EAI address syntax

IETF127, mailmaint

EAI address syntax

- draft-ietf-mailmaint-smtputf8-syntax-07
- draft-ietf-mailmaint-smtputf8-syntax defines simple rules for unicode addresses, yet risk-free even for unicode non-experts.
- draft-ietf-mailmaint-interoperable-addresses-03: parked for focus.

Current status

- Have four implementations now, would like WGLC.
- Settled all former open issues.
- Have checked each code point against UAX31, PRECIS, IDNA2008.
- Do have clarity about U+200D.

Current rules

1. An atom in an address MUST NOT be an A-label.
2. An address uses PRECIS IdentifierClass and adds ASCII space.
3. An address MUST NOT contain more than one script, when ASCII is disregarded. Three extra cases for China, Korea, Japan.
4. An address MUST be in the Unicode NFC form.

Rule changes since IETF126

1. Added a rule that generators must generate NFC.
2. ASCII space is allowed, since RFC733, 822, 2822 and 5322 agree.
3. Japanese is unbroken. I mismerged long ago, fixed now.

Text changes

1. Mention the other rulesets that apply (some readers do not understand that e.g. DNS rules apply even if never mentioned).
2. Add more examples.
3. Many fixes from John Klensin.

Thanks to John and other reviewers!

Open issue: status

```
RCPT TO:<a@example.com>
DATA
To: a@example.com
```

This document updates 6532 (the second address) but not 6531 (the first). Reasonable implementations use the same address for both.

Update both RFCs?

Recent discussion

1. Steffen Nurpmeso wants complete liberty to mix scripts; many people want to prevent homograph attacks. IMO too many people worry about homograph attacks; this document has no choice.
2. Several people (Narendra, Collin, John, more) suggested changes that I’ve largely applied.

Supplementary changes since IETF126

- examples/c/smtputf8-syntax.cpp in the github repository.
- The code is short, the tables are generated.
- The tables are updated to Unicode 18.0.0.
- Updated handling of the Catalan ela geminada (l·l).
- tests.json has grown, with more tests from reviewers.