot below and dot above	XRX
0698	ژ	Jeh	XJ
069A	ښ	seen with dot below and dot above	XXS

[1] Shadda denotes doubling: Latin character or sequence is repeated eg عبّاس becomes E**BB**AS; فضّة becomes F**XDZXDZ**XAH.

Technical Report – Appendix 1

Unicode	Arabic letter	Name	MRZ
06A9	ک	keheh	XKK
06AB	ڪ	kaf with ring	XXK
06AD	ڭ	Ng	XNG
06AF	گ	gaf	XGG
06BA	ں	noon ghunna	XNN
06BC	ڼ	noon with ring	XXN
06BE	ھ	heh doachashmee	XDO
06C0	ۀ	heh with yeh above	XYH
06C1		heh goal	XXG
06C2		heh goal with hamza above	XGE
06C3		teh marbuta goal	XTG
06CC	ی	farsi yeh	XYA
06CD	ۍ	yeh with tail	XXY
06D0	ې	Yeh	Y
06D2	ے	Yeh barree	XYB
06D3	ۓ	yeh barree with hamza above	XBE

Some Arabic characters have near matches

'H' is already assigned to "Heh", so use 'X' as escape

Heh -> 'H', Hah -> 'XH'

"Ain" has no exact Latin equivalent, so assign 'E'

Reiterate: MRZ is different form of name to VIZ

9.4.1 Names in the MRZ are represented differently from those in the VIZ. National characters must be transliterated using only the allowed OCR character set *[A..Z].*

Reiterate: MRZ is different form of name to VIZ

Passport / جواز سفر	Type	Country Code	Passport No.
	P	ASU	A2222222

Name MAHMUT ABDUL RAHIIM   محمود عبدالرحيم   الاسم

Date of Birth   21/08/71   تاريخ الميلاد   ١٣٩١/٠٧/٠١

ID No.   1000146819   الجنس ذكر   Sex   M   السجل المدني

Date of Issue   29/01/2000   تاريخ الإصدار   ١٤٢٠/١٠/٢٢

Date of Expiry   تاريخ الانتهاء   Issuing Authority   مكان الإصدار
04/12/04   ١٤٢٥/١٠/٢٢   مركز المعلومات

P<ASUMXHMWD<EB
A2222222<0ASU7

9.1.3 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 MRPs. **The MRZ is a different representation of the data than is found in the VIZ.**

Transliteration - advantage

Name in MRZ is unique (= Arabic name)

Arabic name direct from MRZ

Chip holds name in DG11 (Unicode), can be compared

Passport / جواز سفر Type P Country Code ASU Passport No. A2222222

Name MAHMUT ABDUL RAHIIM   الاسم   محمود عبدالرحيم

Date of Birth   21/08/71   ١٣٩١/٠٧/٠١   تاريخ الميلاد

ID No.   1000146819   Sex M   الجنس ذكر   السجل المدني

Date of Issue   29/01/2000   ١٤٢٠/١٠/٢٢   تاريخ الإصدار

Date of Expiry   تاريخ الانتهاء   Issuing Authority   مكان الإصدار
04/12/04   ١٤٢٥/١٠/٢٢   مركز المعلومات

P<ASUMXHMWD<EBDALRXHYM<<<<<<<<<<<<<<<<<<<<<<

A2222222<0ASU7108211M0412040100146819<<<<33

محمود   عبدالرحيم

Detail(s) Recorded in MRZ	DG1	Document Type
		Issuing State or organization
		Name (of Holder)
		Document Number
		Check Digit - Doc Number
		Nationality
		Date of Birth
		Check Digit - DOB
		Sex
		Data of Expiry or Valid Until Date
		Check Digit DOE/VUD
		Optional Data
		Check Digit - Optional Data Field
		Composite Check Digit
Encoded Identification Feature(s)		Global Interchange Feature
	DG2	Encoded Face
	DG3	Additional Feature(s)	Encoded Finger(s)
	DG4	Encoded Eye(s)
Displayed Identification Feature(s)	DG5	Displayed Portrait
	DG6	Reserved for Future Use
	DG7	Displayed Signature or Usual Mark
Encoded Security Feature(s)	DG8	Data Feature(s)
	DG9	Structure Feature(s)
	DG10	Substance Feature(s)

محمود   عبدالرحيم

محمود   8,
062F,0627,0644,0631,062D,064A,0645

IMPLEMENTATION ISSUES

IMPLEMENTATION – Legacy Database
and Border Control

Passport / جواز سفر   Type   Country Code   Passport No.
                       P   ASU   A2222222

Name MAHMUT ABDUL RAHIIM   محمود عبدالرحيم   الاسم

Date of Birth   21/08/71   ١٣٩١/٠٧/٠١   تاريخ الميلاد

ID No.   1000146819   الجنس ذكر   Sex   M   السجل المدني

Date of Issue   29/01/2000   ١٤٢٠/١٠/٢٢   تاريخ الإصدار

Date of Expiry   تاريخ الانتهاء   Issuing Authority   مكان الإصدار
04/12/04   ١٤٢٥/١٠/٢٢   مركز المعلومات

P<ASUMXHMWD<EBDALRXHYM<<<<<<<<<<<<<<<<<<<<<
A2222222<0ASU7108211M04120401000146819<<<33

محمود   عبدالرحيم

Original Arabic form

MEHMOOD ABD AL RAHEEM

MAHMOOD ABDUL RAHIM

MAHMUT ABD AR-RAHEEM

MAHMUD ABDALRAHEEM

MAHMOUD ABD-AL-RAHIIM

MAHMUT ABDUL RAHIIM

Other variations can be derived.

IMPLEMENTATION – Name in VIZ

1. **Use Civil Register (transcription)**

   or

2. Status quo transcription

   or

3. Use MRZ transliteration

IMPLEMENTATION - PNR

Only concerned when PNR is used for API, not with PNR for airline use.

IATA/CAWG "API Statement of Principles"

presented to

FACILITATION (FAL) DIVISION
— TWELFTH SESSION
Cairo, Egypt, 2004-3-22/4-2

stated that:

"**Required API data should be limited to the data contained in the machine-readable zone of travel documents** or obtainable from existing government databases, such as those containing visa issuance information."

IMPLEMENTATION – AIRLINE CHECKING

1. PNR = MRZ

2. PNR = both VIZ and MRZ

CONCLUSION

1. Essential – solve identity management issue, required for Interpol, aviation security, etc

2. Using the original name in Arabic is only way

3. Arabic name is now machine readable

4. Will be transitional implementation problems, not impossible, but worthwhile goal

Transliteration Rules - Arabic (WP17)

Mike Ellis
ISO
(JTC1 SC17/WG3 TF3)

on behalf of
New Technologies Working Group
(NTWG)

presented to
ICAO TAG/MRTD 20 – Montreal, September 2011
20th Meeting of the Technical Advisory Group on
Machine Readable Travel Documents

V3.4   09/2011