

# At-Large Comment on Reference Label Generation Rulesets (LGRs) for the Second Level

Bill Jouris

# Security Should be Mandatory Not Discretionary

- Current:

gTLD registry operators **may** consult these reference LGRs while they design their IDN tables to promote consistency. [Emphasis added]

- Proposed

gTLD registry operators **will** incorporate these reference LGRs when they design their IDN variant tables.

# Identifying Variants

- The document says:

Where variants are included, their selection is informed by existing registry practice, as well as by the work performed at ICANN on the script LGRs for the Root Zone. Any cross-repertoire (or cross-script) variants identified in the Root Zone have been retained here for use in zones that support reference LGRs for more than one script or language.

- No documentation on the source of “existing registry practice”
- Use of Root Zone LGRs as a basis is a problem

# Why Not Root Zone LGRs?

- Restricted Repertoires
  - Several scripts allow only a subset of the MSR for TLDs. Other code points, allowed for SLDs but not for TLDs, were never considered for possible variants.
- Narrow variant definitions
  - Homoglyphs and Near Homoglyphs
    - Distinguishable by experts
    - Side by side comparison
  - “Confusables” to be evaluated manually (for TLDs) by Similarity Review Panel
    - Manual review not feasible for SLDs

# Variants vs Confusables

- Criteria (Latin GP):
  - Side by side comparison. Not available to end users looking at a domain name
  - At least 5 of 7 GP experts unable to distinguish – variant
  - If *only* 4 of 7 GP experts unable to distinguish -- confusable

## Example

- Variants:           ć vs ċ           ń vs ñ           ź vs ż
- Not Variants:      é vs è           í vs i