# Transliteration Rules - Arabic (WP17)



Mike Ellis ISO (JTC1 SC17/WG3 TF3)

New Technologies Working Group (NTWG) TAG/MRTD 20 20<sup>th</sup> Meeting of the Technical Advisory Group on Machine Readable Travel Documents

V3.4 09/2011



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

جوال سفر Passport / جوال	Type Country Code Passport No. P ASU A2222222
Jacob Contraction	Name Mahmut Abdul Rahiim محصود عبدالرحيم
(aa)	تاريخ الميك ١٣١١/٠٧/٠١ تاريخ الميك
121	ID No. 1000146819 السجل المدنى Sex M المجنس نكر المدنى
	تاريخ الإصدار ۲۲/۱۰/۲۰ معرين الإصدار ۲۲/۱۰/۲۰ Date of Issue 29/01/2000
	Date of Expiry تاريخ الانتهاء Issuing Authority مكان الإصدار
	مركز المطومات ٥٤/١٤/٥٩ ٥٤
P <asumahmut<< th=""><th>ABDUL<rahiim<<<<<<<<<<< th=""></rahiim<<<<<<<<<<<></th></asumahmut<<>	ABDUL <rahiim<<<<<<<<<<< th=""></rahiim<<<<<<<<<<<>
A2222222<0AS	<b>07</b> 8.3 Languages and characters. <b>When the mandatory elements of Zones I, II and III are</b>
	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

لاسم محمود عب
ناريخ المبلك ١/٠٧/٠١
نجنس نکر Sex M
تاريخ الإصدار ٢٢/١٠/
ority مكان الإصدار
مركز المطومات



# Many phonetic transcriptions of same Arabic name

جواز سار / Passport	Type Cou P AS		Passport No. A2222222	
- State	Name Mahmo	oud Abd-al-I	ود عبدالرحيم Rahiim	וצייא הבה
	Date of Birth	21/08/71	1711/. 4/. 1	تاريخ الميلاد
1251	ID No.	1000146819	Sex N السجل العدني	الجنس نكر ا
	Date of Issue	29/01/2000	117./1./77	تاريخ الإصدار
	Date of Expiry	تاريخ الالتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77	4	مركز المعلومة
P <asumahmoud< td=""><td>&lt; A B D &lt; A L &lt; R A</td><td>\HIIM &lt;&lt;&lt;</td><td></td><td></td></asumahmoud<>	< A B D < A L < R A	\HIIM <<<		
A2222222<0AS	U7108211M	04120401	000146819<<	<<33



جواز سفر / assport	Type Cou P AS		Passport No. A2222222	
1.599.0	Name Mahm	ud Abdalral	ود عبدالرحيم heem	וצייק אבת
(a)	Date of Birth	21/08/71	1791/. 4/. 1	تاريخ المياك
12-1	ID No.	1000146819	Sex N السجل المدنى	الجنس ذكر ا
NE/	Date of Issue	29/01/2000	117./1./77	تاريخ الإصدار
	Date of Expiry	تاريخ الانتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77	4	مركز المطومة
P <asumahmud<< td=""><td>ABDALRAHE</td><td>EM &lt;&lt;&lt;&lt;&lt;</td><td></td><td></td></asumahmud<<>	ABDALRAHE	EM <<<<<		
A2222222<0AS	U7108211M	04120401	1000146819<<	<<33



جواز طر / Passport	Type Cou P AS	intry Code	Passport No. A2222222	
all a	Name Mahm	ut Abd ar-R	ود عبدالرحيم aheem	الأسم مده
laal	Date of Birth	21/08/71	1711/. 4/. 1	تاريخ المباك
1241	ID No.	1000146819	Sex N السجل المدنى	الجنس ذكر
	Date of Issue	29/01/2000	117./1./77	تاريخ الإصدار
	Date of Expiry	تاريخ الالتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77	3	مركز المعلوما
P <asu<mark>MAHMUT&lt;A A22222222&lt;0ASU</asu<mark>	and the second			



جواڑ سار / assport	Type Cou P AS		Passport No. A2222222	
CARGO .	Name Mahm	ood Abdul F	ود عبدالرحيم Rahim	الأسم محم
100)	Date of Birth	21/08/71	1711/. 4/. 1	تاريخ الميلاد
12-1	ID No.	1000146819	Sex M السجل المدنى	الجنس ذكر ا
NE/	Date of Issue	29/01/2000	117./1./77	تاريخ الإصدار
	Date of Expiry	تاريخ الالتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77	3	مركز المطوما
P <asumahmood< td=""><td><abdul <rai<="" td=""><td>HIM &lt;&lt;&lt;&lt;&lt;</td><td></td><td></td></abdul></td></asumahmood<>	<abdul <rai<="" td=""><td>HIM &lt;&lt;&lt;&lt;&lt;</td><td></td><td></td></abdul>	HIM <<<<<		
A2222222<0AS	U7108211M	0412040	1000146819<<	<<33



CAR STAR	Name Mehmo	od Abd al R	aheem محمود عبدالرحيم
11	Date of Birth	21/08/71	and possibly over 9,000 more
12-1	ID No.	1000146819	الجنس نكر Sex M السجل المدنى
Nex.	Date of Issue	29/01/2000	تاريخ الإصدار ۲۲/۱۰/۱۰
	Date of Expiry 04/12/04	تاريخ الالتهاء ١٤٢٥/١٠/٢٢	مكان الإصدار Issuing Authority مركز المعلومات

**MRZ** same as **VIZ** 

## Much variation in (Latin) name



# **IDENTIFICATION: the Arabic name is unique**

	P ASU A2222222
1.000	Name Mahmut Abdul Rahum
100	تاريخ العبان ١٣٩١/٠٧/٠١ تاريخ العبان Date of Birth 21/08/71
12-1	ID No. 1000146819 السجل المدنى Sex M الجنس نكر الم
XEA	تاريخ الإصدار ۲۲/۱۱،۲۰ ۱۱۲۰،۱۱،۲۲ Date of Issue 25/01/2000
	Date of Expiry تاريخ الالتهاء Issuing Authority
	مركز المعلومات ٥٤/١٤/٧٩ ١٤٢٥/١٠/٢٢
P <asuma< td=""><td>مركز المطومات 04/12/04 ١٤٢٥/١٠/٢٢ &gt;</td></asuma<>	مركز المطومات 04/12/04 ١٤٢٥/١٠/٢٢ >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>

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



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

جوال سفر Passport /	Type   Country Code   Passport No.     P   ASU   A2222222
1.000	Name Mahmut Abdul Rankm
last	تاريخ المباط ١٣٩١/٠٧/٠١ تاريخ المباط المرابع المالية المباط
12-1	ID No. 1000146819 السجل المدنى Sex M الجنس نكر Sex M
NE/	تاريخ الإصدار ۲۲/۱۰/۱۰ مناريخ الإصدار ۲۲/۱۰/۱۰ Date of Issue 29/01/2000
	مكان الإصدار Issuing Authority تاريخ الانتهاء Date of Expiry مركز المعلومات 1/1/1/14 ۱۹/۱۰/۱۹
P < A S U M A MUS <	RAHII ACA CARALIA
	U7108211M04120401000146819<<<<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 [AZ].



# Solution: Transliteration of Arabic name into Latin

جراز سار Passport /	and the second second	untry Code SU	Passport No. A2222222	
CARGON .	Name MAHMU	IT ABDUL RA	رد عبدالرحيم HIIM	
100	Date of Birth	21/08/71	1811.41.	تاريخ العياك 1
12-1	ID No.	1000146819	Sex ۱ سېل قىدنى	الچنس ذکر ۷
) el	Date of Issue	29/01/2000	\iT./\./TT	تاريخ الإصدار
	Date of Expiry	تاريخ الانتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77		مركز المعلومان
P <asumxhmwd<< td=""><td>EBDALRXHY</td><td>M&lt;&lt;&lt;&lt;&lt;</td><td>*****</td><td></td></asumxhmwd<<>	EBDALRXHY	M<<<<<	*****	
A2222222<0AS	U7108211M	0412040	1000146819<	<<<33

## Transliteration table based on closest match + 'escape'

# Technical Report – Appendix 1

	Arabic letter	Name	MRZ
Unicode			
0621	۶	hamza	XE
0622	Ĩ	alef with madda above	XAA
0623	ş	alef with hamza above	XAE
0624	ۇ	waw with hamza above	XW
0625	) E	alef with hamza below	I
0626	ئ	yeh with hamza above	XI
0627	1	alef	A
0628	Ļ	beh	В
0629	õ	teh marbuta	XTA/XAH <u>[</u> 1]
062A	ت	teh	Т
062B	ث	theh	ХТН
062C	٤	jeem	J
062D	۲	hah	ХН
062E	Ż	khah	ХКН
062F	د	dal	D
0630	ć	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

	Arabic letter	Name	MRZ
Unicode			
0632	ز	zain	Z
0633	س	seen	S
0634	ش	sheen	XSH
0635	ص	sad	XSS
0636	ض	dad	XDZ
0637	ط	tah	ХТТ
0638	ظ	zah	XZZ
0639	٤	ain	E
063A	ż	ghain	G
0641	ف	feh	F
0642	ق	qaf	Q
0643	ك	kaf	к
0644	ل	lam	L
0645	م	meem	м
0646	ن	noon	N
0647	٥	heh	Н
0648	و	waw	W



# Technical Report – Appendix 1

	Arabic letter	Name	MRZ
Unicode			
0649	ى	alef maksura	ХАҮ
064A	ي	yeh	Y
0651	័	shadda	[DOUBLE][1]
0671	Ĩ	alef wasla	XXA
0679	ڴ	Tteh	ХХТ
067E	پ	Peh	Р
067C	ټ	teh with ring	XRT
0681	ځ	hah with hamza above	ХКЕ
0685	Ś	hah with 3 dots above	ХХН
0686	چ	Tcheh	ХС
0688	۲ د	Ddal	XXD
0689	¢	dal with ring	XDR
0691	رط	Rreh	XXR
0693	ړ	reh with ring	XRR
0696	(د	reh with dot below and dot above	XRX
0698	ژ	Jeh	XJ
069A	ښ	seen with dot below and dot above	XXS



فضّة becomes EBBAS; فضّة becomes FXDZXDZXAH.

# Technical Report – Appendix 1

	Arabic letter	Name	MRZ
Unicode			
06A9	ک	keheh	ХКК
06AB	ګ	kaf with ring	ХХК
06AD	ؿ	Ng	XNG
06AF	گ	gaf	XGG
06BA	ں	noon ghunna	XNN
06BC	ڼ	noon with ring	XXN
06BE	ھ	heh doachashmee	XDO
06C0	δ	heh with yeh above	ХҮН
06C1		heh goal	XXG
06C2		heh goal with hamza above	XGE
06C3		teh marbuta goal	XTG
06CC	ى	farsi yeh	ХҮА
06CD	ى	yeh with tail	ХХҮ
06D0	ې	Yeh	Y
06D2	۷	Yeh barree	ХҮВ
06D3	ź	yeh barree with hamza above	XBE





### **Some Arabic characters have near matches**

جراز سار (Passport	Type Cou P AS	ntry Code U	Passport No. A2222222	Meer
COLUMN.	Name MAHMUT	T ABDUL RA	عبدالرحيم	
(aa)	Date of Birth	21/08/71	1811.41.1	نتريخ لعباك
121	ID No.	1000146819	Sex N	الجنس تعر
	Date of Issue	29/01/2000	117.11.111	تاريخ الإصدار
	Date of Expiry	تاريخ الانتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77	4	مركز المعلومة
P <asu< td=""><td></td><td>*****</td><td></td><td></td></asu<>		*****		
A2222222<0AS	U7108211M	0412040	1000146819<-	<<33



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

جراز بلر Passport /	Type Cou P AS	ntry Code U	Passport No. A2222222		lleb
TO MAR	Name MAHMU	T ABDUL RA	عبدالرحيم	Lass public	Hah محمو
100	Date of Birth	21/08/71	1811.41.1	تاريخ فسالة	
12-1	ID No.	1000146819	Sex M	الچنس ذکر ۸	
X	Date of Issue	29/01/2900	117.11.177	تاريخ الإصدار	
	Date of Expiry	تاريخ الانتهاء	Issuing Authority	مكان الإصدار	
	04/12/04	1170/1./7		مركز المطومة	
P <asum< td=""><td></td><td>*****</td><td>******</td><td></td><td></td></asum<>		*****	******		
A2222222<0AS	U7108211M	0412040	1000146819<-	<<33	

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













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

جراز بار Passport / جراز ما	P ASU	itry Code J	Passport No. A2222222	
The second second	Name MAHMUT	ABDUL RA	are the ere	محمود ع
100	Date of Birth	21/08/71	1==1	تاريخ العباك
121	ID No.	1000146819	Sex ۸ السجل العدني	الچنس تکر ۸
	Date of Issue	29/01/2000	117./1./77	تاريخ الإصدار
	Date of Expiry	تاريخ الانتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./71	Sama a line	مركز المطومة
P <asum td="" xhmwd<=""><td></td><td>&lt;&lt;&lt;&lt;&lt;</td><td>******</td><td></td></asum>		<<<<<	******	
A2222222<0AS	U7108211M0	412040	1000146819<-	<<33

OAC



































# **Reiterate: MRZ is different form of name to VIZ**

جراز بنزر / Passport	Type Cou P AS	intry Code	Passport No. A2222222	
alles.	Name MAHMU	T ABDUL RA	ود عبـدالرحيم HIIM	
(a)	Date of Birth	21/08/71	1811.41.1	تاريخ العياك
12-1	ID No.	1000146819	Sex M السجل العدني	الچنس تکر ۸
LE/	Date of Issue	29/01/2000	117./1./17	تاريخ الإصدار
	Date of Expiry	تاريخ الالتهاء	Issuing Authority	مكان الإصدار
	04/12/04	1170/1./77	-	مركز المطومة
P <asumxhmwd-< td=""><td>EBDALRXHY</td><td>M&lt;&lt;&lt;&lt;&lt;</td><td></td><td></td></asumxhmwd-<>	EBDALRXHY	M<<<<<		
A2222222 < OAS	U7108211M	0412040	1000146819<-	<<33
	941	Names ir	the MRZ are	renrese

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/ برای ا	Type Country Code Passport No.   P ASU A2222222   Name MAHMUT ABDUL RAHIIM محصود عبدالزحيم   Date of Birth 21/08/71   ID No. 1000146819   تاريخ الإصدار ٢٠/٠٠/٠٠   تاريخ الإصدار ٢٠/٠٠/٠٠   Date of Issue 29/01/2000   ١٤٠٠٠ ١٤٠٠٠   تاريخ الإصدار ٢٠/٠٠/٠٠   تاريخ الإصدار ٢٠/٠٠/٠٠
P <asum xhmwd<8<br="">A2222222&lt;0ASU</asum>	9.1.3 The data in the MRZ are formatted in such a way as

representation of the data than is found in the VIZ.



# **Transliteration - advantage**

جراز سار / assport	Type Count P ASU	ry Code Passpo A222		
CARD .	Name MAHMUT	ABDUL RAHIIM	محمود عبدالرحيم	ידיק
las	Date of Birth 21	//08/71	1811/. V/. 1 Sin	ध्रम्म I
12-1	ID No. 10	ھىدنى 000146819	نکر Sex M انسجل	الجنس
) el	Date of Issue	29/01/2000	لاصدار ۲۲/۱۰/۲۲	تاريخ ا
	Date of Expiry	Issuin تاريخ التنهاء	ig Authority المدار	مكان ا
	04/12/04	1170/1./77	المطومات	مركل
P <asumxhmwd<< td=""><td>EBDALRXHYM</td><td></td><td></td><td>&lt;</td></asumxhmwd<<>	EBDALRXHYM			<
A2222222<0AS	U7108211M04	1204010001	46819<<<3	3

# Name in MRZ is unique (= Arabic name)



# Arabic name direct from MRZ

جراز سار / assport	P AS	Contraction of the second s	Passport No. A2222222	
CARDON .	Name MAHMU	T ABDUL RAH	ود عبدالرحيم IIM	11-1-1 aca-
(a)	Date of Birth	21/08/71	1811.41.1	تاريخ العباك
12-1	ID No.	1000146819	Sex ۸ سېل قىدنى	لچنس تکر ۸
121	Date of Issue	29/01/2000	117.11.177	تاريخ الإصدار
	Date of Expiry	تاريخ الانتهاء ١١٢٥/١٠/٢٢	Issuing Authority	مكان الإصدار
P <asumxhmwd<< td=""><td></td><td></td><td></td><td>مركز المطومة &gt;</td></asumxhmwd<<>				مركز المطومة >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
A2222222<0AS	and the second states of the			
عبدالرحيم	100			
عبد الرحيم	مود			



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

چر ( سقر / issport	Type   Country Code   Passport No.     P   ASU   A2222222
	Name MAHMUT ABDUL RAHIIM محمود عبدالرحوم
60	تريخ العين ١٣١١/٠٧/٠١ عنويغ العين Date of Birth 21/08/71
N.J.	ID No. 1000146819 السجل المدتى Sex M المجنس نكر M
NA	تاريخ الإصدار Date of Issue 29/01/2000 ١٤٢٠/١٠/٢٠ تاريخ الإصدار
	مكان الاصدار Issuing Authority تاريخ الانتهاء Date of Expiry
and the state	مركز المعلومات 04/12/04 ١٤٢٥/١٠/٢٧
<pre>&gt;<asumxhmwd< pre=""></asumxhmwd<></pre>	<pre><ebdalrxhym<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<< td=""></ebdalrxhym<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<></pre>
2222222 <oa< td=""><td>su<mark>7108211M04120401000146819&lt;&lt;&lt;&lt;33</mark></td></oa<>	su <mark>7108211M04120401000146819&lt;&lt;&lt;&lt;33</mark>
ر جيم	محمود عبدال



	Provide Line of the		and the second second			
and the second			Docum	ent Type		
		Issuing State or organization				
	1000 M	Name (of Holder)				
		[	Documer	nt Number		
Detailies		Che	ck Digit -	Doc Number		
Detail(s) Recorded	DG1		Natio	onality		
in MRZ		California California	Date of	of Birth		
IVITIZ		(	Check Di	git - DOB		
22		a cher a start	Sex			
and the second		Data of Expiry or Valid Until Date				
		Check Digit DOE/VUD				
	a state	Optional Data				
and the second	a the second	Check Digit - Optional Data Field				
		Co	mposite	Check Digit		
Encoded	Glob	al Interchange Feature	DG2	Encoded Face		
Identification Feature(s)		Additional		Encoded Finger(s)		
	F	Feature(s)	DG4	Encoded Eye(s)		
Displayed	DG5	E	Displayed	d Portrait		
Identification Feature(s)	DG6	Res	erved for	Future Use		
r cuture(s)	DG7	Displayed	Signatu	ure or Usual Mark		
Encoded	DG8		Data Fe	ature(s)		
Security Feature(s)	DG9	St	tructure I	Feature(s)		
(earme(s)	DG10	Su	bstance	Feature(s)		



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

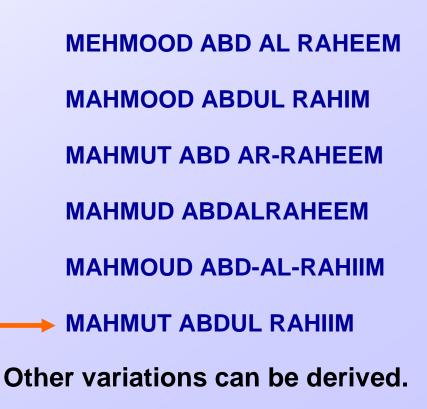


# **IMPLEMENTATION ISSUES**

# IMPLEMENTATION – Legacy Database and Border Control









# **IMPLEMENTATION – Name in VIZ**



- 1. Use Civil Register (transcription)
- or 2. Status quo transcription

or 3. Use MRZ transliteration

assport /	جراز سار	and the second second	untry Code SU	Passport No. A2222222	
5.0		Name MAHMU	T ABDUL RA	ود عبـدالرحيم HIIM	
80	-	Date of Birth	21/08/71	1751/	تاريخ لعباك
12	51	ID No.	1000146819	Sex N السجل العدني	الچنس ذكر ا
	31	Date of Issue	29/01/2000	117./1./17	تاريخ الإصدار
		Date of Expiry	تاريخ الانتهاء	Issuing Authority	مكان الإصدار
		04/12/04	1170/1./77	-	مركز المطومة

A2222222<0ASU7108211M04120401000146819<<<<33

# **IMPLEMENTATION - PNR**



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

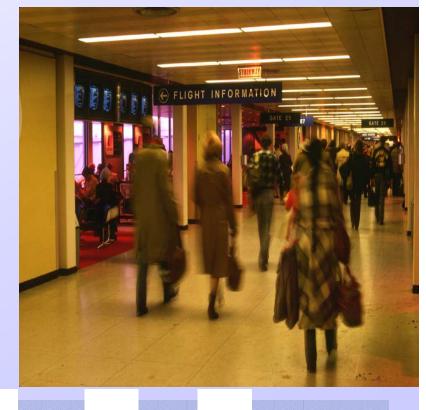
**IATA/CAWG "API Statement of Principles"** 

presented to

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

#### stated that:

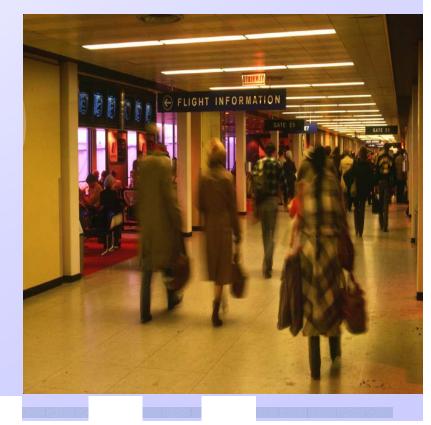
**"Required API data should be limited to the data contained in the machine-readable zone of travel documents** or obtainable from existing government databases, such as those containing visa issuance information."



# **IMPLEMENTATION – AIRLINE CHECKING**

1. PNR = MRZ

2. PNR = both VIZ and MRZ



. OAC

# CONCLUSION



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

2. Using the original name in Arabic is only way

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



TAG is asked to:

1. Note work being undertaken (see also the Technical Report)

2. Provisionally approve, contingent upon comments at the Qatar Regional Seminar in November

**3. Approve eventual inclusion as Recommended Practice in Appendix 9 to Section IV** 



# Transliteration Rules - Arabic (WP17)

# Mike Ellis ISO (JTC1 SC17/WG3 TF3)

#### New Technologies Working Group (NTWG) TAG/MRTD 20

20<sup>th</sup> Meeting of the Technical Advisory Group on Machine Readable Travel Documents

V3.4 09/2011