precis mappings
draft-yoneya-precis-mappings-00

Yoshiro YONEYA

15 Nov 2011

IETF82@Taipei / PRECIS WG

Background

- Preparation of internationalized strings is essential to increase chance of matching.
- Normalization of character representation is one of the important preparation.
- Mappings of characters to canonical form is also one of the important preparation.
- Case mapping is defined in PRECIS Framework.
- How about any other mappings?

Mappings other than case (1/2)

- Width mapping
  - Some characters have fullwidth form and halfwidth form despite of identical.
  - Should be mapped into canonical form before case mapping.

    Ex: ａ → a, ９ → 9, ｱ → ア

  - Will increase backward compatibility with Stringprep (RFC3454) because it uses NFKC.

Mappings other than case (2/2)

- Delimiter mapping
  - Many protocols have delimiters (special meaning characters) which divide a string into components.

    Ex: !, @, :, ., /

  - Some of well known delimiter characters have compatible forms.

    Ex: 。 $\rightarrow$., ： $\rightarrow$:

  - Should be mapped into canonical form after width mapping.

Discussion

- Whether mappings other than case are targets of PRECIS or not?

- If they are target, are they a part of PRECIS Framework or separate ones like IDNA Mapping (RFC5895)?

- Are there another mappings not described in this document? For example, migration from Stringprep (RFC3454) to PRECIS Framework needs some special treatment?