

Introduction to an IDN Revision

John Klensin

Why Make Changes?

- Many reasons, including
 - Some errors in IDNA2003
 - Tie to old version of Unicode – cannot accommodate newly-added scripts
 - Move from
 - Include everything if possible to
 - Letters and digits (extend traditional LDH model)
 - Reduce user confusion
 - Accommodate critical-but-dangerous characters
- Some technical changes

Some Key Goals

- Get a standard that is
 - Unicode version agnostic
 - Easier to understand
 - More predictable wrt what happens
 - More adaptable to local conditions – realistic interoperability
 - Better on other issues to extent possible

DNS Internationalization

- Not just IDNA
- Many areas of responsibility to make things work well
 - Standard Protocol
 - Registries and Registry Restrictions
 - Registrants – Increasing or decreasing confusion
 - Lookup implementers and Applications
 - Common sense by users and others
- Still mnemonics, not literature
- Can't *solve* confusion, but can provide better tools

New Terminology

- Much confusion in IDNA2003
- Standard about labels, not FQDNs...
 - Still some Right-to-left (bidi) questions
 - Hard issues with label separators
- U-label, A-label, LDH label
 - And all kinds of invalid strings

New Concepts - Structure

- Different rules for
 - Registration (similar to IDNA2003, but registry rules are explicit)
 - Lookup (actual rejections if violate rules)

New Concepts II – Contextual Rules

- Characters valid only in context
 - Mechanism for accommodating characters that would be too dangerous in other places
 - Rules specify context
- Joiners and Others Handled Differently
 - Joiners (“contextj”)
 - Verified on lookup and registration
 - Others (“contexto”)
 - Verified on registration,
 - Tested on lookup only for presence of rule
- “No rule” equivalent to “disallowed” until a rule exists

Compatibility

- Generally, if
 - String could actually be registered in IDNA2003 (stored and converted back)
 - String consisted of international letters, digits, and hyphen
 - ... then it is valid and unchanged in IDNA200X
- Applications that depend on character mappings to registered forms
 - Will need compatibility / transition handling
- A few previously mapped-out characters will change interpretation

The Front End

- What does mapping mean/do
- A minimal-variation on-the-wire form
- Compatibility and transition

Summary

- Trying to move ahead to make IDNs far more useful and predictable as mnemonics
- No big-jump revisions in future
- Still can't write literature