

# IDN Protocol Revision

## Overview of details

Cary Karp

John Klensin

**Sygmnd was a poor Austrian  
who'd lost all the vowels in  
his name in a boating accident.**

*Pinchuck's Law*

Woody Allen

# Primary Issues

- Unicode relationships
- Registration-Lookup
  - Detailed explanations
  - Subtle substantive changes
- Characters
  - Internationalize LDH, not “make everything possible”
  - Make to/from ACE more symmetric (eliminate mappings)
  - Cannot get elegance of ASCII model without server-side matching... still making compromises

# New Terminology

- Confusion about IDNA2003
- Standard is about labels, not FQDNs...
  - Still some Right-to-left (bidi) questions
  - Difficult issues with label separators
- U-label, A-label, LDH label
  - Numerous further details

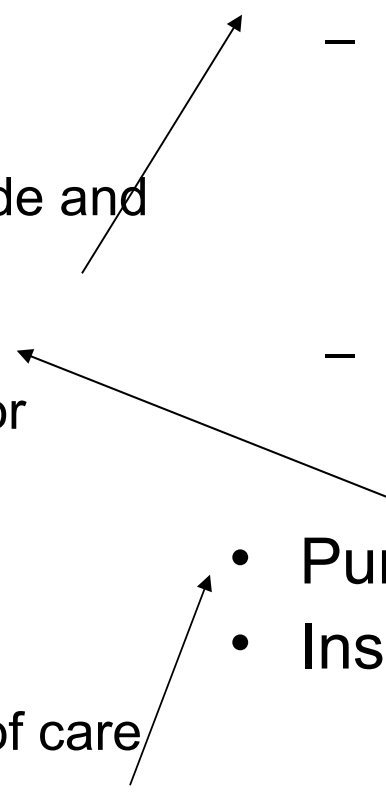
# Summary

- Trying to move ahead to make IDNs far more useful and predictable as mnemonics
- Improve understanding, usability, marketability
- No big-jump revisions in future
- Still not for literary purposes

# Registration Under IDNA2008

- Labels registered are final strings
  - No more mapping in protocol
  - Registries may use mappings optionally, but they will not reliably work at resolution time
- Explicit requirement to develop policies to keep identifiers safe and reliable
  - May want to think more about character exclusions and variants
  - Labels that violate standards will probably not be looked up by applications
- Registries and registrars share responsibility for keeping DNS useful
  - In ASCII DNS, protocol didn't prevent registering unusable strings

# Registration Steps

- No specifics in protocol
  - Pre-protocol
    - Proposed label
    - Conversion to Unicode and normalization check
  - Registry restrictions
    - Defined by registry for
      - Scripts it handles
      - Appropriate local conventions
    - No global rules
    - Just an expectation of care and responsibility
  - Protocol-level Validation
    - Character validation
      - Exclude DISALLOWED
      - Exclude UNASSIGNED
      - Exclude Context-required without rules
    - Label validation
      - Context of characters
      - Bidi rules
  - Punycode conversion
  - Insertion in zone
- 

# Some Scripts Used for Many Languages

- Uses are often different
  - Conventions
  - Rendering
  - Specific characters
- May require trade-offs among languages using script
  - Need for cooperation within script community
  - JET work for CJK – introduced “variants”
  - Current Arabic Script IDN effort



# Effects of the New Model

- By cooperating in making names safe and predictable...
- Registries avoid being punished by applications writers trying to safeguard their users
- No one wants to look at Punycode (!)