



IDN variant TLDs

*A Study of Issues Related to the
Delegation of IDN Variant TLDs*

in San Francisco ■ 13-18 March 2011



**SILICON
VALLEY
Meeting**

Agenda

- Call to order by Dennis Jennings [2 min]
- Introduction by Kurt Pritz [3 min]
- Project Proposal by Francisco Arias [10 min]
- Potential Issues in Arabic by Dr. Sarmad Hussain [10 min]
- Potential Issues in Chinese by Dr. Xiaodong Lee [10 min]
- Potential Issues in Indic by Dr. Mahesh Kulkarni [10 min]
- Potential Issues in Latin by Dr. Cary Karp [10 min]
- Discussion and Q&A [25 min]



Staff team

- Kurt Pritz - Executive sponsor
- Dennis Jennings - Project leader
- Anand Mishra - Project coordinator
- Naela Sarras - Community reps
- Kim Davies - DNS and Security
- Baher Esmat - Linguistic
- Steve Sheng - Policy
- Francisco Arias - Registry Operations



Presenters

- Dr. Sarmad Hussain - Professor and Head, Center for Language Engineering, Al-Khawarizmi Institute of Computer Science, University of Engineering and Technology, Lahore, Pakistan
- Dr. Xiaodong Lee - VP, CTO of CNNIC
- Dr. Mahesh Kulkarni - Programme Coordinator & HEAD GIST Centre for Development of Advanced Computing, Pune, India
- Dr. Cary Karp - Director of Internet Strategy and Technology at the Swedish Museum of Natural History



Cyrillic case

- Andrei Kolesnikov - CEO
Coordination Center for TLD .RU
- Has advised that IDN variants are
"not applicable in our case":

http://www.iana.org/domains/idn-tables/tables/ru_ru-ru_1.0.html



Project proposal

*A Study of Issues Related to the
Delegation of IDN Variant TLDs*

Why this project?



- Long standing request from a number of IDN user communities.
- Board instructed staff to develop an issues report on the subject.

<http://www.icann.org/en/minutes/resolutions-25sep10-en.htm#2.5>



IDN variant TLDs as of today



- Applicants may declare variant strings for the TLD in its application.
- No variant TLD strings are delegated until variant management solutions are developed and implemented.



What is being proposed?

TLD proposal
arabic
indian
cyrillic
chinese
variant
IDN

- Conduct five study cases with participants from the community:
 - 1) Arabic
 - 2) Chinese
 - 3) Cyrillic
 - 4) Indic
 - 5) Latin



Case study composition



- Each study case composed of community members* experienced in:
 - DNS,
 - Security,
 - Policy,
 - Linguistics,
 - Registry operations, and
 - Community representatives

* Some members are shared across case study teams



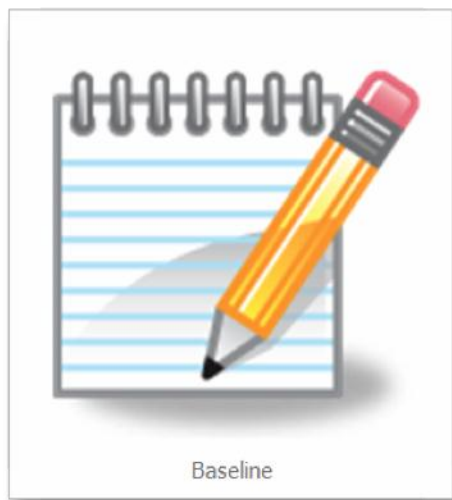
Avoid reinventing the wheel



- Survey of current TLDs' policy on IDN variants.
- Coordination with other groups working on the subject (DNSEXT, JIG, etc.)
- Use documents available from other groups as input.



Project scope



1. Create a glossary of terms vetted with technical and linguistic communities.
2. Identify challenges (requirements) of IDN variant TLDs based on:
 - a) linguistic accuracy,
 - b) technical feasibility & accuracy,
 - c) usability (including minimize user confusion),
 - d) accessibility, and
 - e) security and stability



Outcome of the project



- Issues Report from each case study
- Issues Report integrating common and case-specific issues



Timeline



Task	Time
Goal setting	31 March 2011
Recruit case study team members	29 April 2011
Complete case studies	30 September 2011
Synthesize issues across cases	30 November 2011
Issues report	15 December 2011



We value your feedback



- Submit feedback through the public comments box: <http://www.icann.org/en/public-comment/#idn-variant-tld>
- Talk with project coordination team in SF





Thank You

in San Francisco ■ 13-18 March 2011



**SILICON
VALLEY**
Meeting

Questions

One World

One Internet



Definitions of Variant TLDs

- No universally accepted definition.
- In our proposal, the following* is used:
 - “Variant characters occur where a single conceptual character can be identified with two or more different Unicode Code Points with graphic representations that may or may not be visually similar. IDN variant TLDs contain one or more characters that have such variants.”



* Based on definition from RFC 3743

Problem space



		Unilingual		Multilanguage	
		<i>Uni-script</i>	<i>Multi-script</i>	<i>Uni-script</i>	<i>Multi-script</i>
Equivalence	<i>Char-level</i>	A	B	C	D
	<i>String-level</i>	E	F	G	H
Visual similarity	<i>Char-level</i>	I	J	K	L
	<i>String-level</i>	M	N	O	P

