

Proposal for an Armenian Script Root Zone LGR

(TLDCON-2015, IDN Session)

*I.Mkrtumyan,
Armenian Generation Panel*



General information

Armenian IDN

- Code: Armn
- N°: 230
- English Name: Armenian
- English name of the script: Hye
- Native name of the script: huj

The announcement for the successful completion of Armenia's string evaluation has been posted at <https://www.icann.org/news/announcement-2014-11-20-en>.



Armenian language

- The **Armenian language** is an Indo-European language spoken by the Armenians. It is the official language of the Republic of Armenia and the self-proclaimed Nagorno-Karabakh Republic. It has historically been spoken throughout the Armenian Highlands and today is widely spoken in the Armenian Diaspora.



Armenian language

- Armenians has its own unique script, the Armenian alphabet, invented in 405 AD by Mesrop Mashtots.
- Linguists classify Armenian as an independent branch of the Indo-European language family.
- There are two standardized modern literary forms, Eastern Armenian and Western Armenian, with which most contemporary dialects are mutually intelligible.
- Total Armenian population in the world is about 10mln.



Geographic territories or countries with significant user communities for the script

Official language in Armenia and Nagorno Karabakh Republic
Big Diaspora using the Armenian language in

Argentina	Poland
Brazil	Romania
Cyprus	Russia
France	Syria
Georgia	Turkey
Hungary	Ukraine
Iran	United States
Iraq	Uruguay
Lebanon	



Armenian GP

- Armenian GP was created
- Armenian GP mailing list was created
Armeniangu@icann.org
- General information about the mailing list is at:
<https://mm.icann.org/mailman/listinfo/armeniangu>



Armenian MSR-2 table

- There are two standardized mutually intelligible modern literary forms, Eastern Armenian and Western Armenian with some difference in orthographies. But as both are written with the same script it doesn't influence the LGR document.

The following web site contains a full description of the Armenian repertoire for use in the Armenian writing system:










<http://www.omniglot.com/writing/armenian.htm>

MSR-2 Armenian Repertoire

0530

Armenian

058F

	053	054	055	056	057	058
0		Հ 0540	Ր 0550		Տ 0570	Ր 0580
1	Ա 0531	Զ 0541	Յ 0551	Մ 0561	Ճ 0571	Գ 0581
2	Բ 0532	Ղ 0542	Ի 0552	Բ 0562	Պ 0572	Լ 0582
3	Գ 0533	Ճ 0543	Փ 0553	Գ 0563	Ճ 0573	Փ 0583
4	Դ 0534	Մ 0544	Ք 0554	Դ 0564	Ճ 0574	Ք 0584
5	Ե 0535	Յ 0545	Օ 0555	Ե 0565	Ե 0575	Օ 0585
6	Զ 0536	Ն 0546	Ֆ 0556	Զ 0566	Ն 0576	Ֆ 0586
7	Է 0537	Շ 0547		Է 0567	Շ 0577	Է 0587
8	Ը 0538	Ո 0548		Ը 0568	Ո 0578	
9	Թ 0539	Չ 0549	Շ 0559	Թ 0569	Չ 0579	Շ 0589
A	Ժ 053A	Պ 054A	՝ 055A	Ժ 056A	Պ 057A	՝ 058A
B	Ի 053B	Ջ 054B	՛ 055B	Ի 056B	Ջ 057B	
C	Լ 053C	Ռ 054C	~ 055C	Լ 056C	Ռ 057C	
D	Խ 053D	Ս 054D	՝ 055D	Խ 056D	Ս 057D	
E	Վ 053E	Վ 054E	՝ 055E	Վ 056E	Վ 057E	
F	Կ 053F	Տ 054F	՛ 055F	Կ 056F	Տ 057F	Գ 058F

MSR-2 Armenian Repertoire

0531

Armenian

058F

Uppercase letters

0531	Ա	ARMENIAN CAPITAL LETTER AYB
0532	Բ	ARMENIAN CAPITAL LETTER BEN
0533	Գ	ARMENIAN CAPITAL LETTER GIM
0534	Դ	ARMENIAN CAPITAL LETTER DA
0535	Ե	ARMENIAN CAPITAL LETTER ECH
0536	Զ	ARMENIAN CAPITAL LETTER ZA
0537	Է	ARMENIAN CAPITAL LETTER EH
0538	Ը	ARMENIAN CAPITAL LETTER ET
0539	Թ	ARMENIAN CAPITAL LETTER TO
053A	Ճ	ARMENIAN CAPITAL LETTER ZHE
053B	Ի	ARMENIAN CAPITAL LETTER INI
053C	Լ	ARMENIAN CAPITAL LETTER LIWN
053D	Խ	ARMENIAN CAPITAL LETTER XEH
053E	Վ	ARMENIAN CAPITAL LETTER CA
053F	Կ	ARMENIAN CAPITAL LETTER KEN
0540	Հ	ARMENIAN CAPITAL LETTER HO
0541	Ջ	ARMENIAN CAPITAL LETTER JA
0542	Ղ	ARMENIAN CAPITAL LETTER GHAD
0543	Ճ	ARMENIAN CAPITAL LETTER CHEH
0544	Ս	ARMENIAN CAPITAL LETTER MEN
0545	Գ	ARMENIAN CAPITAL LETTER YI
0546	Ն	ARMENIAN CAPITAL LETTER NOW
0547	Շ	ARMENIAN CAPITAL LETTER SHA
0548	Ո	ARMENIAN CAPITAL LETTER VO
0549	Չ	ARMENIAN CAPITAL LETTER CHA
054A	Պ	ARMENIAN CAPITAL LETTER PEH
054B	Ջ	ARMENIAN CAPITAL LETTER JHEH
054C	Ռ	ARMENIAN CAPITAL LETTER RA
054D	Ս	ARMENIAN CAPITAL LETTER SEH
054E	Վ	ARMENIAN CAPITAL LETTER VEV
054F	Տ	ARMENIAN CAPITAL LETTER TIWN
0550	Ր	ARMENIAN CAPITAL LETTER REH
0551	Ծ	ARMENIAN CAPITAL LETTER CO
0552	Ի	ARMENIAN CAPITAL LETTER YIWN
0553	Փ	ARMENIAN CAPITAL LETTER PIWR
0554	Ք	ARMENIAN CAPITAL LETTER KEH
0555	Օ	ARMENIAN CAPITAL LETTER OH
0556	Ֆ	ARMENIAN CAPITAL LETTER FEH

Modifier letters

0559	◌ ^ʻ	ARMENIAN MODIFIER LETTER LEFT HALF RING • technical use
055A	◌ ^ˆ	ARMENIAN APOSTROPHE
055B	◌ ^ˆ	ARMENIAN EMPHASIS MARK
055C	◌ ^ˆ	ARMENIAN EXCLAMATION MARK
055D	◌ ^ˆ	ARMENIAN COMMA
055E	◌ ^ˆ	ARMENIAN QUESTION MARK
055F	◌ ^ˆ	ARMENIAN ABBREVIATION MARK

Lowercase letters

0561	ա	ARMENIAN SMALL LETTER AYB
0562	բ	ARMENIAN SMALL LETTER BEN
0563	գ	ARMENIAN SMALL LETTER GIM
0564	դ	ARMENIAN SMALL LETTER DA
0565	ե	ARMENIAN SMALL LETTER ECH
0566	զ	ARMENIAN SMALL LETTER ZA
0567	է	ARMENIAN SMALL LETTER EH
0568	ը	ARMENIAN SMALL LETTER ET
0569	թ	ARMENIAN SMALL LETTER TO
056A	ճ	ARMENIAN SMALL LETTER ZHE
056B	ի	ARMENIAN SMALL LETTER INI

056C	լ	ARMENIAN SMALL LETTER LIWN
056D	խ	ARMENIAN SMALL LETTER XEH
056E	վ	ARMENIAN SMALL LETTER CA
056F	կ	ARMENIAN SMALL LETTER KEN
0570	հ	ARMENIAN SMALL LETTER HO
0571	ձ	ARMENIAN SMALL LETTER JA
0572	ղ	ARMENIAN SMALL LETTER GHAD
0573	ճ	ARMENIAN SMALL LETTER CHEH
0574	ս	ARMENIAN SMALL LETTER MEN
0575	յ	ARMENIAN SMALL LETTER YI
0576	ն	ARMENIAN SMALL LETTER NOW
0577	շ	ARMENIAN SMALL LETTER SHA
0578	ո	ARMENIAN SMALL LETTER VO
0579	չ	ARMENIAN SMALL LETTER CHA
057A	պ	ARMENIAN SMALL LETTER PEH
057B	ջ	ARMENIAN SMALL LETTER JHEH
057C	ր	ARMENIAN SMALL LETTER RA
057D	ս	ARMENIAN SMALL LETTER SEH
057E	վ	ARMENIAN SMALL LETTER VEV
057F	տ	ARMENIAN SMALL LETTER TIWN
0580	ր	ARMENIAN SMALL LETTER REH
0581	ղ	ARMENIAN SMALL LETTER CO
0582	ւ	ARMENIAN SMALL LETTER YIWN
0583	փ	ARMENIAN SMALL LETTER PIWR
0584	ք	ARMENIAN SMALL LETTER KEH
0585	օ	ARMENIAN SMALL LETTER OH
0586	ֆ	ARMENIAN SMALL LETTER FEH
0587	և	ARMENIAN SMALL LIGATURE ECH YIWN

Punctuation

0589	։	ARMENIAN FULL STOP
058A	~	ARMENIAN HYPHEN

Religious symbols

058D	Ⲁ	RIGHT-FACING ARMENIAN ETERNITY SIGN
058E	ⲁ	LEFT-FACING ARMENIAN ETERNITY SIGN

Currency symbol

058F	֏	ARMENIAN DRAM SIGN
------	---	--------------------



Armenian MSR-2 table

- This defines a formal reference for the justification of the repertoire for this proposal.
- The following subset of code points identified as the Unicode range U+0561..U+0586 will be available for the root zone:

Selected code points

Code	Script	Name	Code	Script	Name
0561	ա	ARMENIAN SMALL LETTER AYB	0574	մ	ARMENIAN SMALL LETTER MEN
0562	բ	ARMENIAN SMALL LETTER BEN	0575	յ	ARMENIAN SMALL LETTER YI
0563	գ	ARMENIAN SMALL LETTER GIM	0576	ն	ARMENIAN SMALL LETTER NOW
0564	դ	ARMENIAN SMALL LETTER DA	0577	շ	ARMENIAN SMALL LETTER SHA
0565	ե	ARMENIAN SMALL LETTER ECH	0578	ո	ARMENIAN SMALL LETTER VO
0566	զ	ARMENIAN SMALL LETTER ZA	0579	չ	ARMENIAN SMALL LETTER CHA
0567	է	ARMENIAN SMALL LETTER EH	057A	պ	ARMENIAN SMALL LETTER PEH
0568	ր	ARMENIAN SMALL LETTER ET	057B	ղ	ARMENIAN SMALL LETTER JHEH
0569	թ	ARMENIAN SMALL LETTER TO	057C	ռ	ARMENIAN SMALL LETTER RA
056A	ժ	ARMENIAN SMALL LETTER ZHE	057D	ս	ARMENIAN SMALL LETTER SEH
056B	ի	ARMENIAN SMALL LETTER INI	057E	վ	ARMENIAN SMALL LETTER VEW
056C	լ	ARMENIAN SMALL LETTER LIWN	057F	տ	ARMENIAN SMALL LETTER TIWN
056D	խ	ARMENIAN SMALL LETTER XEH	0580	ր	ARMENIAN SMALL LETTER REH
056E	ծ	ARMENIAN SMALL LETTER CA	0581	ց	ARMENIAN SMALL LETTER CO
056F	կ	ARMENIAN SMALL LETTER KEN	0582	ւ	ARMENIAN SMALL LETTER YIWN
0570	հ	ARMENIAN SMALL LETTER HO	0583	փ	ARMENIAN SMALL LETTER PIWR
0571	ձ	ARMENIAN SMALL LETTER JA	0584	ք	ARMENIAN SMALL LETTER KEH
0572	ղ	ARMENIAN SMALL LETTER GHAD	0585	օ	ARMENIAN SMALL LETTER OH
0573	ճ	ARMENIAN SMALL LETTER CHEH	0586	ֆ	ARMENIAN SMALL LETTER FEH



Variants

There is some similarity between Armenian letters GIM and ZA but we considered that it is not enough to include them into variants. So our conclusion is:

The Armenian repertoire does not contain in-script variants.



Cross-script homoglyphs

- Cyrillic, Greek and Latin scripts are examples of related scripts where some cross-script homoglyphs exist.

Armenian-Cyrillic homoglyphs

Armenian Letter	Cyrillic Letter
հ U+0570 Armenian small letter HO	һ U+04BB Cyrillic small letter SHHA
օ U+0585 Armenian small letter OH	о U+043E Cyrillic small letter O

‘ն’ U+0578 Armenian small letter VO and ‘п’ U+43F Cyrillic small letter PE are usually somewhat distinct, unless both are in italic serif style, like ‘*п*’ vs. ‘*н*’, where they would be homoglyphs. However, they are not included into cross-script variants because italic font styles are insufficient justification for creating a cross-script variant, especially as that is not the most common style for presenting domain names to the user.



Armenian and Latin homoglyphs

Armenian Letter	Latin Letter
q U+0566 Armenian small letter ZA	q U+0071 Latin small letter Q
h U+0570 Armenian small letter HO	h U+0068 Latin small letter H
n U+0578 Armenian small letter VO	n U+006E Latin small letter N
u U+057D Armenian small letter SHE	u U+0075 Latin small letter U
g U+0581 Armenian small letter CO	g U+0067 Latin small letter G
Լ U+0582 Armenian small letter YIWN	ı U+0269 Latin small letter IOTA
o U+0585 Armenian small letter OH	o U+006F Latin small letter O



Armenian and Latin homoglyphs

The ARMENIAN SMALL LETTER ZA and the LATIN SMALL LETTER Q are not perfect homoglyphs but the difference is may not be perceivable at normal size. A label, such as .quo would be readily accepted by users as “the same” as the label .quo (in Latin).

The ‘g’ homoglyph situation only exists in sans-serif style, which, however, is a very common choice for user interfaces.

Armenian YIWN and Latin IOTA are considered homoglyphs as their visual perception is quite similar.

The ARMENIAN LETTER YIWN and j U+006A LATIN SMALL LETTER J are not considered homoglyphs as “dot” changes the visual perception of the letter. That is why this case is not included into the cross-script variants.

Armenian and Greek homoglyphs

Armenian Letter	Greek Letter
ղ U+0572 Armenian small letter GHAD	η U+03B7 Greek small letter ETA
լ U+0582 Armenian small letter YIWN	ι U+03B9 Greek small letter IOTA
օ U+0585 Armenian small letter OH	ο U+03BF Greek small letter OMICRON

Armenian YIWN and Greek IOTA as well as Armenian GHAD and Greek ETA are considered homoglyphs as their visual perception is quite similar. They are included in the cross-script variants.

The case of OH/OMICRON/Cyrillic O/Latin O is considered cross-script homoglyphs because of strong similarity.



Conclusions

As result of integration, blocked variants can exist between script-specific repertoires.

Even if a label is applied for in one script, the existence of certain other labels in another script may block it from being delegated .

For example, both .uno and .uno would be well-formed TLDs in Latin and Armenian,

If one of these two labels is applied for and allocated, the other would be blocked. Otherwise, both could be independently allocated to different applicants causing user confusion.

Armenian GP is hardly expecting any other application for an Armenian TLD but in cases when there is a homoglyph similarity in related scripts it should be blocked.

The Armenian GP has added out of repertoire blocked variants for all homoglyphs investigated above, except for cases identified as merely "similar".



Examples from the XML file

```
<data>
```

```
<char cp="0068" comment="Cross-script homoglyph">
```

```
  <var cp="0068" type="out-of-repertoire-var" comment="identity" />
```

```
  <var cp="04BB" type="blocked" comment="Homoglyph" />
```

```
  <var cp="0570" type="blocked" comment="Homoglyph" />
```

```
</char>
```

```
<char cp="0570" tag="sc:Armn" ref="0 100" comment="Armenian  
small letter HO">
```

```
  <var cp="0068" type="blocked" comment="Latin small letter H" />
```

```
  <var cp="04BB" type="blocked" comment="Cyrillic small letter  
SSHA" />
```

```
</char>
```



Second level domains

Visual similarities in the second level domains will be solved by registration rules. There will be restrictions in the policy of domain registration. The problem will be solved by limiting Armenian domain names strictly to the Armenian MSR table, Latin dash (code point '2d') and Latin numbers (code points '30' - '39')".

Visual similarities of related scripts will be blocked by the domain registration program as it will check the scripts for the correspondence to the Armenian MSR table and will not allow domains names with the visually similar code points of related scripts.



Enabling a Multilingual Internet: Public Comments open on Armenia Label Generation Rules proposal

<https://www.icann.org/public-comments/proposal-armenian-lgr-2015-07-22-en>

<https://www.icann.org/public-comments/proposal-armenian-lgr-22jul15-hy-2015-07-22-hy>

We encourage you to comment the document.



Many thanks for support to:

Michael Yakushev

Sarmad Hussain

Asmus Freytag

Michel Suignard

Yuri Kargopolov

Dusan Stojicevic

IDN Integration Panel



Thank you!