

IDNs in Greece

Segredakis Vaggelis

Administrator

Registry of [.gr] Domain Names

Institute of Computer Science

Foundation for Research and Technology – Hellas

Email: segred@ics.forth.gr



IDN Facts

- The registration of .gr IDNs has started on the **4th of July 2005**.
- We have registered about **8500** IDN domain names to date.
- We allow all Greek characters, current and ancient.
- Single script registrations are the only allowed.
- We register variants and homograph names under a special procedure.



IDN Facts (ii)

- In Greek language we have a tonic accent sign called “Tonos” which is used in most Greek words.
- The punctuation of the ancient Greek words is even more complicated.
- The possibility to put the tonos on a different letter of a word creates different domain names in the Punycode protocol e.g. πέρνα – περνά. These words however in capital letters should be the same: ΠΕΡΝΑ. We consider them variants.
- The similarity of some Greek characters with latin characters creates an even more delicate problem: Homographs!



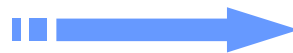
Examples of Homograph Names

eett (**EETT**)



Εεττ (**EETT**)

Ote (**OTE**)



Οτε (**OTE**)

Etba (**ETBA**)



Ετβα (**ETBA**)

Ana (**ANA**)



Ava (**ANA**)

Ava (**AVA**)



Ava (**ANA**)

Even more Homograph names could be created if you mix scripts. We decided to register domain names containing only a single character set.



The case of Homograph Domains

- Homograph domain names can mislead a user to wrong addresses with Phishing issues.

www.EETT.gr (www.eett.gr)

is very similar to www.EETT.gr (www.εεττ.gr)!



Was it necessary to regulate this?

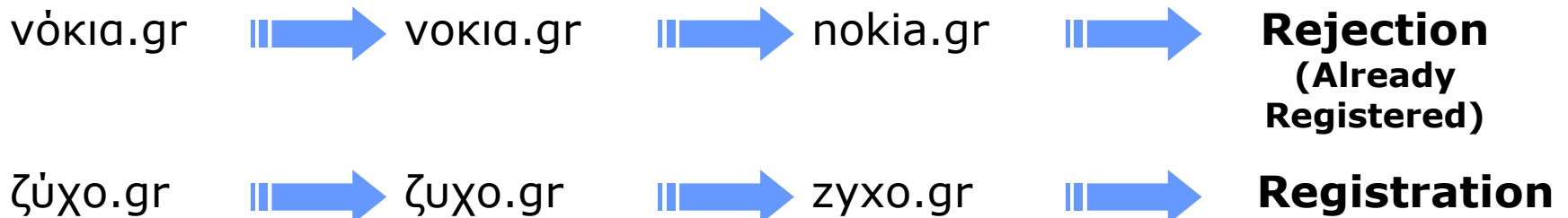
- If this issue was not regulated the user would end up with visually similar but different domain names.
- The possible misuse of these domain names had already created issues and problems to other registries by the time we decided to provide IDNs to our registrants.
- Both ICANN and CENTR had responded to this issue proposing to the registries to act against it.
- Because of the close similarity of Latin characters to the Greek characters the problem would soon become substantial.



Which was the solution?

- The solution selected was to check each character in a domain name against Latin characters. If each character of a word is a homograph character we check to see if this Latin based domain name is registered in the registry database. If it is not we allow the registration, in any other case the domain name is considered reserved only for the registrant of the original domain name. (the same applies if you first register an IDN and then try to register the homograph in Latin characters).

Example:



Activation of a Homograph Domain Name

- The Homograph Domain Names can only be activated to the registrant who is using the initial Domain name.
- The Homograph Domain Names are reserved for each registrant if his choice of Domain Name is eligible in creating homographs (e.g. the name maria.gr does not have any homographs because of the letter R. Instead, the name mapia.gr -> MAPIA.gr -> μαρία.gr).



Activation of a Homograph Domain Name (ii)

- For Domain names that pre-existed this rule and they should be able to register the same homograph domain name, no registrant was allowed to register this homograph.
- Each Homograph activation is charged and the bundle of domain names created acts as a single registration (deleted, renewed, expired on the same day).



Character Substitution

- The substitution of characters is like this $\alpha \rightarrow a$, $\beta \rightarrow b$, $\epsilon \rightarrow e$ e.t.c.
- Two special groups of characters are the following where one character can look like many, depending on the fonts used

h,n,v,η,v

u,y,u

We substitute each letter with each other and reserve more domain names:

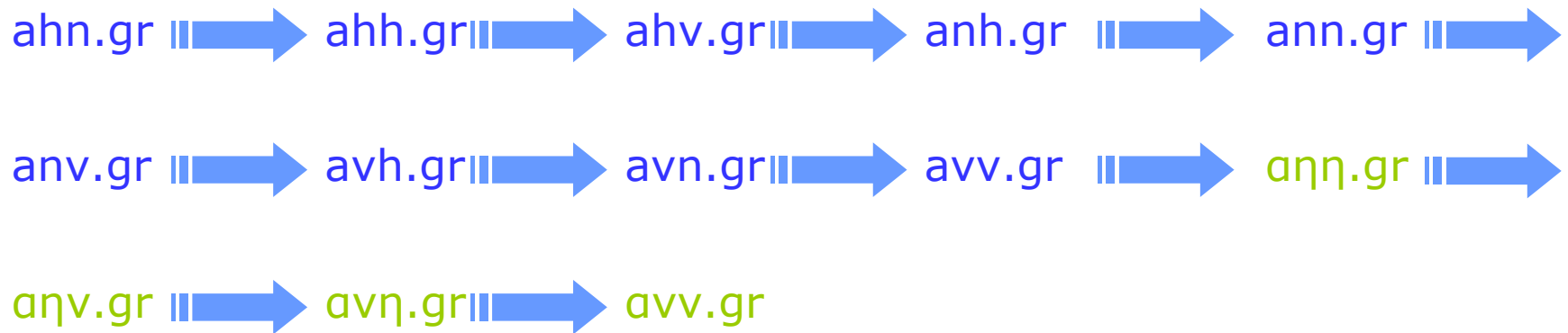
aha.gr \Rightarrow ana.gr \Rightarrow ava.gr \Rightarrow αηα.gr \Rightarrow αβα.gr

(In Green the Greek characters, in blue the Latin ones)



Character Substitution(2)

A more complicated substitution:



These substitutions create some domain names that no one would wish to use but you have to exclude them because they are visually similar:
 ahh.gr -> AHH.gr -> AHH.gr -> ανη.gr -> ann.gr -> ANN.gr -> ANN.gr -> avv.gr e.t.c.

(In Green the Greek characters, in blue the Latin ones)



Character Substitution Table

GREEK HOMOGRAPH CHARACTER TABLE (SIMPLIFIED)				
GREEK CHARACTER		LATIN CHARACTER		
α	A	a	A	a
β	B	b	B	a
ε	E	e	E	e
ζ	Z	z	Z	z
ι	I	i	I	i
κ	K	k	K	k
μ	M	m	M	m
ο	O	o	O	o
ρ	P	p	P	p
τ	T	t	T	t
χ	X	x	X	x
η	H	n	N	
η	H	h	H	
υ	N	v	V	
υ	Y	u	U	
υ	Y	y	Y	



The Punycode Problem

- For the Greek language to be correctly represented we should have each domain name corresponding both in the capital letter word and in the small caps word.
- If for example we use the word “δοκιμή” as a domain name we expect that the capital letters translation of this domain should be the corresponding capital word.
- The word “δοκιμή” becomes “ΔΟΚΙΜΗ”. Instead the domain “δοκιμή.gr” becomes “ΔΟΚΙΜΉ.gr” instead of the correct “ΔΟΚΙΜΗ.gr”. Why??



The Punycode Problem (ii)

- The domain name “ΔΟΚΙΜΉ.gr” is actually the xn--jxalpdlp.gr where the domain name “ΔΟΚΙΜΗ.gr” is the xn--pxagfdlp.gr.
- The Punycode translation is not a very efficient way to translate the Greek language to domain names since for each word a user has to register two domain names to achieve the representation.
- This fact was recognized and implemented in our IDN system. For each domain name with a tonos the user gets it's variant without the tonos for free. The variant without tonos is considered the main Domain.



The DNAME solution

- Both these domain names are inserted in the .gr zone file as DNAMEs so that the user has nothing more to do for setting up this domain name than asking his ISP to set up a zone for the domain name without the tonos, in the case of “ΔOKIM´H.gr” and “ΔOKIMH.gr” the xn--pxagfdlp.gr.

```

xn--pxagfdlp.gr    IN      NS      ns1. xn--pxagfdlp.gr.
                   IN      NS      ns2. xn--pxagfdlp.gr.
xn--jxalpdlp.gr   IN      DNAME  xn--pxagfdlp.gr.
  
```



The DNAME solution (ii)

- Unfortunately the use of DNAME is not the total solution. It works fine if you want to visit `www.δοκιμή.gr` but you could never visit `δοκιμή.gr`. You could send an email to `info@www.δοκιμή.gr` but never to `info@δοκιμή.gr` because this DNAME works only for the branches below the link.
- To solve the DNAME problem we decided to provide the choice to registrants to use their variants or homographs as normal domain names if they decided that the DNAME solution is not working for them.
- Every other variant of the domain name is inserted in the zone file either as a DNAME or as a Fully Qualified Domain Name.
- In case a Greek IDN is a homograph of a Latin domain name, all the homographs and variants are DNAMED to the first (as in First Come First Serve) of the two registrations.



The Variants

- Why only register δοκιμή.gr and not δόκιμη.gr? “δόκιμη” is a word as well.
- Both these words in capital are the same “ΔΟΚΙΜΗ” and so they correspond to the same word without tonos “δοκιμη”.
- We allow this registrations with the same procedure we use for Homographs. They can only be registered by the same registrant.
- The initial domain name, the homographs and the variants become an object we call “Bundle”. They can be either DNAMED or inserted in the Zone file with NS records but the same NS records for all these names.



IDN.IDN

Why did we choose to present our IDN policy today and not in 2005?

- We believe our system for IDN registrations is one of the more complicated ones but it is actually working very well. We had to let time pass before we could be certain.
- We would like to express our interest in IDN.IDN entries in the Root Zone File. We are one more example that shows that local policy in registrations is the best way to move forward, especially when so many local parameters come in play in each IDN registration.
- Since our character set is not the Latin character set we are very confident that a multilingual Internet should start at the very top and all character sets should be allowed in the root zone.



IDN.IDN (ii)

- We would welcome participating in a test phase of IDN.IDN in the root to let our users decide what domain name they wish to have as .IDN. We already have proposed to ICANN to use the following domains in the test phase:
 - .xn--qxam (.ελ)
 - .xn--mxahsa5b (.ελλας)
 - .xn--hxarsa5b (.ελλάς)
 - .xn--mxaaic4aa (.ελλαδα)
 - .xn--hxakic4aa (.ελλάδα)
- Since no ISO-3166 could be adopted this time for the form of the .IDN domains we believe that user experience along with consultations with the Local Internet Community could be a solution for the fast adoption of the IDNs in the Root Zone.



Your Questions

?



Thank you for your attention and patience

Segredakis Vaggelis
Administrator
Registry of [.gr] Domain Names
Institute of Computer Science
Foundation for Research and Technology – Hellas
Email: segred@ics.forth.gr

