

International Searching and Sorting

Comparator Registry for Application Protocols

Comparators

- An abstract named function which takes two input octet strings and is used to perform one or more of three matching operations: equality, substring and ordering.
- Comparator names have some structure and can be matched with wildcards.
- Application protocols use comparator IANA registry to abstract details of comparison from protocol specification.

Relationship to Stringprep

- A stringprep profile by itself is not a comparator.
- Can construct a comparator from a stringprep profile, but this is not done automatically due to versioning.
- The ordering function is generally not useful if the comparator is constructed from a stringprep profile.

Comparator Definition

- Which Functions: equality, substring, order
- Canonicalization rules (may be identity)
- Precise lookup tables (for versioning)
- Equality: match / no-match / error
- Substring: match start&end, no-match, error
- Order: -1, 0, +1 (avoid “0”)

Application Protocol

- Mechanism to select comparator
- What happens if no explicit selection
- Way to list comparators (wildcards)
- For substring, start & end match positions
- Relevant app protocol error codes
- Disconnected operation: table fetching?

Initial Comparators

- i;octet
- i;ascii-numeric
- en;ascii-casemap
(legacy i;ascii-casemap)
- i;nameprep-v=1-uv=3.2
- i;basic-uca=3.1.1-uv=3.2 (customizable)

Basic Comparator

- Based on Unicode Collation Algorithm version 3.1.1 and version 3.2 of Unicode tables.
- Provides “adequate” ordering for international use.
- Can be customized for localized use.
- Open source implementation included in ICU project.

References

- Mail List:
 - ietf-comparator@imc.org
- Comparator Registry:
 - [draft-newman-il8n-comparator-00.txt](#)
- Example use in App Protocol:
 - [draft-ietf-imapext-il8n-00.txt](#)

Feedback:

- ISO Collation Registry
- “Maybe” response for equality & order
 - Like “undefined” in LDAP
- Check for version upgrade via Errata
- Consider “other” functions in registry:
scope creep vs. one place for similar things.