

**TLD:** sa  
**Language Tag:** AR  
**Language Description:** Arabic  
**Version:** 2.0  
**Effective Date:** 07 Dec 2010

**Registry:** Saudi Network Information Center  
**Contact:** Abdulaziz Al-Zoman [azoman\[at\]citc.gov.sa](mailto:azoman[at]citc.gov.sa)  
**Address:** SaudiNIC, CITC, P.O. Box 75606,  
**Riyadh** 11588, Saudi Arabia  
**Telephone:** +966-1-263-9233 Fax: +966-1-263-9393  
**Website:** <http://www.nic.sa/>

### Relevant Policy Documents:

- Saudi Domain Name Registration Regulation:  
[http://nic.sa/docs/Saudi\\_Domain\\_Name\\_Registration\\_Regulation\\_V3.0\\_EN.pdf](http://nic.sa/docs/Saudi_Domain_Name_Registration_Regulation_V3.0_EN.pdf)
- Guideline Rules for writing Arabic IDNs under the IDN ccTLD (السعودية):  
[http://nic.net.sa/docs/Guidelines\\_for\\_writing\\_Arabic\\_IDNs\\_under\\_the\\_IDN\\_ccTLD\\_V1.2-en.pdf](http://nic.net.sa/docs/Guidelines_for_writing_Arabic_IDNs_under_the_IDN_ccTLD_V1.2-en.pdf)

### Introduction:

This document provides the IDN (Internationalized Domain Names) Language Table and guideline rules to be used for writing and registering Arabic Domain names under xn--mgberp4a5d4ar (السعودية) IDN ccTLD. These are based on the recommendations outlined in the RFC 5564: "Linguistic Guidelines for the Use of the Arabic Language in Internet Domains", that can be found in the following URL:

<http://tools.ietf.org/html/rfc5564>.

### Permitted Characters and Symbols in writing an Arabic Label:

code	Character/symbol	Unicode name
0621	ء	ARABIC LETTER HAMZA
0622	أ	ARABIC LETTER ALEF WITH MADDA ABOVE
0623	آ	ARABIC LETTER ALEF WITH HAMZA ABOVE
0624	ؤ	ARABIC LETTER WAW WITH HAMZA ABOVE
0625	إ	ARABIC LETTER ALEF WITH HAMZA BELOW
0626	ئ	ARABIC LETTER YEH WITH HAMZA ABOVE
0627	ا	ARABIC LETTER ALEF
0628	ب	ARABIC LETTER BEH
0629	ة	ARABIC LETTER TEH MARBUTA
062A	ت	ARABIC LETTER TEH
062B	ث	ARABIC LETTER THEH
062C	ج	ARABIC LETTER JEEM
062D	ح	ARABIC LETTER HAH
062E	خ	ARABIC LETTER KHAH
062F	د	ARABIC LETTER DAL
0630	ذ	ARABIC LETTER THAL
0631	ر	ARABIC LETTER REH
0632	ز	ARABIC LETTER ZAIN
0633	س	ARABIC LETTER SEEN
0634	ش	ARABIC LETTER SHEEN
0635	ص	ARABIC LETTER SAD
0636	ض	ARABIC LETTER DAD
0637	ط	ARABIC LETTER TAH

0638	ظ	ARABIC LETTER ZAH
0639	ع	ARABIC LETTER AIN
063A	غ	ARABIC LETTER GHAIN
0641	ف	ARABIC LETTER FEH
0642	ق	ARABIC LETTER QAF
0643	ك	ARABIC LETTER KAF
0644	ل	ARABIC LETTER LAM
0645	م	ARABIC LETTER MEEM
0646	ن	ARABIC LETTER NOON
0647	هـ	ARABIC LETTER HEH
0648	و	ARABIC LETTER WAW
0649	ى	ARABIC LETTER ALEF MAKSURA
064A	ي	ARABIC LETTER YEH
0660	٠	ARABIC-INDIC DIGIT ZERO
0661	١	ARABIC-INDIC DIGIT ONE
0662	٢	ARABIC-INDIC DIGIT TWO
0663	٣	ARABIC-INDIC DIGIT THREE
0664	٤	ARABIC-INDIC DIGIT FOUR
0665	٥	ARABIC-INDIC DIGIT FIVE
0666	٦	ARABIC-INDIC DIGIT SIX
0667	٧	ARABIC-INDIC DIGIT SEVEN
0668	٨	ARABIC-INDIC DIGIT EIGHT
0669	٩	ARABIC-INDIC DIGIT NINE
0030	0	DIGIT ZERO
0031	1	DIGIT ONE
0032	2	DIGIT TWO
0033	3	DIGIT THREE
0034	4	DIGIT FOUR
0035	5	DIGIT FIVE
0036	6	DIGIT SIX
0037	7	DIGIT SEVEN
0038	8	DIGIT EIGHT
0039	9	DIGIT NINE
002D	-	HYPHEN-MINUS

## Variant Table for the Arabic Language:

```
#####
#
# This file describes IDN table of Variants for Arabic language that belong to this label
# xn--mgberp4a5d4ar. Each row in this file represents the relation of Arabic character
# (character which under the domain of Arabic language) with variant characters across
# Arabic script (based on UNICODE standard). Each relation has type of either Exact or Typo
# (Exact for mirror matching of both characters while Typo for Similarity look).
#
# Structure of this file:
# <CHAR>; <VCHAR1>(<POS>:<REL>), <VCHAR2>(<POS>:<REL>), <VCHAR3>(<POS>:<REL>), ...
#
# <CHAR> a unicode for character which is part of Arabic language domain
# <VCHAR#> a unicode for variant character which is part of Arabic script domain and has relation with <CHAR>
# <POS> a combination of B,M,F and I (B: Beginning, M: Medial, F: Final, I: Isolated)
# <REL> relation type of both <CHAR> and <VCHAR#> either E for Exact of T for Typo
#
# Example:
# 0646; 06BA(BM:E)
# The previous line shows that the Unicode char(06BA) is exactly similar to the Unicode Char (0646) in
# both "Beginning" and "Medial" positions.
#
```

```
# Authors:
# Abdulaziz Al-Zoman (azoman[at]citc.gov.sa)
# Raed Al-Fayez (rfayez[at]citc.gov.sa)
# Abdulrahman I. AL-Ghadir (aghadir[at]citc.gov.sa)
#
# SaudiNIC
#
#####

0621;
0622; 0671 (FI:T)
0623; 0672 (FI:T), 0675 (FI:T)
0624; 0676 (FI:T)
0625; 0673 (FI:T)
0626; 06D3 (FI:T), 0678 (FI:T), 0678 (BM:E)
0627; 0622 (FI:T), 0623 (FI:T), 0625 (FI:T)
0628;
0629; 06C3 (F:T), 06C3 (I:E)
062A; 067A (BMFI:T)
062B; 067D (BMFI:T), 06BD (FI:T), 06BD (BM:E)
062C;
062D;
062E;
062F;
0630;
0631;
0632;
0633;
0634;
0635;
0636;
0637;
0638;
0639;
063A;
0641; 06A7 (FI:T), 06A7 (BM:E)
0642;
0643; 06A9 (FI:T), 06A9 (BM:E), 06AA (BMFI:T)
0644;
0645;
0646; 06BA (BM:E)
0647; 06BE (BME:E), 06C1 (I:E), 06C1 (MF:T), 06D5 (FI:E)
0648;
0649; 06CD (FI:T), 06D2 (FI:T), 06CC (FI:E)
064A; 067B (BM:T), 06D0 (BMFI:T), 06CC (BM:E)

0660; 0030 (BMFI:T), 06F0 (BMFI:E)
0661; 0031 (BMFI:T), 06F1 (BMFI:E)
0662; 0032 (BMFI:T), 06F2 (BMFI:E)
0663; 0033 (BMFI:T), 06F3 (BMFI:E)
0664; 0034 (BMFI:T), 06F4 (BMFI:T)
0665; 0035 (BMFI:T), 06F5 (BMFI:T)
0666; 0036 (BMFI:T), 06F6 (BMFI:T)
0667; 0037 (BMFI:T), 06F7 (BMFI:E)
0668; 0038 (BMFI:T), 06F8 (BMFI:E)
0669; 0039 (BMFI:T), 06F9 (BMFI:E)

0030; 0660 (BMFI:T), 06F0 (BMFI:T)
0031; 0661 (BMFI:T), 06F1 (BMFI:T)
0032; 0662 (BMFI:T), 06F2 (BMFI:T)
0033; 0663 (BMFI:T), 06F3 (BMFI:T)
0034; 0664 (BMFI:T), 06F4 (BMFI:T)
0035; 0665 (BMFI:T), 06F5 (BMFI:T)
0036; 0666 (BMFI:T), 06F6 (BMFI:T)
0037; 0667 (BMFI:T), 06F7 (BMFI:T)
0038; 0668 (BMFI:T), 06F8 (BMFI:T)
0039; 0669 (BMFI:T), 06F9 (BMFI:T)

002D;
002E;

#EOF
```

## Guideline Rules for writing Arabic Labels:

**Rule 1:** Diacritics are not allowed

Example: سَجَل.السعودية Not Accepted  
سجل.السعودية Accepted

**Rule 2:** No mixing between scripts (Arabic & Latin)

Example: موقع.SaudiNIC-السعودية Not Accepted  
SaudiNIC.السعودية Not Accepted

**Rule 3:** Use of hyphen (instead of space) between words particularly if the 2 words will get connected

Example: هيئةالاتصالات.السعودية Accepted  
هيئة-الاتصالات.السعودية Accepted  
مدارس-خيف.السعودية Accepted  
مدارسخيف.السعودية Not Accepted

**Rule 4:** Hyphen cannot be used at the beginning or end of a label, and 2 or more consecutive hyphens are not allowed

Example: -هيئةالاتصالات.السعودية Not Accepted  
هيئةالاتصالات-.السعودية Not Accepted  
هيئة--الاتصالات.السعودية Not Accepted

**Rule 5:** Digits cannot be used at the beginning of a label. Digits can be used inside or at the end of a label from the 2 sets (Arabic-Indic digits and Arabic digits) without mixing.

Example: ٩٩٩.السعودية Not Accepted  
٩٩٩يساعدك.السعودية Not Accepted  
اتصل٩٩٩.السعودية Accepted  
٩٩وللنجدة.السعودية Not Accepted  
اتصل٩٩٩للنجدة.السعودية Not Accepted  
اتصل٩٩٩للنجدة.السعودية Accepted

It is permitted to register other variants of the domain name that can be formulated just by changing the digit set:

Example: اتصل٩٩٩للنجدة.السعودية Accepted

**Rule 6:** After registering a domain name that contains one or more of the confusingly similar characters (within the Arabic Language), e.g.:

- ALEF MAKSURA and YEH at the end of a word;
- All forms of ALEF (ALEF, ALEF WITH MADDA ABOVE, ALEF WITH HAMZA ABOVE, and ALEF WITH HAMZA BELOW);
- TEH MARBUTA and TEH at the end of a word,

it is permitted to register any of the domain name variants that can be created just by changing the confusingly similar characters provided that it does not infringe on the rights of others. This option is offered to provide security and stability to the registry as well as to domain name holders.

Example: شبكة-الأخبار.السعودية Accepted

It is permitted to register any of the domain name variants taking care of the consequences of the abovementioned condition:

Example: شبكة-الاخبار.السعودية Accepted with condition  
شبكة-الأخبار.السعودية Accepted with condition  
شبكة-الاخبار.السعودية Accepted with condition

**Rule 7:** It is permitted to register the domain name variants within the Arabic Script so that the Arabic domain names have presence and use at the global level by overcoming the problem of confusingly similarity of some of the Arabic language characters with characters from other languages using the Arabic script (such as Urdu, Farsi, Pashto, ...). The confusingly similar characters are identified in the above-mentioned language table. This take care of that fact that there are different input devices (e.g., keyboards) being used within the Arabic script communities.

Example: مكة.السعودية Accepted (all characters are from the Arabic Language)

It is permitted to register the other variants:

Example: مكة.السعودية Accepted (the letter KAF and TEH MARBUTA are from Urdu)