

Proposal for a Greek Script Root Zone Label Generation Ruleset (LGR)

Greek Generation Panel
LGR Version: 5
1 February 2022
Document Version: 6.7

Table of Contents

| | | |
|------------|--|-----------|
| 1 | <i>General Information / Overview / Abstract</i> | 3 |
| 2 | <i>Script for which the LGR is proposed</i> | 3 |
| 3 | <i>Background on Script and Principal Languages Using it</i> | 3 |
| 3.1 | The Pomak Language | 4 |
| 3.2 | The Greek Language – Useful key points | 4 |
| 3.2.1 | The Greek Language Question..... | 5 |
| 3.2.2 | The Greek Orthography | 6 |
| 4 | <i>Overall Development Process and Methodology</i> | 7 |
| 5 | <i>Repertoire</i> | 7 |
| 6 | <i>Variants</i> | 9 |
| 6.1 | In-script variants | 9 |
| 6.1.1 | Polytonic Domain Name Variants | 9 |
| 6.1.2 | Vowels with/without Tonos and/or Dialytika..... | 9 |
| 6.1.3 | The sigma and final sigma..... | 11 |
| 6.2 | Cross-script variants | 12 |
| 6.2.1 | Greek and Latin | 13 |
| 6.2.2 | Greek and Cyrillic | 17 |
| 6.2.3 | Greek and Armenian | 19 |
| 6.2.4 | The issue of uppercase Greek characters, similarity included by transitivity and cross-script variants | 21 |
| 6.2.5 | Cross-script variants – conclusions and recommendations | 25 |
| 6.3 | Summary of all variants | 28 |
| 7 | <i>Whole Label Evaluation Rules (WLE) and Actions</i> | 32 |
| 8 | <i>Contributors</i> | 33 |
| 9 | <i>References</i> | 34 |
| | <i>Appendix: Cross-script homoglyphs – possible confusability issues</i> | 36 |
| A.1 | Cross-script similarity between code-points included in MSR-4 | 36 |
| A.2 | Uppercase Similarity Table | 39 |
| A.3 | Effect of Transitivity on Uppercase Similarity..... | 41 |

Proposal for a Greek Script Root Zone Label Generation Rules (LGR)

1 General Information / Overview / Abstract

The purpose of this document is to give an overview of the proposed Greek LGR in the XML format and the rationale behind the design decisions taken. It includes a discussion of relevant features of the script, the process and methodology used and information on the contributors. The formal specification of the Greek LGR can be found in the accompanying XML document:

proposal-greek-lgr-01feb22-en.xml

Labels for testing can be found in the accompanying text document:

greek-test-labels-01feb22-en.txt

2 Script for which the LGR is proposed

ISO 15924 Code: Grek

ISO 15924 Key No: 200

ISO 15924 English Name: Greek

Property Value Alias: Greek

Native Name of the script: Ελληνικά

Maximal Starting Repertoire (MSR) version: MSR-5

3 Background on Script and Principal Languages Using it

The Greek script is used to write the Greek language, which is the official language of Greece and Cyprus. It is also used by Greeks who live in other countries. Significant Greek population can be found in the European Union Member-States, the United States of America, Australia, Canada, Albania, Turkey, Egypt, South Africa, Russia and other Member-States of the former Soviet Union etc.

Historically, the Greek script was used for some other dialects/forms/variants/languages (e.g. most recently, Karamanli Turkish¹, Albanian dialects in Greece (Arvanitika²), Ladino/Sephardic, a Spanish-Jewish dialect used in

¹ Source: Wikipedia, https://en.wikipedia.org/wiki/Karamanli_Turkish

² Source: Wikipedia, <https://en.wikipedia.org/wiki/Arvanitika>

| | | |
|---|--------|------------------------|
| Proposal for a Greek Script Root Zone LGR | 3 / 41 | Greek Generation Panel |
|---|--------|------------------------|

Greece, etc.³), however today that use has become obsolete⁴. Nowadays, the only “non-Greek” language that actually uses Greek characters is the Pomak language.

3.1 The Pomak Language

The Pomak language⁵ (Greek: πομακική γλώσσα, pomakiki glossa or πομακικά, pomakika, Bulgarian: помашки език, romaški ezik, Turkish: Pomakça) is a term used in Greece and Turkey to refer to some of the Rup dialects of the Bulgarian language spoken by the Pomaks in Western Thrace in Greece and Eastern Thrace in Turkey. These dialects are native also in Bulgaria and are classified as part of the Smolyan subdialect.

The Pomak language is a living language, used mainly in oral communication and is spoken by about 30,000 people in Greece today. Currently, in the formation of the literary standard, the Pomaks attempt to create script based on the Greek, Cyrillic or Latin characters.

The most important attempt towards the representation of the Pomak language with the use of Greek characters of the monotonic system took place with the publication of the “Pomak-Greek Dictionary” in 1996. As stated in its introduction, the intention was to achieve the greatest possible correspondence between the Greek and the Pomak alphabet in the way of pronouncing the sounds.

3.2 The Greek Language – Useful key points

Katharevousa⁶ (Greek: Καθαρεύουσα, [katha' revusa], lit. "puristic [language]"): It is a form of the Greek language conceived in the early 19th century as a compromise between Ancient Greek and the Modern Greek of the time, with a vocabulary largely based on ancient forms, but a much-simplified grammar. Originally, it was widely used both for literary and official purposes, though seldom in daily language. In the 20th century, it was increasingly used for official and formal purposes, until Dimotiki became the official language of Greece in 1976.

Dimotiki⁷ (Greek: δημοτική [γλώσσα] [ðimoti'ci], "[language] of the people"): It is the modern vernacular form of the Greek language. The term has been in use since 1818. Dimotiki refers particularly to the form of the language that evolved naturally from ancient Greek, in opposition to the artificially archaic Katharevousa, which was the official standard until 1976. The two complemented each other in a

³ Source: Wikipedia, https://en.wikipedia.org/wiki/Greek_alphabet#Use_for_other_languages

⁴ Source: "Character codes for Greek: Problems and modern solutions", Macrakis Stavros M., 1996

⁵ Source: Wikipedia, https://en.wikipedia.org/wiki/Pomak_language

⁶ Source: Wikipedia, <http://en.wikipedia.org/wiki/Katharevousa>

⁷ Source: Wikipedia, <http://en.wikipedia.org/wiki/Dimotiki>

typical example of "diglossia" until the resolution of the Greek language question in favour of Dimotiki.

3.2.1 The Greek Language Question⁸

The Greek language question was a dispute discussing the question whether the language of the Greek people (Dimotiki) or a cultivated imitation of Ancient Greek (Katharevousa) should be the official language of the Greek state. It was a highly controversial topic in the 19th and 20th centuries and was finally resolved in 1976, when Dimotiki was made the official language. The language phenomenon in question — which also occurs elsewhere in the world — is called “**diglossia**”, i.e. a word consisting of “di-” (from “dis”, twice) and “glossa” (language, in Greek), which is used for cases where two forms of one language (or two languages) are used under different conditions by the same community.

While Dimotiki was the mother tongue of the Greeks, Katharevousa was an archaic and formal variant that was pronounced like Modern Greek, but adopted both lexical and morphological features of ancient Greek that the spoken language had lost over time, like: morphological features, phonological features, syntactic features and lexical features. These differences meant that Katharevousa was only partly intelligible to a Greek without higher education. In addition, there was no single Katharevousa. Instead, proponents of the formal language utilized ever-changing variants that were never standardized. Some of these variants (the most extreme ones) were very close to the Attic variant of ancient Greek while others could be closer to spoken Greek and could be understood by a portion of the people.

Diglossia: For a person who does not speak Greek and whose mother tongue (e.g. English) exhibits no comparable form of diglossia, it is hard to understand the motivation of the Greek language question, as it concerns the coexistence of two — in extreme cases — completely different forms of Greek that greatly exceeds the usual stylistic difference between written and spoken language.

Katharevousa was established as the official language of Greece after the recognition of the country's independence in 1830, since the "unpolished" language of the people was not considered able to fit the needs of a modern state.

By 1900, the discussion about the form of the Greek language had become a matter of public interest, deeply dividing the people.

The educational system was in an alarming state and completely ineffective: the children were completely unable to express themselves in the unfamiliar formal language, which severely harmed their speech acquisition instead of educating them.

⁸ Source: Wikipedia, http://en.wikipedia.org/wiki/Greek_language_question

| | | |
|--|---------------|-------------------------------|
| <i>Proposal for a Greek Script Root Zone LGR</i> | <i>5 / 41</i> | <i>Greek Generation Panel</i> |
|--|---------------|-------------------------------|

By 1917 the Dimotiki had been successfully introduced into a limited number of primary schools; but again, it was repeatedly replaced with Katharevousa. Only on April 30, 1976 was the era of linguistic purism ended in Greece when the Greek Government banned Katharevousa from use in schools and, only a few months later, passed a law introducing the use of the Dimotiki in official texts and documents, which effectively terminated the diglossia. Ironically, the law in question was formulated in Katharevousa.

3.2.2 The Greek Orthography⁹

Greek orthography has used a variety of diacritics starting in the Hellenistic period. The complex polytonic (multi-accent) orthography notates Ancient Greek phonology. The simple monotonic (single-accent) orthography, introduced in 1982, corresponds to Modern Greek phonology and requires only two diacritics:

Tonos: Greek accent mark, acute accent (Greek Tonos, U+0384).

Dialytika (diaeresis): Greek accent mark (appears on the letters “ι̇” (e.g. Greek small letter iota with dialytika, U+03CA) and “υ̇” (e.g. Greek small letter upsilon with dialytika, U+03CB) to show that a pair of vowel letters is pronounced separately within a syllable, as a phonetic diphthong; rather than as an orthographic digraph (which would be realized as a monophthong phonetically). See, e.g., <https://el.wikipedia.org/wiki/Δίφθογγος>, and Greek: *An Essential Grammar of the Modern Language*, Holton, Mackridge, Philippaki, Routledge 2004, .¹⁰ It can also be combined with tonos over the same letters, Greek small letter iota with dialytika and tonos, U+0390 and Greek small letter upsilon with dialytika and tonos, U+03B0.

Polytonic orthography (πολύς, "many", τόνος "accent") is the standard system for Ancient Greek. The acute accent (´), the grave accent (`), and the circumflex (^) indicate different kinds of pitch accent. The rough breathing (ρ̣) indicates the presence of an /h/ sound before a letter, while the smooth breathing (ρ̣̄) indicates the absence of /h/.

Since in Modern Greek the pitch accent was replaced by a dynamic accent, and the /h/ was lost, most polytonic diacritics have no phonetic significance, and merely reveal the underlying Ancient Greek etymology.

Monotonic orthography (μόνος "single", τόνος "accent") is the standard system for Modern Greek. It retains only the acute accent (tonos) to indicate stress

⁹ Source: Wikipedia, http://en.wikipedia.org/wiki/Greek_diacritics

¹⁰ Source: Greek: *An Essential Grammar of the Modern Language*, Holton, Mackridge, Philippaki, Routledge 2004, §§ 1.1-1.3. Also, Wikipedia, <http://el.wikipedia.org/wiki/Δίφθογγος>. [Note that Greek diphthongs, whether orthographic or phonetic, are not described at all in the English wikipedia <http://en.wikipedia.org/wiki/Diphthong>.] .

| | | |
|---|--------|------------------------|
| Proposal for a Greek Script Root Zone LGR | 6 / 41 | Greek Generation Panel |
|---|--------|------------------------|

and the diaeresis (¨) to indicate a diphthong. Both the accent and the dialytika differentiate otherwise identical words in Modern Greek: compare the minimal pairs

- γέρος [ˈjeros], with the accent on the first syllable, meaning "old man" vs
- γερός [jeˈros], with the accent on the second syllable, meaning "strong"

and

- παῖδάκια [paɪˈðaca] "lamb chops", with a diphthong, and
- παιδάκια /peˈðaca/ "little children" with a simple vowel.

Tonos and dialytika can be combined on a single vowel.

4 Overall Development Process and Methodology

The process followed by the Panel was according to work plan given in “Proposal for the Generation Panel for the Greek Script Label Generation Ruleset for the Root Zone”.

5 Repertoire

Following discussions between the members of the Panel, the opinion mostly shared between the members is to disallow the ancient special characters and only allow characters that are today used in the spelling of Greek words. Furthermore, the group recommends allowing only the characters of the monotonic system.

The Panel reached this decision on the basis that TLD IDN registrations should be useful for as many people as possible. The use of ancient characters or characters of the polytonic system is not easy, even for people using Greek keyboards. Furthermore, the users of the Greek language are accustomed in the contemporary monotonic way of writing and this choice of characters will create a good use experience for the IDN TLDs in Greek. However, this recommendation stands only for the top-level labels and it does not prohibit registrations in characters of the polytonic system or ancient characters for the lower levels where some users could be expected to be accustomed to the use of these characters.

Regarding the Pomak language, due to the fact that the aim is to achieve the greatest possible correspondence between the Greek and the Pomak alphabet in the way of pronouncing the sounds, some additional characters introduced in the Pomak-Greek Dictionary, such as "SMALL ALPHA WITH DIALYTIKA" and "SMALL OMICRON WITH DIALYTIKA", which are not included in the Greek character set in Unicode, are used for better performance of the Pomak sounds, but as such, the Panel has concluded that they do not actually affect the process of the formation of domain

| | | |
|---|--------|------------------------|
| Proposal for a Greek Script Root Zone LGR | 7 / 41 | Greek Generation Panel |
|---|--------|------------------------|

names with Greek characters and therefore the available allowed set of Greek characters for use in the domain names.

Following the above, the code points recommended for inclusion in the root zone LGR for the Greek script are shown below:

Table 1: Greek LGR Repertoire

| # | Unicode Code Point | Glyph | Unicode Name |
|----|--------------------|-------|---|
| 1 | 0390 | ϊ | GREEK SMALL LETTER IOTA WITH DIALYTIKA AND TONOS |
| 2 | 03AC | ά | GREEK SMALL LETTER ALPHA WITH TONOS |
| 3 | 03AD | έ | GREEK SMALL LETTER EPSILON WITH TONOS |
| 4 | 03AE | ή | GREEK SMALL LETTER ETA WITH TONOS |
| 5 | 03AF | ί | GREEK SMALL LETTER IOTA WITH TONOS |
| 6 | 03B0 | ϋ | GREEK SMALL LETTER UPSILON WITH DIALYTIKA AND TONOS |
| 7 | 03B1 | α | GREEK SMALL LETTER ALPHA |
| 8 | 03B2 | β | GREEK SMALL LETTER BETA |
| 9 | 03B3 | γ | GREEK SMALL LETTER GAMMA |
| 10 | 03B4 | δ | GREEK SMALL LETTER DELTA |
| 11 | 03B5 | ε | GREEK SMALL LETTER EPSILON |
| 12 | 03B6 | ζ | GREEK SMALL LETTER ZETA |
| 13 | 03B7 | η | GREEK SMALL LETTER ETA |
| 14 | 03B8 | θ | GREEK SMALL LETTER THETA |
| 15 | 03B9 | ι | GREEK SMALL LETTER IOTA |
| 16 | 03BA | κ | GREEK SMALL LETTER KAPPA |
| 17 | 03BB | λ | GREEK SMALL LETTER LAMDA |
| 18 | 03BC | μ | GREEK SMALL LETTER MU |
| 19 | 03BD | ν | GREEK SMALL LETTER NU |
| 20 | 03BE | ξ | GREEK SMALL LETTER XI |
| 21 | 03BF | ο | GREEK SMALL LETTER OMICRON |
| 22 | 03C0 | π | GREEK SMALL LETTER PI |

| # | Unicode Code Point | Glyph | Unicode Name |
|----|--------------------|-------|---|
| 23 | 03C1 | ρ | GREEK SMALL LETTER RHO |
| 24 | 03C2 | ς | GREEK SMALL LETTER FINAL SIGMA |
| 25 | 03C3 | σ | GREEK SMALL LETTER SIGMA |
| 26 | 03C4 | τ | GREEK SMALL LETTER TAU |
| 27 | 03C5 | υ | GREEK SMALL LETTER UPSILON |
| 28 | 03C6 | φ | GREEK SMALL LETTER PHI |
| 29 | 03C7 | χ | GREEK SMALL LETTER CHI |
| 30 | 03C8 | ψ | GREEK SMALL LETTER PSI |
| 31 | 03C9 | ω | GREEK SMALL LETTER OMEGA |
| 32 | 03CA | ϊ | GREEK SMALL LETTER IOTA WITH DIALYTIKA |
| 33 | 03CB | ϋ | GREEK SMALL LETTER UPSILON WITH DIALYTIKA |
| 34 | 03CC | ό | GREEK SMALL LETTER OMICRON WITH TONOS |
| 35 | 03CD | ύ | GREEK SMALL LETTER UPSILON WITH TONOS |
| 36 | 03CE | ώ | GREEK SMALL LETTER OMEGA WITH TONOS |

6 Variants

6.1 In-script variants

6.1.1 Polytonic Domain Name Variants

As stated before, the conclusion of the Panel is that the use of polytonic characters for TLD registration presents no significant advantage for the user and the use of only monotonic characters is preferable. However, this might not be the case for second-level registrations and the registries should be encouraged to decide upon the use of polytonic characters in their policy. Certain user groups (e.g. the Greek Orthodox Church, scholars etc.) might have interest in registering and using domain names in polytonic characters in the second and lower levels.

6.1.2 Vowels with/without Tonos and/or Dialytika

Almost all Greek words when written in lower case (excluding only monosyllabic words) must have the tonos over one of their vowels to be spelled

correctly. This way, as already shown in paragraph 3.1.2, the person using this word is able to identify the correct pronunciation and, often, the correct meaning of the word.

Regarding domain names, a domain name with the tonos sign over a vowel is a different domain name than the one without the tonos due to the different encoding of the specific code points used. However, the domain name without and with tonos can be perceived as the ‘same’ domain name.

Thus, it is considered by the Panel that vowels with and without tonos and/or dialytika should be handled as “in-script variants” respectively. Table 2 shows these variants arranged by their respective non-accented vowel:

Table 2: Non-Accented Vowels and Respective Accented Vowels

| Non-accented vowel | Vowel with tonos | Vowel with dialytika | Vowel with tonos and dialytika |
|--------------------|------------------|----------------------|--------------------------------|
| α U+03B1 | ά U+03AC | | |
| ε U+03B5 | έ U+03AD | | |
| η U+03B7 | ή U+03AE | | |
| ι U+03B9 | ί U+03AF | ϊ U+03CA | ϊ̇ U+0390 |
| ο U+03BF | ό U+03CC | | |
| υ U+03C5 | ύ U+03CD | ϋ U+03CB | ϋ̇ U+03B0 |
| ω U+03C9 | ώ U+03CE | | |

The Panel took under consideration and examined carefully IP’s recommendations as stated in “IP Review of Variants for the Proposed Greek LGR dated 2019-09-01”, dated 2019-12-04, where it was recommended, as far as possible, not to introduce variant sets made up of a base character and the same base character with a tonos and/or dialytika, stating that this introduces complexity in the root zone as may create undesirable in-script variants in other scripts.

The Panel, being aware of the fact that since 2005 the Greek speaking Internet community has been able to register domain names in Greek at the second level in the [.gr] domain, has concluded that despite the relatively small number of these domain names (in relation to Internationalized Domain Names in other scripts), the current rules have formed the respective user experience at a level that cannot be changed. The Panel has also taken into consideration that these rules have been

updated recently¹¹ and give the users the choice to also register Greek domain names in the Greek TLD [. $\epsilon\lambda$].

In order to minimize the number of allocatable variant labels as per the LGR procedure, the GP decides to allow for any applied-for label and its non-accented form as an allocatable variant label. All the other forms with arbitrary mixtures of accented and non-accented code points would be blocked.

The GP recommends that both forms of the TLD labels should be allocated as part of a single application for usability reasons and to maintain the experience of Greek community of Internet users.

In the light of the above, the Panel recommends that the code points mentioned at the table above should be considered as *“in-script variants”*. *The variant type from an accented code point to a non-accented code point is “allocatable” while the variant type from a non-accented code points to an accented code point is “blocked”*.

6.1.3 The sigma and final sigma

‘Sigma’ is the 18th character of the Greek alphabet, corresponding to /s/. There are three types of sigma: ‘Lower Case Sigma’ (Unicode U+03C3), ‘Upper Case Sigma’ (Unicode U+03A3) and small letter ‘Final Sigma’ (Unicode U+03C2). The use of the first two is obvious. The small letter Final Sigma is used only at the end of words, instead of the lowercase sigma.

In IDNA2003, the three letters were mapped to one another and the small letter final sigma was replaced by the IDNA2003 processor by the letter small sigma. When a final sigma was used, the translation of a word from the Punycode was not satisfactory because the Nameprep translation had converted it to a small letter sigma.

In IDNA2008 a different approach was followed. The following paragraphs¹² explain the rationale behind the decision:

“In IDNA2003, all characters are case folded and mapped by clients in a standardized step.

¹¹ Decision 843/2/01 March 2018 of the Hellenic Telecommunications and Post Commission: “Regulation on Management and Assignment of [. gr] or [. $\epsilon\lambda$] Domain Names

¹² Source: IETF, RFC5894, “Internationalized Domain Names for Applications (IDNA) Background, Explanation, and Rationale”, August 2010

Even in scripts that generally support case distinctions, some characters do not have uppercase forms. For example, the Unicode case-folding operation maps Greek Final Form Sigma (U+03C2) to the medial form (U+03C3) and maps Eszett (German Sharp S, U+00DF) to “ss”. Neither of these mappings is reversible because the uppercase of U+03C3 is the uppercase Sigma (U+03A3) and “ss” is an ASCII string. IDNA2008 permits, at the risk of some incompatibility, slightly more flexibility in this area by avoiding case folding and treating these characters as themselves. Approaches to handling one-way mappings are discussed in Section 7.2.

Because IDNA2003 maps Final Sigma and Eszett to other characters, and the reverse mapping is never possible, neither Final Sigma nor Eszett can be represented in the ACE form of IDNA2003 IDN nor in the native character (U-label) form derived from it. With IDNA2008, both characters can be used in an IDN and so the A-label used for lookup for any U-label containing those characters is now different. See Section 7.1 for a discussion of what kinds of changes might require the IDNA prefix to change; after extended discussions, the IDNABIS Working Group came to consensus that the change for these characters did not justify a prefix change.”

Taking into consideration the above and given the fact that Upper-Case letters are not part of IDNs in IDNA2008, **the Panel recommends that the ‘Lower Case Sigma’ (U+03C3) and the ‘Final Sigma’ (U+03C2) should be treated as “in-script variants”. The variant type from the Final Sigma to the Lower-Case Sigma is “allocatable” while the variant type from the Lower Case Sigma to the Final Sigma is “blocked”.**

If these two variant scenarios in 6.1.2 and 6.1.3 (accents and sigma) are in the same label, the maximal number of allocatable variants per original label is four.

The GP recommends that all of the allocatable TLD labels should be allocated as part of a single application for usability reasons and to maintain the experience of Greek community of Internet users.

6.2 Cross-script variants

Greek, Latin and Cyrillic scripts are examples of related scripts where some cross-script homoglyphs exist. In addition, the Panel is of the opinion that there are also other cases of homoglyphs – identical (i) or merely similar (s) – between Greek and other scripts included in the MSR-5, like Armenian for example. In order to conclude whether a code point of a given MSR-5 script and a Greek code point (small-case only) are considered as variants, the Panel examined carefully each pair of the code points under question, using the same font and the same size for each pair, where possible.

The Panel also examined the case of the application of the underline style, which is commonly done to highlight hyperlinks. As the proposal is for the root-zone, the proposal does not include the underlining cases into the variant consideration. These can be listed as similar cases.

Finally, the Panel took into account the user experience of the Greek Internet users and the “Regulation on Management and Assignment of .GR Domain Names”, which initially introduced Greek at the second level domain in 2005 and deals with it without any significant problems since then. In the light of the above, the findings of the Panel’s analysis per script are shown below:

6.2.1 Greek and Latin

Table 3: Greek and Latin Cross-Script Variant Analysis

| Greek Letter | Latin Letter | i / s |
|--|---|---------------------------------------|
| U+03B1 α Greek small letter ALPHA | a U+0061 LATIN small letter A ɑ U+0251 Latin small letter ALPHA Ɽ U+0061+U+0331 LATIN small letter A + COMBINING MACRON BELOW | i s |
| U+03B2 β Greek small letter BETA | ß U+00DF Latin small letter SHARP S | i |
| U+03B3 γ Greek small letter GAMMA | γ U+0079 Latin small letter Y Ƴ U+0263 Latin small letter GAMMA | i i |
| U+03B4 δ Greek small letter DELTA | | |
| U+03B5 ε Greek small letter EPSILON | ε U+025B Latin small letter OPEN E | i |
| U+03B6 ζ Greek small letter ZETA | | |
| U+03B7 η Greek small letter ETA | n U+006E Latin small letter N ñ U+0146 Latin small letter N WITH CEDILLA ŋ U+014B Latin small letter ENG Ɱ U+1E49 Latin small letter N WITH LINE BELOW Ɐ U+1E4B Latin small letter N WITH CIRCUMFLEX BELOW | i s i s s |
| U+03B8 θ Greek small letter THETA | ø U+0275 Latin small letter BARRED O | s |

| Greek Letter | Latin Letter | i / s |
|--|---|---------------------------------------|
| U+03B9 ι Greek small letter IOTA | ι U+0269 Latin small letter IOTA | i |
| U+03BA κ Greek small letter KAPPA | k U+006B Latin small letter K ķ U+0199 Latin small letter K WITH HOOK | s s |
| U+03BB λ Greek small letter LAMDA | | |
| U+03BC μ Greek small letter MU | | |
| U+03BD ν Greek small letter NU | v U+0076 Latin small letter V ʋ U+1E7F Latin small letter V WITH DOT BELOW | i s |
| U+03BE ξ Greek small letter XI | | |
| U+03BF ο Greek small letter OMICRON | o U+006F Latin small letter O ø U+006F+U+0331 Latin small letter O + COMBINING MACRON BELOW ȯ U+1ECD Latin small letter O WITH DOT BELOW | i s s |
| U+03C0 π Greek small letter PI | | |
| U+03C1 ρ Greek small letter RHO | ρ U+0070 Latin small letter P | s |
| U+03C2 ς Greek small letter FINAL SIGMA | s U+0073 Latin small letter S ç U+00E7 Latin small letter C WITH CEDILLA ș U+015F Latin small letter S WITH CEDILLA ŝ U+0219 Latin small letter S WITH COMMA BELOW ſ U+1E63 Latin small letter S WITH DOT BELOW | s s s s s |
| U+03C3 σ Greek small letter SIGMA | σ U+01A1 Latin small letter O WITH HORN ø U+1EE3 Latin small letter O WITH HORN AND DOT BELOW | s s |
| U+03C4 τ Greek small letter TAU | | |

| Greek Letter | Latin Letter | i / s |
|---|--|--|
| U+03C5 ς Greek small letter UPSILON | u U+0075 Latin small letter U υ U+028B Latin small letter V WITH HOOK ϋ U+1EE5 Latin small letter U WITH DOT BELOW | i s |
| U+03C6 ϕ Greek small letter PHI | | |
| U+03C7 χ Greek small letter CHI | x U+0078 Latin small letter X | s |
| U+03C8 ψ Greek small letter PSI | | |
| U+03C9 ω Greek small letter OMEGA | w U+0077 Latin small letter W | s |
| U+03AC ᾱ Greek small letter ALPHA WITH TONOS | á U+00E1 Latin small letter A WITH ACUTE | i |
| U+03AD ἑ Greek small letter EPSILON WITH TONOS | | |
| U+03AE ῆ Greek small letter ETA WITH TONOS | ń U+0144 Latin small letter N WITH ACUTE ñ U+01F9 Latin small letter N WITH GRAVE ň U+1E45 Latin small letter N WITH DOT ABOVE | i s i |
| U+03AF ῖ Greek small letter IOTA WITH TONOS | i U+0069 Latin small letter I ï U+1ECB Latin small letter I WITH DOT BELOW ĭ U+0069+U+0331 Latin small letter I + COMBINING MACRON BELOW ì U+00EC Latin small letter I WITH GRAVE í U+00ED Latin small letter I WITH ACUTE ï U+1EC9 Latin small letter I WITH HOOK ABOVE | i s s s i s |

| Greek Letter | Latin Letter | i / s |
|---|---|--|
| U+03CC Greek small letter OMICRON WITH TONOS | ò U+00F2 Latin small letter O WITH GRAVE ò U+1ECD+U+0300 Latin small letter O WITH DOT BELOW + COMBINING GRAVE ACCENT ó U+00F3 Latin small letter O WITH ACUTE ó U+1ECD+U+0301 Latin small letter O WITH DOT BELOW + COMBINING ACUTE ACCENT ò U+022F Latin small letter O WITH DOT ABOVE ò U+1ECF Latin small letter O WITH HOOK ABOVE | s s i s i s |
| U+03CD Greek small letter UPSILON WITH TONOS | ù U+00F9 Latin small letter U WITH GRAVE ú U+00FA Latin small letter U WITH ACUTE ů U+1EE7 Latin small letter U WITH HOOK ABOVE ú U+1EE9 Latin small letter U WITH HORN AND ACUTE ù U+1EEB Latin small letter U WITH HORN AND GRAVE | s i s s s |
| U+03CE Greek small letter OMEGA WITH TONOS | ń U+1E83 Latin small letter W WITH ACUTE ŵ U+1E87 Latin small letter W WITH DOT ABOVE | s s |
| U+03CA Greek small letter IOTA WITH DIALYTIKA | ï U+00EF Latin small letter I WITH DIAERESIS | i |
| U+03CB Greek small letter UPSILON WITH DIALYTIKA | ü U+00FC Latin small letter U WITH DIAERESIS ű U+0171 Latin small letter U WITH DOUBLE ACCUTE | i s |
| U+0390 Greek small letter IOTA WITH DIALYTIKA AND TONOS | | |
| U+03B0 Greek small letter UPSILON WITH DIALYTIKA AND TONOS | ű U+01D8 Latin small letter U WITH DIAERESIS AND ACUTE ű U+01D8 Latin small letter U WITH DIAERESIS AND GRAVE | i s |

6.2.2 Greek and Cyrillic

Table 4: Greek and Cyrillic Cross-Script Variant Analysis

| Greek Letter | | Cyrillic Letter | | i / s |
|--|------------------------------------|-----------------|--|-------|
| U+03B1 | α Greek small letter ALPHA | a | U+0430 Cyrillic small letter A | i |
| U+03B2 | β Greek small letter BETA | в | U+0432 Cyrillic small letter VE | s |
| U+03B3 | γ Greek small letter GAMMA | γ | U+0443 Cyrillic small letter U | i |
| | | γ | U+04AF Cyrillic small letter STRAIGHT U | i |
| | | ϣ | U+04B1 Cyrillic small letter STRAIGHT U WITH STROKE | s |
| U+03B4 | δ Greek small letter DELTA | | | |
| U+03B5 | ε Greek small letter EPSILON | є | U+0454 Cyrillic small letter UKRAINIAN IE | s |
| U+03B6 | ζ Greek small letter ZETA | | | |
| U+03B7 | η Greek small letter ETA | | | |
| U+03B8 | θ Greek small letter THETA | ө | U+04E9 Cyrillic small letter BARRED O | s |
| U+03B9 | ι Greek small letter IOTA | | | |
| U+03BA | κ Greek small letter KAPPA | κ | U+043A Cyrillic small letter KA | i |
| | | κ | U+049B Cyrillic small letter KA WITH DESCENDER | s |
| | | κ | U+04A1 Cyrillic small letter BASHKIR KA | s |
| U+03BB | λ Greek small letter LAMDA | | | |
| U+03BC | μ Greek small letter MU | | | |
| U+03BD | ν Greek small letter NU | ϣ | U+04B1 Cyrillic small letter STRAIGHT U WITH STROKE | s |
| U+03BE | ξ Greek small letter XI | | | |
| U+03BF | ο Greek small letter OMICRON | о | U+043E Cyrillic small letter O | i |
| U+03C0 | π Greek small letter PI | π | U+043F Cyrillic small letter PE | s |
| | | π | U+0525 Cyrillic small letter PE WITH DESCENDER | s |
| <i>Proposal for a Greek Script Root Zone LGR</i> | | <i>17 / 41</i> | <i>Greek Generation Panel</i> | |

| Greek Letter | Cyrillic Letter | i / s |
|--|--|--------------|
| U+03C1 ρ Greek small letter RHO | ρ U+0440 Cyrillic small letter ER | s |
| U+03C2 ζ Greek small letter FINAL SIGMA | ς U+0455 Cyrillic small letter DZE ς U+04AB Cyrillic small letter ES WITH DESCENDER | s s |
| U+03C3 σ Greek small letter SIGMA | | |
| U+03C4 τ Greek small letter TAU | τ U+0442 Cyrillic small letter TE τ U+04AD Cyrillic small letter TE WITH DESCENDER | i s |
| U+03C5 υ Greek small letter UPSILON | | |
| U+03C6 φ Greek small letter PHI | φ U+0444 Cyrillic small letter EF | i |
| U+03C7 χ Greek small letter CHI | х U+0445 Cyrillic small letter HA χ U+04B3 Cyrillic small letter HA WITH DESCENDER | s s |
| U+03C8 ψ Greek small letter PSI | | |
| U+03C9 ω Greek small letter OMEGA | | |
| U+03AC á Greek small letter ALPHA WITH TONOS | | |
| U+03AD é Greek small letter EPSILON WITH TONOS | | |
| U+03AE ή Greek small letter ETA WITH TONOS | | |
| U+03AF í Greek small letter IOTA WITH TONOS | і U+0456 Cyrillic small letter BYELORUSSIAN-UKRAINIAN I | i |
| U+03CC ó Greek small letter OMICRON WITH TONOS | | |
| U+03CD ú Greek small letter UPSILON WITH TONOS | | |
| U+03CE ώ Greek small letter OMEGA WITH TONOS | | |

| Greek Letter | Cyrillic Letter | i / s |
|---|---|--------------|
| U+03CA ĭ Greek small letter IOTA WITH DIALYTIKA | ĭ U+0457 Cyrillic small letter YI | i |
| U+03CB ü Greek small letter UPSILON WITH DIALYTIKA | | |
| U+0390 ĭ̄ Greek small letter IOTA WITH DIALYTIKA AND TONOS | | |
| U+03B0 ǖ Greek small letter UPSILON WITH DIALYTIKA AND TONOS | | |

6.2.3 Greek and Armenian

Table 5: Greek and Armenian Cross-Script Variant Analysis

| Greek Letter | Armenian Letter | |
|---|---|---------------------|
| U+03B1 α Greek small letter ALPHA | | |
| U+03B2 β Greek small letter BETA | | |
| U+03B3 γ Greek small letter GAMMA | | |
| U+03B4 δ Greek small letter DELTA | | |
| U+03B5 ε Greek small letter EPSILON | | |
| U+03B6 ζ Greek small letter ZETA | | |
| U+03B7 η Greek small letter ETA | ղ U+0572 Armenian small letter GHAD ն U+0578 Armenian small letter VO դ U+0564 Armenian small letter DA | i i s |
| U+03B8 θ Greek small letter THETA | | |
| U+03B9 ι Greek small letter IOTA | Լ U+0582 Armenian small letter YIWN | i |
| U+03BA κ Greek small letter KAPPA | | |
| U+03BB λ Greek small letter LAMDA | | |

| Greek Letter | Armenian Letter | |
|---|--|---|
| U+03BC μ Greek small letter MU | | |
| U+03BD ν Greek small letter NU | | |
| U+03BE ξ Greek small letter XI | | |
| U+03BF ο Greek small letter OMICRON | ο U+0585 Armenian small letter OH | i |
| U+03C0 π Greek small letter PI | | |
| U+03C1 ρ Greek small letter RHO | | |
| U+03C2 ς Greek small letter FINAL SIGMA | | |
| U+03C3 σ Greek small letter SIGMA | | |
| U+03C4 τ Greek small letter TAU | | |
| U+03C5 υ Greek small letter UPSILON | u U+057D Armenian small letter SEH | i |
| U+03C6 φ Greek small letter PHI | | |
| U+03C7 χ Greek small letter CHI | | |
| U+03C8 ψ Greek small letter PSI | | |
| U+03C9 ω Greek small letter OMEGA | | |
| U+03AC ἄ Greek small letter ALPHA WITH TONOS | | |
| U+03AD ἕ Greek small letter EPSILON WITH TONOS | | |
| U+03AE ἦ Greek small letter ETA WITH TONOS | | |
| U+03AF ἱ Greek small letter IOTA WITH TONOS | | |
| U+03CC ὀ Greek small letter OMICRON WITH TONOS | | |

| Greek Letter | Armenian Letter | |
|---|-----------------|--|
| U+03CD $\acute{\upsilon}$ Greek small letter UPSILON WITH TONOS | | |
| U+03CE $\acute{\omega}$ Greek small letter OMEGA WITH TONOS | | |
| U+03CA $\grave{\iota}$ Greek small letter IOTA WITH DIALYTIKA | | |
| U+03CB $\ddot{\upsilon}$ Greek small letter UPSILON WITH DIALYTIKA | | |
| U+0390 $\grave{\iota}$ Greek small letter IOTA WITH DIALYTIKA AND TONOS | | |
| U+03B0 $\ddot{\upsilon}$ Greek small letter UPSILON WITH DIALYTIKA AND TONOS | | |

6.2.4 The issue of uppercase Greek characters, similarity included by transitivity and cross-script variants

According to the IDNA2008 protocol for PVALID characters, all Greek capital letters are disallowed. However, the applications still allow URLs in uppercase Greek letters if they are translated to lowercase. This introduces the potential hazard of confusable uppercase labels, e.g. “NIKE” and “ΝΙΚΕ” in Latin and even some labels in the Cyrillic alphabet. In lowercase characters, these labels differ considerably (in our example “nike” and “νικε”) and they do not present confusability issues but in uppercase form they are identical, even for a trained eye. In theory, the application should help the user differentiate such cases (e.g. by translating the Greek label to small letters) but this is not always easy or straightforward, for example a link in uppercase could be presented to the user with intention to mislead.

As a panel we thought of proposing the inclusion of the uppercase variant characters in the XML method, to be considered as cross-script variants and become allocatable only to the same entity. On second thought though, this entity would most of the times ask for the allocation of such an alternate domain name only in order to protect itself from malicious acts and since a cost is associated with each allocation, we suggest the Similarity Panel should deal with the exclusion of such variant domain names (block) to avoid imposing unnecessary economic burden to the registrant. If two labels consisting only of confusingly similar characters according to the following

table are to be allocated for any reason, we suggest they should be allocated to the same entity only, to avoid confusion by applying special measures for the second level labels.

Table 6: Uppercase Similarity Table

| Uppercase Similarity Table | | |
|--|------------------------------------|---|
| Greek Letter | Latin Letter | Cyrillic Letter |
| U+0391 A GREEK CAPITAL LETTER ALPHA | A U+0041 LATIN CAPITAL LETTER A | A U+0410 CYRILLIC CAPITAL LETTER A |
| U+0392 B GREEK CAPITAL LETTER BETA | B U+0042 LATIN CAPITAL LETTER B | B U+0412 CYRILLIC CAPITAL LETTER VE |
| U+0393 Γ GREEK CAPITAL LETTER GAMMA | | Г U+0413 CYRILLIC CAPITAL LETTER GHE |
| U+0394 Δ GREEK CAPITAL LETTER DELTA | | |
| U+0395 E GREEK CAPITAL LETTER EPSILON | E U+0045 LATIN CAPITAL LETTER E | E U+0415 CYRILLIC CAPITAL LETTER IE |
| U+0396 Ζ GREEK CAPITAL LETTER ZETA | Z U+005A LATIN CAPITAL LETTER Z | |
| U+0397 Η GREEK CAPITAL LETTER ETA | H U+0048 LATIN CAPITAL LETTER H | H U+041D CYRILLIC CAPITAL LETTER EN |
| U+0398 Θ GREEK CAPITAL LETTER THETA | | |
| U+0399 Ι GREEK CAPITAL LETTER IOTA | I U+0049 LATIN CAPITAL LETTER I | |
| U+039A Κ GREEK CAPITAL LETTER KAPPA | K U+004B LATIN CAPITAL LETTER K | K U+041A CYRILLIC CAPITAL LETTER KA |
| U+039B Λ GREEK CAPITAL LETTER LAMDA | | |
| U+039C Μ GREEK CAPITAL LETTER MU | M U+004D LATIN CAPITAL LETTER M | M U+041C CYRILLIC CAPITAL LETTER EM |
| U+039D Ν GREEK CAPITAL LETTER NU | N U+004E LATIN CAPITAL LETTER N | |

| | | |
|--|----------------|-------------------------------|
| <i>Proposal for a Greek Script Root Zone LGR</i> | <i>22 / 41</i> | <i>Greek Generation Panel</i> |
|--|----------------|-------------------------------|

| Uppercase Similarity Table | | |
|--|---|--|
| Greek Letter | Latin Letter | Cyrillic Letter |
| U+039E Ξ GREEK CAPITAL LETTER XI | | |
| U+039F Ο GREEK CAPITAL LETTER OMICRON | O U+004F LATIN CAPITAL LETTER O | О U+041E CYRILLIC CAPITAL LETTER O |
| U+03A0 Π GREEK CAPITAL LETTER PI | | П U+041F CYRILLIC CAPITAL LETTER PE |
| U+03A1 Ρ GREEK CAPITAL LETTER RHO | P U+0050 LATIN CAPITAL LETTER P | Р U+0420 CYRILLIC CAPITAL LETTER ER |
| U+03A3 Σ GREEK CAPITAL LETTER SIGMA | | |
| U+03A4 Τ GREEK CAPITAL LETTER TAU | T U+0054 LATIN CAPITAL LETTER T | Т U+0422 CYRILLIC CAPITAL LETTER TE |
| U+03A5 Υ GREEK CAPITAL LETTER UPSILON | Y U+0059 LATIN CAPITAL LETTER Y | |
| U+03A6 Φ GREEK CAPITAL LETTER PHI | | Ф U+0424 CYRILLIC CAPITAL LETTER EF |
| U+03A7 Χ GREEK CAPITAL LETTER CHI | X U+0058 LATIN CAPITAL LETTER X | Х U+0425 CYRILLIC CAPITAL LETTER HA |
| U+03A8 Ψ GREEK CAPITAL LETTER PSI | | |
| U+03A9 Ω GREEK CAPITAL LETTER OMEGA | | |

As a panel, we feel obliged to present an even trickier issue, regarding the letters in the following table. The colored fields are identical to each other, depending on the case of the letter. However, transitivity would also imply similarity across all the same-colored letters:

| Table 7: Transitivity of Uppercase Similarity | | | | | |
|---|--|----------------------------------|------------------------------------|--------------------------------------|--|
| Greek small Letter | Greek Capital Letter | Latin small Letter | Latin Capital Letter | Cyrillic small Letter | Cyrillic Capital Letter |
| U+03B7 η GREEK SMALL LETTER ETA | U+0397 Η GREEK CAPITAL LETTER ETA | U+006E n LATIN SMALL LETTER N | U+004E N LATIN CAPITAL LETTER N | | |
| U+03B7 η GREEK SMALL LETTER ETA | U+0397 Η GREEK CAPITAL LETTER ETA | U+0068 h LATIN SMALL LETTER H | U+0048 H LATIN CAPITAL LETTER H | U+043D Η CYRILLIC SMALL LETTER EN | U+041D Η CYRILLIC CAPITAL LETTER EN |
| U+03BD ν GREEK SMALL LETTER NU | U+039D Ν GREEK CAPITAL LETTER NU | U+0076 v LATIN SMALL LETTER V | U+0056 V LATIN CAPITAL LETTER V | | |
| U+03C5 υ GREEK SMALL LETTER UPSILON | U+03A5 Υ GREEK CAPITAL LETTER UPSILON | U+0075 u LATIN SMALL LETTER U | U+0055 U LATIN CAPITAL LETTER U | | |
| U+03C5 υ GREEK SMALL LETTER UPSILON | U+03A5 Υ GREEK CAPITAL LETTER UPSILON | U+0079 y LATIN SMALL LETTER Y | U+0059 Y LATIN CAPITAL LETTER Y | U+0443 y CYRILLIC SMALL LETTER U | |
| | | U+0079 y LATIN SMALL LETTER Y | U+0059 Y LATIN CAPITAL LETTER Y | U+0443 y CYRILLIC SMALL LETTER U | U+0423 Y CYRILLIC CAPITAL LETTER U |

A simplified form of the table above which clearly shows the effect of transitivity on similarities of quite different letters:

Table 8: Transitivity of Uppercase Similarity Simplified

| Greek small Letter | Greek Capital Letter | Latin small Letter | Latin Capital Letter | Cyrillic small Letter | Cyrillic Capital Letter |
|--------------------|----------------------|--------------------|----------------------|-----------------------|-------------------------|
| H | H | n | N | | |
| H | H | h | H | Η | Η |
| N | N | v | V | | |
| Y | Υ | u | U | | |
| Y | Υ | y | Y | y | |
| | | y | Y | y | Y |

This type of similarity has to be taken into account in combination with the similarity of uppercase characters before allowing the allocation of a label in these three scripts. If two labels consisting only of confusingly similar characters according to the tables above are to be allocated for any reason, the Panel suggests they should be allocated to the same entity only, to avoid confusion by requesting the registry to apply special anti-similarity measures for the second level labels.

6.2.5 Cross-script variants – conclusions and recommendations

Following the analysis shown in previous paragraphs regarding the relationship between the code points of the Greek script and the “corresponding” code-points of the Latin, Cyrillic and Armenian scripts, the Panel concludes that there are cases of “identical” code points and cases of “similar” code points. Taking into account that the Greek script is much older than the other ones and that the Greek civilization has influenced all surrounding populations since the ancient times, the Panel considers this as logical and unavoidable.

For the purpose of this Proposal, only the “identical” code points will be considered as “cross-script variants” and recommended to be marked as “blocked”. However, the “similar” code points will be presented in the Appendix in order to be taken into account should any confusability issues might arise. It is expected that all scripts’ code points – including Latin ASCII ones – will be treated equally and that all scripts will enjoy the same level of protection as far as confusability issues are to be concerned.

In the light of the above, **the Panel recommends that the code points mentioned at the table below should be considered as “cross-script variants”.**

Table 9: Cross-Script Variant Code Points

| Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|--------------------------------------|---|-------------------------------------|-----------------|
| U+03B1 α Greek small letter ALPHA | a U+0061 LATIN small letter A α U+0251* Latin small letter ALPHA | a U+0430 Cyrillic small letter A | |

| Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|--|--|---|--|
| U+03B2 β Greek small letter BETA | ß U+00DF Latin small letter SHARP S ss U+0073 U+0073 Latin small letter S Latin small letter S (Variant defined by Latin GP) | ss U+0455 U+0455 Cyrillic small letter Dze Cyrillic small letter Dze (Variant defined by Latin GP) | |
| U+03B3 γ Greek small letter GAMMA | γ U+0079 Latin small letter Y Υ U+0263 Latin small letter GAMMA | γ U+0443 Cyrillic small letter U Υ U+04AF Cyrillic small letter STRAIGHT U | |
| U+03B5 ε Greek small letter EPSILON | ε U+025B Latin small letter OPEN E | | |
| U+03B7 η Greek small letter ETA | η U+006E Latin small letter N η U+014B Latin small letter ENG | | ղ U+0572 Armenian small letter GHAD ն U+0578 Armenian small letter VO |
| U+03B9 ι Greek small letter IOTA | ι U+0269 Latin small letter IOTA | | Լ U+0582 Armenian small letter YIWN |
| U+03BA κ Greek small letter KAPPA | | κ U+043A Cyrillic small letter KA | |
| U+03BD ν Greek small letter NU | ν U+0076 Latin small letter V | | |
| U+03BF ο Greek small letter OMICRON | ο U+006F Latin small letter O | ο U+043E Cyrillic small letter O | օ U+0585 Armenian small letter OH |
| U+03C4 τ Greek small letter TAU | | т U+0442 Cyrillic small letter TE | |
| <i>Proposal for a Greek Script Root Zone LGR</i> | | 26 / 41 | <i>Greek Generation Panel</i> |

| Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|--|---|---|--|
| U+03C5 ϰ Greek small letter UPSILON | u U+0075 Latin small letter U ϰ U+028B Latin small letter V WITH HOOK | | u U+057D Armenian small letter SEH |
| U+03C6 ϰ Greek small letter PHI | | ϰ U+0444 Cyrillic small letter EF | |
| U+03AC ᾱ Greek small letter ALPHA WITH TONOS | ᾱ U+00E1 Latin small letter A WITH ACUTE | | |
| U+03AE ῆ Greek small letter ETA WITH TONOS | ῆ U+0144 Latin small letter N WITH ACUTE ῆ U+1E45 Latin small letter N WITH DOT ABOVE | | |
| U+03AF ῖ Greek small letter IOTA WITH TONOS | i U+0069 Latin small letter I í U+00ED Latin small letter I WITH ACUTE I U+0131 Latin small letter DOTLESS I (Variant defined by Latin GP) ï U+1EC9 Latin small letter I WITH HOOK ABOVE (Variant defined by Latin GP) | i U+0456 Cyrillic small letter BYELORUSSIAN-UKRAINIAN I | |
| U+03C1 ϱ Greek small letter RHO | p U+0070 Latin small letter P (Variant defined by Latin GP) | p U+0440 Cyrillic small letter ER (Variant defined by Latin GP) | |
| Proposal for a Greek Script Root Zone LGR | | 27 / 41 | Greek Generation Panel |

| Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|--|---|-------------------------------------|-----------------|
| U+03C3 σ Greek small letter SIGMA | σ U+01A1Latin small letter O WITH HORN (Variant defined by Latin GP) | | |
| U+03CC ó Greek small letter OMICRON WITH TONOS | ó U+00F3Latin small letter O WITH ACUTE ò U+022F*Latin small letter O WITH DOT ABOVE | | |
| U+03CD ú Greek small letter UPSILON WITH TONOS | ú U+00FALatin small letter U WITH ACUTE | | |
| U+03CA ï Greek small letter IOTA WITH DIALYTIKA | ï U+00EFLatin small letter I WITH DIAERESIS | й U+0457Cyrillic small letter YI | |
| U+03CB ü Greek small letter UPSILON WITH DIALYTIKA | ü U+00FCLatin small letter U WITH DIAERESIS | | |
| U+03B0 ű Greek small letter UPSILON WITH DIALYTIKA AND TONOS | ű U+01D8*Latin small letter U WITH DIAERESIS AND ACUTE | | |

*Code points not in draft Latin LGR. These are listed here for information but will not be included in the XML.

6.3 Summary of all variants

Based on sections 6.1.2, 6.1.3, and 6.2.5 the overall sets of variants are presented in table 10 below. These includes all the in-script and cross-script variants and any variants generated due to transitivity. These variant sets are captured in the XML.

Table 10: Overall Sets of Variant Mappings

| | Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|---|---|---|--|---|
| 1 | U+03B1 α Greek small letter ALPHA U+03AC ᾶ Greek small letter ALPHA WITH TONOS | a U+0061 LATIN small letter A á U+00E1 Latin small letter A WITH ACUTE | a U+0430 Cyrillic small letter A | |
| 2 | U+0390 ῀ Greek small letter IOTA WITH DIALYTIKA AND TONOS U+03AF ῁ Greek small letter IOTA WITH TONOS U+03B9 ι Greek small letter IOTA U+03CA ῂ Greek small letter IOTA WITH DIALYTIKA | i U+0069 Latin small letter I í U+00ED Latin small letter I WITH ACUTE ï U+00EF Latin small letter I WITH DIAERESIS I U+0131 Latin small letter DOTLESS I (Variant defined by Latin GP) ı U+0269 Latin small letter IOTA İ U+1EC9 Latin small letter I WITH HOOK ABOVE (Variant defined by Latin GP) | i U+0456 Cyrillic small letter BYELORUSSIAN-UKRAINIAN I й U+0457 Cyrillic small letter YI | Ի U+0582 Armenian small letter YIWN |

| | Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|---|---|--|--|--|
| 3 | U+03B7 η Greek small letter ETA U+03AE ῆ Greek small letter ETA WITH TONOS | n U+006E Latin small letter N η U+014B Latin small letter ENG ń U+0144 Latin small letter N WITH ACUTE ň U+1E45 Latin small letter N WITH DOT ABOVE | | ղ U+0572 Armenian small letter GHAD ռ U+0578 Armenian small letter VO |
| 4 | U+03BF ο Greek small letter OMICRON U+03CC ό Greek small letter OMICRON WITH TONOS | o U+006F Latin small letter O ó U+00F3 Latin small letter O WITH ACUTE | o U+043E Cyrillic small letter O | օ U+0585 Armenian small letter OH |
| 5 | U+03C1 ρ Greek small letter RHO | p U+0070 Latin small letter P (Variant defined by Latin GP) | р U+0440 Cyrillic small letter ER (Variant defined by Latin GP) | |
| 6 | U+03B2 β Greek small letter BETA | ß U+00DF Latin small letter SHARP S ss U+0073 U+0073 Latin small letter S Latin small letter S | ss U+0455 U+0455 Cyrillic small letter Dze Cyrillic small letter Dze | |

| | Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|----|--|--|---|--|
| 7 | U+03C5 ϰ Greek small letter UPSILON U+03CD ϱ Greek small letter UPSILON WITH TONOS U+03CB ϰ̄ Greek small letter UPSILON WITH DIALYTIKA U+03B0 ϰ̄̄ Greek small letter UPSILON WITH DIALYTIKA AND TONOS | u U+0075 Latin small letter U v U+028B Latin small letter V WITH HOOK ú U+00FA Latin small letter U WITH ACUTE ü U+00FC Latin small letter U WITH DIAERESIS | | ս U+057D Armenian small letter SEH |
| 8 | U+03BD ν Greek small letter NU | v U+0076 Latin small letter V | | |
| 9 | U+03B3 γ Greek small letter GAMMA | y U+0079 Latin small letter Y γ U+0263 Latin small letter GAMMA | у U+0443 Cyrillic small letter U Ү U+04AF Cyrillic small letter STRAIGHT U | |
| 10 | U+03C3 σ Greek small letter SIGMA U+03C2 ϣ Greek small letter FINAL SIGMA | ⱪ U+01A1 Latin small letter O WITH HORN (Variant defined by Latin GP) | | |

| | Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|----|---|---------------------------------------|--------------------------------------|-----------------|
| 11 | U+03B5 ε Greek small letter EPSILON U+03AD έ Greek small letter EPSILON WITH TONOS | ε U+025B Latin small letter OPEN E | | |
| 12 | U+03BA κ Greek small letter KAPPA | | κ U+043A Cyrillic small letter KA | |
| 13 | U+03C4 τ Greek small letter TAU | | τ U+0442 Cyrillic small letter TE | |
| 14 | U+03C6 φ Greek small letter PHI | | φ U+0444 Cyrillic small letter EF | |
| 15 | U+03C9 ω Greek small letter OMEGA U+03CE ώ Greek small letter OMEGA WITH TONOS | | | |

7 Whole Label Evaluation Rules (WLE) and Actions

The Greek LGR does not need any script specific WLE rules.

A set of Greek-specific actions has been defined that limit the number of allocatable variants for labels with multiple accented vowels (Section 6.1.2) or multiple final sigmas (Section 6.1.3).

8 Contributors

The Greek Generation Panel is composed of individuals actively involved in policy development processes in the areas of electronic communication, research and development related to the Greek language, standardization, computing and maintenance of the domain name system in Greece and Cyprus. The panelists come from governmental and regulatory authorities, academia, private sector, ccTLD registries of Greece and Cyprus. Moreover, the Panel was officially formed with the Decision 54020/1088/25 Nov 2015 of the Minister of Infrastructure, Transport and Networks of the Hellenic Republic. The formation of the Greek Generation Panel is as follows:

| Name | Role | Organization | Country | Field of Expertise |
|-----------------------------|------------|--|---------|---|
| Panagiotis PAPASPILIOPOULOS | Chair | Hellenic Ministry of Digital Governance | Greece | Technical, Policy Development, GAC Rep |
| Vaggelis SEGREDAKIS | Vice Chair | Registry of [.gr] & [.ελ] domain names, FORTH-ICS | Greece | Technical (DNS) ccNSO Rep |
| Konstantinos VALASIS | Member | Hellenic Ministry of Foreign Affairs | Greece | Policy Development |
| Ioannis KATSARAS | Member | Hellenic Ministry of Foreign Affairs | Greece | Policy Development |
| Asimina GIANNOPOULOU | Member | Hellenic Telecommunications and Post Commission | Greece | Technical, Regulatory |
| Giorgos KOLYVAS | Member | Hellenic Telecommunications and Post Commission | Greece | Technical, Regulatory |
| Michael VASSILAKIS | Member | Registry of [.gr] & [.ελ] domain names, FORTH-ICS | Greece | Technical, ccNSO Rep |
| Antonis ANTONIADES | Member | Office of Electronic Communications and Postal Regulations | Cyprus | Technical, Policy Development, Regulatory |
| Constantinos STYLIANOU | Member | Office of Electronic Communications and Postal Regulations | Cyprus | Technical, Policy Development, Regulatory |

| | | | | |
|--------------------|--------|---|--------|---------------------|
| Alexandros PSYRRIS | Member | Hellenic Body for Standardization | Greece | Technical (Unicode) |
| Maria GAVRIILIDOU | Member | Institute for Language and Speech Processing / Research Centre "Athena" | Greece | Linguistics |
| Penny LABROPOULOU | Member | Institute for Language and Speech Processing / Research Centre "Athena" | Greece | Linguistics |

The Panel would also like to express its gratitude for the valuable contribution of the following persons (in alphabetic order according to their surname):

- Sarmad HUSSAIN, ICANN IDN Team, IDN Project Coordinator
- Ritvan KARA-CHOTZA, Educationist, Lexicographer
- Pitinan KOOARMORNPATANA, ICANN IDN Team
- Chrisoula MICHALIDOU, Legal Expert
- Antonios PERPATIDIS, Interpreter

9 References

1. "The Unicode Standard", the Unicode Consortium
URL: <http://www.unicode.org/standard/standard.html>.
2. "Maximal Starting Repertoire – MSR-5", ICANN IDN Integration Panel
URL: <https://www.icann.org/en/system/files/files/msr-5-overview-24jun21-en.pdf>.
3. "Requirements for LGR Proposals from Generation Panels", Integration Panel, ICANN IDN Variant Issues Project
URL: <https://www.icann.org/en/system/files/files/Requirements-for-LGR-Proposals-20150424.pdf>.
4. "Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels", ICANN
URL: <http://www.icann.org/en/resources/idn/variant-tlds/draft-lgr-procedure-20mar13-en.pdf>.
5. "IP Review of Variants for the Proposed Greek LGR dated 2019-09-01", dated 2019-12-04.
6. "Study of the issues present in the registration of IDN TLDs in Greek characters", Greek Case Study Team, ICANN IDN Variant Issues Project
URL: <http://archive.icann.org/en/topics/new-gtlds/greek-vip-issues-report-07oct11-en.pdf>.

7. "Internationalized Domain Names for Applications (IDNA) Background, Explanation, and Rationale", IETF, RFC5894, August 2010.
8. "Greek language question", Wikipedia, http://en.wikipedia.org/wiki/Greek_language_question.
9. "Katharevousa", Wikipedia, <http://en.wikipedia.org/wiki/Katharevousa>.
10. "Demotic Greek", Wikipedia, <http://en.wikipedia.org/wiki/Dimotiki>.
11. "Greek diacritics", Wikipedia, http://en.wikipedia.org/wiki/Greek_diacritics.
12. "Pomak language", Wikipedia, https://en.wikipedia.org/wiki/Pomak_language.
13. "Πομακική γλώσσα", Wikipedia, https://el.wikipedia.org/wiki/Πομακική_γλώσσα (in Greek).
14. "Greek alphabet", Wikipedia, https://en.wikipedia.org/wiki/Greek_alphabet_-_Use_for_other_languages.
15. "Character codes for Greek: Problems and modern solutions", Macrakis Stavros M., 1996.
16. "Greek: *An Essential Grammar of the Modern Language*", Holton, Mackridge, Philippaki, Routledge 2004, ISBN 0-203-64521-9, Master e-book ISBN: http://www.verbalplanet.com/resourcedownload.asp?re_id=100001646935.
17. "Pomak-Greek Dictionary", Kara-Chotza Ritvan et al., 4th Army Corps, Army General Staff, Hellenic Ministry of Defense, 1996.
18. Email correspondence (in Greek) between the Greek Generation Panel chair and Mr. Ritvan Kara-Chotza, 2017.
 "Regulation on Management and Assignment of [.gr] or [.ελ] Domain Names", Decision No.843/2/01 Mar 2018 of the Hellenic Telecommunications and Post Commission (Issue No. 973/B'/19 March 2018 of the Hellenic Official Governmental Gazette),
http://www.eett.gr/opencms/export/sites/default/EETT_EN/Electronic_Communications/DomainNames/AP843-002.pdf).

Appendix: Cross-script homoglyphs – possible confusability issues

A.1 Cross-script similarity between code-points included in MSR-5

| Greek Letter | Latin Letter | Cyrillic Letter | Armenian Letter |
|---|---|---|---|
| U+03B1 α Greek small letter ALPHA | ā U+0061+U+0331 LATIN small letter A + COMBINING MACRON BELOW | | |
| U+03B2 β Greek small letter BETA | | ѵ U+0432 Cyrillic small letter VE | |
| U+03B3 γ Greek small letter GAMMA | | у U+04B1 Cyrillic small letter STRAIGHT U WITH STROKE | |
| U+03B5 ε Greek small letter EPSILON | | є U+0454 Cyrillic small letter UKRAINIAN IE | |
| U+03B7 η Greek small letter ETA | ñ U+0146 Latin small letter N WITH CEDILLA ņ U+1E49 Latin small letter N WITH LINE BELOW ḥ U+1E4B Latin small letter N WITH CIRCUMFLEX BELOW | | դ U+0564 Armenian small letter DA |
| U+03B8 θ Greek small letter THETA | ø U+0275 Latin small letter BARRED O | ѐ U+04E9 Cyrillic small letter BARRED O | |
| U+03BA κ Greek small letter KAPPA | k U+006B Latin small letter K ķ U+0199 Latin small letter K WITH HOOK ķ̇ U+0137 Latin small letter K WITH CEDILLA ķ̆ U+1E33 Latin small letter K WITH DOT BELOW | к U+049B Cyrillic small letter KA WITH DESCENDER к̆ U+04A1 Cyrillic small letter BASHKIR KA | |
| U+03BD ν Greek small letter NU | ν̇ U+1E7F Latin small letter V WITH DOT BELOW | ў U+04B1 Cyrillic small letter STRAIGHT U WITH STROKE | |
| U+03BF ο Greek small letter OMICRON | ȯ U+006F+U+0331 | | |

| | | | |
|---|--|---|--|
| Greek small letter OMICRON | Latin small letter O + COMBINING MACRON BELOW ø U+1ECD Latin small letter O WITH DOT BELOW | | |
| U+03C0 π Greek small letter PI | | п U+043F Cyrillic small letter PE п U+0525 Cyrillic small letter PE WITH DESCENDER | |
| U+03C1 ρ Greek small letter RHO | ρ U+0070 Latin small letter P | р U+0440 Cyrillic small letter ER | |
| U+03C2 ς Greek small letter FINAL SIGMA | s U+0073 Latin small letter S ş U+1E63 Latin small letter S WITH DOT BELOW ç U+00E7 Latin small letter C WITH CEDILLA ș U+015F Latin small letter S WITH CEDILLA ş U+0219 Latin small letter S WITH COMMA BELOW | s U+0455 Cyrillic small letter DZE ç U+04AB Cyrillic small letter ES WITH DESCENDER | |
| U+03C3 σ Greek small letter SIGMA | σ U+01A1 Latin small letter O WITH HORN σ U+1EE3 Latin small letter O WITH HORN AND DOT BELOW | | |
| U+03C4 τ Greek small letter TAU | | т U+04AD Cyrillic small letter TE WITH DESCENDER | |
| U+03C5 υ Greek small letter UPSILON | υ U+1EE5 Latin small letter U WITH DOT BELOW | | |
| U+03C7 χ Greek small letter CHI | x U+0078 Latin small letter X | x U+0445 Cyrillic small letter HA χ U+04B3 | |

| | | | |
|--|---|---|--|
| | | Cyrillic small letter HA WITH DESCENDER | |
| U+03C9 ω Greek small letter OMEGA | w U+0077 Latin small letter W | | |
| U+03AE η Greek small letter ETA WITH TONOS | ñ U+01F9 Latin small letter N WITH GRAVE | | |
| U+03AF ί Greek small letter IOTA WITH TONOS | ï U+1ECB Latin small letter I WITH DOT BELOW ı U+0069+U+0331 Latin small letter I + COMBINING MACRON BELOW ì U+00EC Latin small letter I WITH GRAVE ï U+1EC9 Latin small letter I WITH HOOK ABOVE | | |
| U+03CC ó Greek small letter OMICRON WITH TONOS | ò U+00F2 Latin small letter O WITH GRAVE ò̇ U+1ECD+U+0300 Latin small letter O WITH DOT BELOW + COMBINING GRAVE ACCENT ò̇ U+1ECD+U+0301 Latin small letter O WITH DOT BELOW + COMBINING ACUTE ACCENT ỏ U+1ECF Latin small letter O WITH HOOK ABOVE | | |
| U+03CD ú Greek small letter UPSILON WITH TONOS | ù U+00F9 Latin small letter U WITH GRAVE ủ U+1EE7 Latin small letter U WITH HOOK ABOVE ư U+1EE9 Latin small letter U WITH HORN AND ACUTE | | |

| | | | | |
|---|---|--------|--|--|
| | ǔ | U+1EEB | | |
| | Latin small letter U WITH HORN AND GRAVE | | | |
| U+03CE | ώ | U+1E83 | | |
| Greek small letter OMEGA WITH TONOS | Latin small letter W WITH ACUTE | | | |
| | ŵ | U+1E87 | | |
| | Latin small letter W WITH DOT ABOVE | | | |
| U+03CB | ü | U+0171 | | |
| Greek small letter UPSILON WITH DIALYTIKA | Latin small letter U WITH DOUBLE ACCUTE | | | |
| U+03B0 | ÿ | U+01D8 | | |
| Greek small letter UPSILON WITH DIALYTIKA AND TONOS | Latin small letter U WITH DIAERESIS AND GRAVE | | | |

A.2 Uppercase Similarity Table

| Uppercase Similarity Table | | | |
|------------------------------|---|------------------------|-----------------------------|
| Greek Letter | | Latin Letter | Cyrillic Letter |
| U+0391 | A | A | U+0041 |
| GREEK CAPITAL LETTER ALPHA | | LATIN CAPITAL LETTER A | U+0410 |
| | | | CYRILLIC CAPITAL LETTER A |
| U+0392 | B | B | U+0042 |
| GREEK CAPITAL LETTER BETA | | LATIN CAPITAL LETTER B | U+0412 |
| | | | CYRILLIC CAPITAL LETTER VE |
| U+0393 | Γ | | U+0413 |
| GREEK CAPITAL LETTER GAMMA | | | CYRILLIC CAPITAL LETTER GHE |
| U+0394 | Δ | | |
| GREEK CAPITAL LETTER DELTA | | | |
| U+0395 | E | E | U+0045 |
| GREEK CAPITAL LETTER EPSILON | | LATIN CAPITAL LETTER E | U+0415 |
| | | | CYRILLIC CAPITAL LETTER IE |
| U+0396 | Z | Z | U+005A |
| GREEK CAPITAL LETTER ZETA | | LATIN CAPITAL LETTER Z | |
| U+0397 | H | H | U+0048 |
| GREEK CAPITAL LETTER ETA | | LATIN CAPITAL LETTER H | U+041D |
| | | | CYRILLIC CAPITAL LETTER EN |
| U+0398 | Θ | | |

| | | |
|--|------------------------------------|--|
| GREEK CAPITAL LETTER THETA | | |
| U+0399 I GREEK CAPITAL LETTER IOTA | I U+0049 LATIN CAPITAL LETTER I | |
| U+039A K GREEK CAPITAL LETTER KAPPA | K U+004B LATIN CAPITAL LETTER K | K U+041A CYRILLIC CAPITAL LETTER KA |
| U+039B Λ GREEK CAPITAL LETTER LAMDA | | |
| U+039C M GREEK CAPITAL LETTER MU | M U+004D LATIN CAPITAL LETTER M | M U+041C CYRILLIC CAPITAL LETTER EM |
| U+039D N GREEK CAPITAL LETTER NU | N U+004E LATIN CAPITAL LETTER N | |
| U+039E Ξ GREEK CAPITAL LETTER XI | | |
| U+039F O GREEK CAPITAL LETTER OMICRON | O U+004F LATIN CAPITAL LETTER O | O U+041E CYRILLIC CAPITAL LETTER O |
| U+03A0 Π GREEK CAPITAL LETTER PI | | Π U+041F CYRILLIC CAPITAL LETTER PE |
| U+03A1 P GREEK CAPITAL LETTER RHO | P U+0050 LATIN CAPITAL LETTER P | P U+0420 CYRILLIC CAPITAL LETTER ER |
| U+03A3 Σ GREEK CAPITAL LETTER SIGMA | | |
| U+03A4 T GREEK CAPITAL LETTER TAU | T U+0054 LATIN CAPITAL LETTER T | T U+0422 CYRILLIC CAPITAL LETTER TE |
| U+03A5 Y GREEK CAPITAL LETTER UPSILON | Y U+0059 LATIN CAPITAL LETTER Y | |
| U+03A6 Φ GREEK CAPITAL LETTER PHI | | Φ U+0424 CYRILLIC CAPITAL LETTER EF |
| U+03A7 X GREEK CAPITAL LETTER CHI | X U+0058 LATIN CAPITAL LETTER X | X U+0425 CYRILLIC CAPITAL LETTER HA |
| U+03A8 Ψ GREEK CAPITAL LETTER PSI | | |

| | | |
|--|--|--|
| U+03A9 Ω GREEK CAPITAL LETTER OMEGA | | |
|--|--|--|

A.3 Effect of Transitivity on Uppercase Similarity

| Transitivity of Uppercase Similarity | | | | | |
|---|---|-------------------------------------|---------------------------------------|---|---|
| Greek small Letter | Greek Capital Letter | Latin small Letter | Latin Capital Letter | Cyrillic small Letter | Cyrillic Capital Letter |
| U+03B7 η GREEK SMALL LETTER ETA | U+0397 Η GREEK CAPITAL LETTER ETA | U+006E n LATIN SMALL LETTER N | U+004E N LATIN CAPITAL LETTER N | | |
| U+03B7 η GREEK SMALL LETTER ETA | U+0397 Η GREEK CAPITAL LETTER ETA | U+0068 h LATIN SMALL LETTER H | U+0048 H LATIN CAPITAL LETTER H | U+043D ѡ CYRILLIC SMALL LETTER EN | U+041D Ѣ CYRILLIC CAPITAL LETTER EN |
| U+03BD υ GREEK SMALL LETTER NU | U+039D Ν GREEK CAPITAL LETTER NU | U+0076 v LATIN SMALL LETTER V | U+0056 V LATIN CAPITAL LETTER V | | |
| U+03C5 υ GREEK SMALL LETTER UPSILON | U+03A5 Υ GREEK CAPITAL LETTER UPSILON | U+0075 u LATIN SMALL LETTER U | U+0055 U LATIN CAPITAL LETTER U | | |
| U+03C5 υ GREEK SMALL LETTER UPSILON | U+03A5 Υ GREEK CAPITAL LETTER UPSILON | U+0079 y LATIN SMALL LETTER Y | U+0059 Y LATIN CAPITAL LETTER Y | U+0443 ѣ CYRILLIC SMALL LETTER U | |
| | | U+0079 y LATIN SMALL LETTER Y | U+0059 Y LATIN CAPITAL LETTER Y | U+0443 ѣ CYRILLIC SMALL LETTER U | U+0423 Ѣ CYRILLIC CAPITAL LETTER U |