

The Use of Non-ASCII Characters in RFCs -bis

Paul Hoffman

IETF 116, Yokohama

March 2023

Introduction

- [draft-hoffman-rfc7997bis](#) -02 and -03 published after lots of discussion on the RSWG mailing list
- Considers current policies being used by the RPC to include non-ASCII characters in RFCs
- Simplifies some of the rules in RFC 7997
- Based on the wording from RFC 7997, so some parts are still out of place

Proposed changes to RFC 7997 (1 of 4)

- Added the role of the RFC stream approving bodies and the RFC Production Center (RPC) as described in RFC 9280. All responsibilities that were held by the RFC Editor in RFC 7997 are handled by the RPC. When the RPC has questions about the content of RFCs, the RPC asks the RFC stream approving bodies for input.

Proposed changes to RFC 7997 (2 of 4)

- Removed requirements of marking non-ASCII characters with XML markup. Clarified that person and names with non-Latin characters should have Latin transliterations.
- Removed the suggestion that character names for non-ASCII characters be included, instead relying on the "U+NNNN" syntax from BCP 137.

Proposed changes to RFC 7997 (3 of 4)

- Added the suggestion that non-ASCII characters generally appear in NFC.
- The basic requirements were softened to reflect the realities of the variability of search engines and web browsers.
- Changed "Rules for the Use of Non-ASCII Characters" to "Use of Non-ASCII Characters" because the RPC has used, and should continue to use, their own discretion based on what makes the RFC most useful.

Proposed changes to RFC 7997 (4 of 4)

- Language about the future was changed to the past tense to indicate that RFC 7997 was already implemented. For example, "RFCs will switch" was changed to "RFCs switched", and so on. Also added more acknowledgement of the use of non-ASCII characters outside the narrow scope that was defined in RFC 7997.
- Greatly rearranged and simplified the text about person, company, and postal names.

Proposal

- Adopt as a WG document
- Make sure the WG has consensus on each change, and modify or revert as necessary
- Clean up more language left over from RFC 7997
- Go through WG and IETF last calls, and publish as an RFC