

Kerberos Working Group

UTF-8 Stringprep Profile

UTF-8 Stringprep Profile

- Goals and Principles
- Applicability - Kerberos strings
- Character Repertoire - Unicode 3.1
- Unassigned Code Points - From Unicode 3.1
- Mappings
- Normalization
- Prohibited Output

Goals and Principles

- Improve interoperability
- Insure consistent representation
- Affects principals, realms, passwords
- Be compatible with IDN
- Don't reinvent the wheel

Mappings

Special-case mappings prior to normalization

- Mapped Out
 - Hyphens
 - Zero-Width Spaces
 - Variation Selectors and Joiners

- Mapped to ASCII space
 - Various visible spaces
 - Should we not do this?

Normalization

Standard conversion of unicode to a consistent form

- Unicode Normalization Form KC
 - Produces "compressed" representation
 - Consistent with IDN

- No special handling for Hanguul
 - IDN has decided this is inappropriate for them
 - We are not experts in this area
 - Not IETF's job to fix Unicode brokenness

Prohibited Output

These characters may not be used

- Control Characters
 - C0, C1, DELETE
 - Musical Controls
 - Other Controls

- Private-Use & Replacement Characters
 - Private-Use code points for nonstandard characters
 - Replacement for non-renderable characters

- Non-Character Code Points
 - Reserved code points which are not characters

- UTF-16 Surrogate Codes
 - Used as part of UTF-16 encoding

Prohibited Output

These characters may not be used

- Inappropriate for Plain Text
 - Special characters for structured text
- Inappropriate for Canonical Representation
 - Ideographic description characters
- Change Display Properties
 - Change display format
 - Change rendering order
- Tagging Characters
 - Used for tagging text