

25 July 2011

# ICANN IDN variant project

Educational session on existing  
variant practices

Dr. Andrzej Bartosiewicz, Yonita Inc.

**Yonita Inc.**

800 West El Camino Real, Suite 180  
Mountain View, CA 94040, United States

<http://Yonita.com>  
[andrzej@Yonita.com](mailto:andrzej@Yonita.com)  
+1 650 2493707





## Dr. Andrzej Bartosiewicz, CEO of Yonita Inc.

- Ph.D. in Informatics, Warsaw University of Technology
- Career in Domain Names related business:
  - Director and Chairman of CENTR (2006-2010)
  - Rapporteur and Chairman of ITU-T in the field of “Internationalized Domain Names” (2005-2008)
  - Head of .PL Registry (2001-2010)
- Ongoing Senior Advisor to the International Telecommunication Union
- Over 200 presentations and publications

# Experience with variants

- Based on my submission\* of the IETF Internet Drafts<sup>†</sup> and ICANN “IDN tables”<sup>‡</sup>, .PL Registry has introduced (2003-2005) registration of the domains in the Latin character set based on the language tables
- **No variants** management techniques have been implemented so far.
- No techniques on language sets mixing within Latin script has been introduced.
- **Identical approach has been taken successfully by other European ccTLDs.**

\*) WITH CO-AUTHOR

†) IETF INTERNET DRAFT (EXPIRED):

[HTTP://TOOLS.IETF.ORG/HTML/DRAFT-BARTOSIEWICZ-IDN-PLTLD-08](http://tools.ietf.org/html/draft-bartosiewicz-idn-pltld-08)

‡) IDN TABLES REPOSITORY (IDN TABLES USED IN REGISTRIES):

[HTTP://WWW.IANA.ORG/DOMAINS/IDN-TABLES/](http://www.iana.org/domains/idn-tables/)

# Experience with variants (2)

- Language covered in 2003-2004 deployment:  
Catalan, Czech, Danish, German, Esperanto, Spanish, Estonian, Finnish, French, Irish, Croatian, Hungarian, Icelandic, Italian, Luxembourgish, Lithuanian, Latvian, Moldavian (Romanian script), Maltese, Dutch, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Albanian, Swedish and Turkish.

# Question

- Are the variant issues at the root level (1<sup>st</sup>) and the 2<sup>nd</sup> level similar?
- If yes, why not to focus on existing experience of ccTLDs and gTLDs?

# Feedback since 2003

- The policy has been widely presented by Bartosiewicz during ICANN, CENTR, IETF and other meetings.
- Since the deployment of the whole Latin character set, no complains (until March 2010) regarding variants or look-alike domains has been reported neither from community nor from other cc/gTLDs.
- Many of other ccTLDs has adopted similar policies.

# Problems

- The only problems reported were related to the **script mixing** between Latin and other scripts (Latin & Cyrillic's and Latin & Greek) → look alike domains issue.
- Solution:
  - script mixing blocked +
  - no registration of look-alike Cyrillic domains to the already registered Latin variants (should apply to Greek too)

# Suggestion:

## Next steps for IDN Variant Project

- Based on the existing experience of cc/gTLDs with IDN deployment at 2<sup>nd</sup> level, it's recommended to base strategy of Latin group further work on the existing, already identified problems with variants and homographs etc.
- List of issues should be identified, discussed and conclusions reached based on the **real-life examples**, not the theoretical one.



# Confusable variants to be subject for analysis

- It's always a difficult decision what should be treated as variant and what shouldn't. Common sense should always prevail.
- Some variants are confusable (look-alike domains or visually confusing), some not (or confusable in very specific situations):
  - Confusable: U+0443 у “Cyrillic Small Letter U” and U+0079 y “Latin Small Letter Y”
  - Semi-confusable: U+00E6 æ “Latin Small Letter Æ” and combination of “ae”

# Results

- Complete list of characters or combination of characters (pairs) that can be treated as **real confusable** variants should be created and publicized
- Such list should be taken into consideration when new TLDs are being assigned.
- Rules for new TLDs should apply to the domain name (as stated in the application) as well as all **reasonable variants of the name** in the particular **context**, only **if** variants can create confusion.

# Contact details

Dr. Andrzej Bartosiewicz

andrzej@Yonita.com

<http://yonita.com>

US mobile: +1 650 2493707

Polish mobile: +48 518 235209