Make char16/32_t literals UTF-16/32

P1041 – Targeting Core Language

R. Martinho Fernandes | cpp@rmf.io | https://wg21.link/p1041

Presented By: JeanHeyd Meneide | phdofthehouse@gmail.com

Motivation

- p1025r1 – Update the Reference to the Unicode Standard formally in the Working Draft
  - No longer need copy-pasted, handwaved definitions of UTF-16 and UTF-32 for char16_t and char32_t ("surrogate"?)
  - No longer need to reference UTF-8 through 2003 ECMAScript standard

Solution – Standardize Existing Practice

- Every surveyed compiler so far has just implemented UTF-8/16/32 as specified by the Unicode Consortium's Unicode Standard
  - One of the most existing-iest, de-facto practices we are aware of

- Implementers have 0 work to actually do
  - Gives users confidence they can work with UTF-16/32 without being afraid of possible handwaving implementations

Wording

- The paper has wording and is ready to send to Core