

# **Task Force on Arabic Script IDNs: Overview and Progress**

ICANN Singapore Meeting

Task Force on Arabic Script IDNs (TF-AIDN)  
Middle East Strategy Working Group (MESWG)

[tf-aidn@meswg.org](mailto:tf-aidn@meswg.org)

# History of Community Work on Arabic Script IDNs

- Very active Arabic script community; time line of Arabic script community
  - 2003: Arabic Domain Names Task Force (ADN-TF) by UNESCWA and LAS
  - 2005: Persian (Arabic script) SLD domain names offered
  - 2005: Arabic Domain Names Pilot Project
  - 2008: Arabic Script IDN Working Group (ASIWG)
  - 2009: ICANN IDN ccTLDs Fast Track Program initiated
  - 2010: Arabic Language in Internet Domains (RFC 5564)
  - 2010: First IDN ccTLDs delegated
  - 2011: Arabic Script Case Study for 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

# IDN TLDs Assigned or Delegated

1. الجزائر
2. عمان
3. ايران
4. امارات
5. بازار
6. باكستان
7. الاردن
8. بهارت
9. المغرب
10. السعودية
11. سودان
12. مليسيا
13. شبكة
14. سورية
15. تونس
16. مصر
17. قطر
18. فلسطين

# Code Points for Arabic Script IDNs

- Arabic script has the following specifications
  - ISO 15924 code: **Arab**
  - ISO 15924 no.: 160
  - English Name: Arabic
- Relevant sub-set
  - The complete set of code points in the Arabic script fall in the following Unicode ranges
    - Arabic U+0600 – U+06FF
    - Arabic Supplement U+0750 – U+075F
    - Arabic Extended A U+08A0 – U+08FF
  - Additional code points to be considered (as per Arabic Variant Issues Report)
    - Zero Width Non-Joiner U+200C
    - Zero Width Joiner U+200D

# Typology of Arabic IDN Variants

- Same
  - Identical U+06CC (ى)/U+0649 (ي)
  - In Context U+06A9 (كـبـك)/U+0643 (كـبـك)
  - Normalization U+0632 (ز)/U+0631+U+06EC (ر + ّ = ز)
- Similar
  - Character U+06AA (ـكـ)/U+06A9 (كـ)
  - Diacritic U+062A (تـ)/U+067A (تـ)
- Different
  - Shape U+0629 (بـة)/U+06C3 (بـة)
  - Character U+0629 (ة)/U+06C1 (هـ)

# IDN Variants Needs and Challenges

## Security and Stability Needs

پاکستان

U+0643

xn--mgbai9a5eva00b

پاکستان

U+06A9

xn--mgbai9azgqp6j

- 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 innumerability
  - 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.

# Community Driven Way forward: Task Force on Arabic Script IDNs

- Creation and oversight by community based Middle East Strategy Working Group (MESWG;  
<https://community.icann.org/display/MES/MESWG+Members> )
- TF-AIDN 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
  - Universal acceptability of Arabic script IDNs
  - Technical challenges around registration of Arabic script IDNs
  - Operational software for registry and registrar operations
  - DNS security matters specifically related to Arabic script IDNs
  - Technical training material around Arabic script IDNs

# Timeline of Formation

Aug 13

- Call for membership of TF-AIDN (released at Arab IGF II) by MESWG

Oct 13

- Formation of TF-AIDN by MESWG
- Initiation of Proposal Development for Formation of Generation Panel

Dec 13

- Informal consultation on final draft of proposal with ICANN/IP

Jan 13

- Submission of proposal to ICANN for formation of Arabic Script Generation Panel (GP)

Feb 13

- Acceptance of Proposal by ICANN → formation of Arabic Script GP
- All members of TF-AIDN members of Arabic Script GP



# Membership

- Currently **21 members** – applications still being received
- From **11 countries** – Egypt, Germany, Iran, Jordan, Lebanon, Malaysia, Morocco, Pakistan, Palestine, Saudi Arabia, Sudan
- Speaking more than **nine languages** – Arabic, Malay, Saraiki, Sindhi, Pashto, Persian, Punjabi, Torwali, Urdu, and African Languages
- With **expertise** in use of Arabic script from East Asia, South Asia, Middle East, North Africa, Africa
- Coming from **diverse disciplines** – academia (linguistics and technical), registries, registrars, national and regional policy bodies, community based organizations, technical community

# 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>

# Method of Work

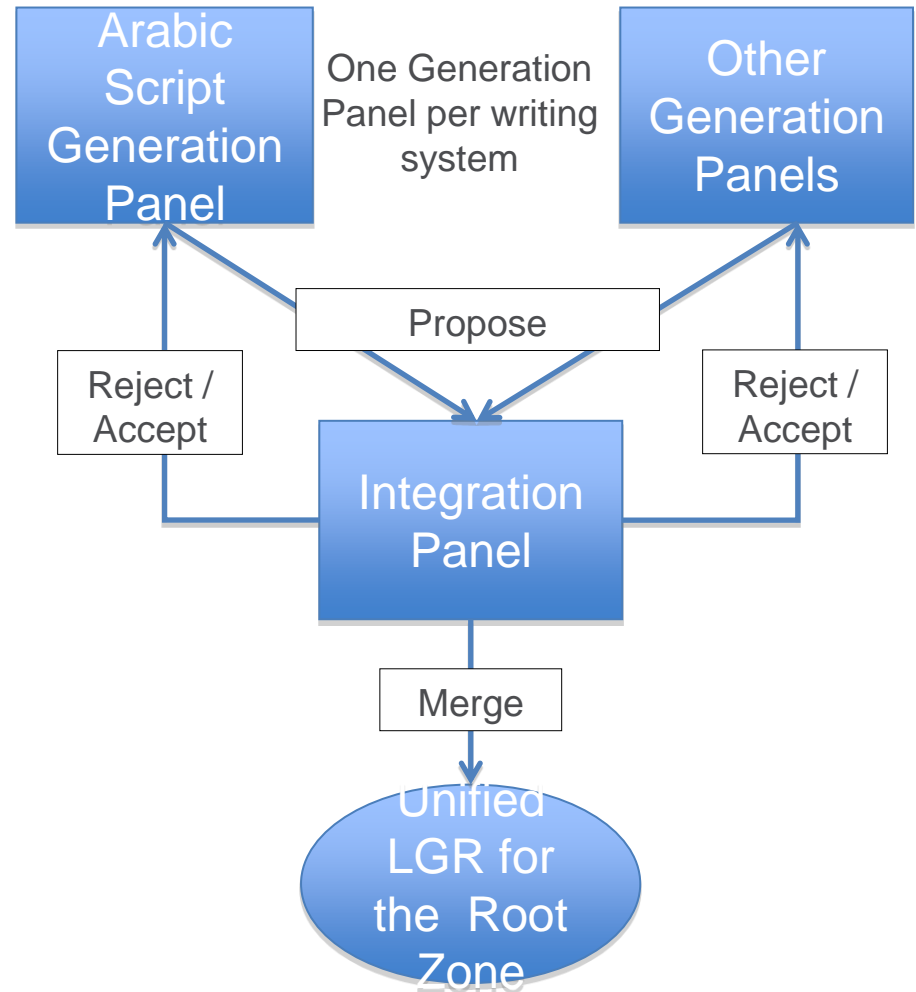
- Open call for each work item within TF-AIDN
- Volunteers develop the work item
- Work item presented to TF-AIDN for discussion
- Work item finalized with consensus
- Discussion within TF-AIDN archived at public list
- All teleconferences recorded and posted at public wiki page of TF-AIDN under MESWG wiki page
- All materials finalized after a public comment period
- URLs
  - Background and Introduction to TF-AIDN
    - <https://community.icann.org/display/MES/Task+Force+on+Arabic+Script+IDNs>
  - Workspace, news and document archive
    - <https://community.icann.org/display/MES/TF-AIDN+Work+Space>
  - Email Archive
    - <http://lists.meswg.org/pipermail/tf-aidn/>

# Proposal for Arabic Script GP

- Development of Proposal as per the ICANN guidelines
  - General Information
    - Target Script
    - Principle Languages
    - Countries with Significant Use
  - Composition of Panel
    - Panel Chair and Members
    - Panel Diversity
    - Relationship with Past Work
  - Work Plan
    - Suggested Timelines with Significant Milestones
    - Schedule of Meetings and Teleconferences
    - Sources of Funding
- Submission of proposal to ICANN

# Role of GP in IDN Project

- Generation Panels generate proposals for script specific LGRs, based on community expertise and requirements
- Integration Panel reviews proposals and integrates them into common Root Zone LGR while minimizing the risk to root zone as shared resource



# Current Work in Arabic Script GP

- Completed principles for inclusion, exclusion or deferral of code points for Arabic script LGR for the root zone - released for public comment
- Reviewing Code Points for Inclusion and exclusion from LGR
- Reviewing of MSR
- Discussing IDN Variants and defining general principles

# Principles for Code Points

- Inclusion principles, e.g.:
  - *Letter code point* which is a *letter* and has *established contemporary use* in a language
  - *Mark code point* which represents a *required mark*, where at least one of the *letter* it forms has *established contemporary use* in a language
- Exclusion principles, e.g.:
  - Code point either deprecated or not recommended for use in Unicode Standard; exception being it meets one of the inclusion criteria with no alternative code point(s)
  - Code point specifically for historic use with no *established contemporary use*
- Deferral principles, e.g.:
  - Code point which can neither be confirmed for inclusion nor for exclusion based on principles
  - Such code points will be considered in future versions of LGR, when more concrete information is available

# F2F Meeting in Singapore

- 10 participants funded by ICANN
- F2F Meeting Program: 20-22 Mar
  - Character repertoire (based on principles) [1.5 Days]
  - Review of general principles based [0.25 Days]
  - Review of MSR released by Integration Panel [0.5 Days]
  - Review of Variant and Whole Label Issues [0.75 Days]
- ICANN Meeting Program: 23-26 Mar
  - Training on IDN P1
  - Public sessions by TF-AIDN during the ICANN meeting
  - ICANN IDN Variant Program Update Session
  - Meetings with MESWG, Integration Panel, ICANN IDN Team, ICANN staff and policy makers



# Next Steps for GP

- Characters – Jan-Apr 14
  - Finalize code point inclusion/exclusion principles
  - Determine code point for inclusion/exclusion/deferral
  - Release for Public Comments
- Variants – Apr – Jul 14
  - Document principles for variants
  - Define variants
  - Release for Public Comments
- Whole Label Rules – Aug – Oct 14
  - Document principles for whole label variants
  - Define whole label variants
  - Release for Public Comments
- Finalization – Nov –Dec 14
  - Finalize LGR for Arabic script
  - Submit to ICANN/IP
  - Release for Public Comments

# Middle East DNS Forum

- Feb. 2013 – Stakeholders from the region
- Presented issues pertaining to IDNs to the community
- Engaged stakeholders and introduce TF-AIDN and the work been undertaken
- Received community feedback; Priority
  - Arabic Script IDN Universal Acceptance
  - Arabic Script Email
- Enlisted candidates to join the TF-AIDN from the community

# Next Steps for TF-AIDN

- Complete work on Arabic Script LGR
- Initiate additional work identified
  - Arabic Script Email
  - Universal Acceptability of Arabic Script IDNs

# ICANN's Support for the TF-AIDN

- Facilitating the work of the TF-AIDN
  - Dedicated wiki space at <https://community.icann.org/display/MES/TF-AIDN+Work+Space>
  - Arranging for conference calls when needed, and posting call recordings on the wiki space
  - Provided support to the first F2F meeting in Singapore; both logistically and financially
  - Arranging for meetings with relevant ICANN staff when needed
  - Addressing issues and concerns raised by TF-AIDN members
- Financial support for the next TF-AIDN meeting scheduled for September 2014
- Ensuring public visibility of the TF-AIDN at relevant fora, gatherings, and events.

تریمہ کاسیہ

شکراً

شکریہ

با سپاس

تشکر

Thank You