“Whois” Internationalization Issues

John C Klensin
Purpose of Panel

• Raise issues and questions for thought and policy development
• Not to recommend particular solutions
IDNs: A Remedial Course

- No actual non-ASCII characters in DNS – strings meet “hostname” constraints.
- Special encoding, called “punycode”
  - Applied as last step in conversion procedure
- Label: “xn—” plus gibberish. “xn—” is the hint that the decoding rules should be invoked.
- The real label – after decoding or before coding – is some Unicode form.
Internationalization Changes
Many Rules and Assumptions

• Port 43 Whois is defined as ASCII only
  – So can’t query using Unicode or get a response in it.

• Characters for query:
  – IDNA punycode or
  – Unicode (UTF-8) or
  – Local coded character set
  – A combination?? (multiple keys??)
  – One standard would be a good idea.
The Response

• Not much good if receiver can’t read it
  – All English?
  – All local language?
  – Local language plus English?
  – English… or choice of that or French, Russian, Chinese,…?

• Is it ok to expect someone to hire a translator?
Queries and Responses Again

• If can’t type the query, it will be hard to get an answer.

• Getting an answer in Klingon won’t help most of us, even if the query and database chars were to stay ASCII.
Variants

• Reserved names and their implications
• How much information about names in the package if one asks for one of them? If the one asked for is not the primary one? Or is reserved?
Summary

- Time to take this seriously
- Waiting will increase risk
- People who expect the problem to solve itself are going to be disappointed
- Use of “unusual” languages could make Whois useless
- The NVT constraint for this may kill Port 43 Whois
- Plan now, rather than having to clean up later.