IDNs and Variants for the Root

Sarmad Hussain
Center for Language Engineering
Al-Khawarizmi Institute of Computer Science
University of Engineering and Technology
Lahore, Pakistan
Recent Community Work on Arabic IDNs

• Very active Arabic script community; time line of recent Arabic script community work
  – 2008: Arabic Script IDN Working Group (ASIWG)
  – 2009: ICANN IDN ccTLDs Fast Track Program initiated
  – 2010: Arabic Language in Internet Domains (RFC 5564)
  – 2011: Arabic Script Case Study for IDN Variants in the Root
  – 2012: User Experience Study on IDN Variants in the Root
  – 2013: Task Force on Arabic Script IDNs by MESWG

• Significant expertise in a variety of relevant domains: Unicode, DNS, IDNA and Registry and Registrar Operations
  – experience from SLD deployments in ccTLDs
  – experience from IDN ccTLD deployments
Typology of Arabic IDN Variants

• Same
  – Identical U+06CC (ی/ى) ی (U+0649)
  – In Context U+06A9 (ک/ک) کبک (U+0643)
  – Normalization U+0632 (ز/ر) ز = ُ + ر (U+0631+U+06EC)

• Similar
  – Character U+06AA (ک/ک) ک (U+06A9)
  – Diacritic U+062A (ت/ت) ت (U+067A)

• Different
  – Shape U+06C3 (ب/ة) بة (U+0629)
  – Character U+06C1 (ه/ة) ه (U+0629)
IDN Variants Needs and Challenges

Security and Stability Needs

- 120+ cases of visually same or similar Arabic script characters identified by case study team
  - Variants must not be allocated independently
  - Variants may need activation to allow user access (w/ different KB)
- 16 IDN ccTLD applications with 4 applications with variants

Security and Stability Challenges

- Consistency and innumerabley
  - Consistent across and within TLDs
  - Minimal activation for manageability
- Management tools
  - Registration
  - Configuration and Maintenance
  - Security and Monitoring
- Usability in applications
  - Browsing, emailing, etc.
  - Searching, privacy, etc.
Way forward for Community: Task Force on Arabic Script IDNs

• Creation and Oversight by community based MESWG
• Objectives: a holistic approach
  – Arabic Script Label Generation Ruleset (LGR) for the Root Zone
  – Second level LGRs for the Arabic script
  – Arabic script Internationalized Registration Data Protocol and Practice
  – Universal acceptability of Arabic script IDNs and variants
  – Technical challenges around registration of Arabic IDNs and variants
  – Operational software for Arabic script IDN registry and registrar operations
  – DNS security matters specifically related to Arabic IDNs and variants
  – Technical training material around Arabic script IDNs
Task Force on Arabic Script IDNs

- Membership open, community based
- Details and interests of members posted by MESWG
- Discussions publicly archived
- Details at http://lists.meswg.org/mailman/listinfo/tf-aidn

- Open Call for Participation in August-Sept. 2013
- Announced at Arab IGF II in Algeria on 2\textsuperscript{nd} Oct. 2013
- 20 members from 15 different countries
- Covering a diverse set of languages which use Arabic script from East Asia, South Asia, Middle East, North Africa, Sub-Saharan Africa
Thank you