Unicode liaison report to ISO/IEC JTC 1/SC 22 #44

To: ISO/IEC JTC 1/SC 22
From: Robin Leroy (eggrobin@unicode.org)
Date: 2025-08-14

The liaison representative has worked with the Ada Rapporteur Group (under WG 9) and with SG 16 (under WG 21) to assist in interpreting and applying Unicode standards, and welcomes questions from other Working Groups, should such questions arise in their work.

Unicode Standard Annex #31, Unicode Identifiers and Syntax, which provides definitions for use in programming language syntaxes, continues to be actively maintained; a new version will be published as part of Unicode Version 17.0, to be released 2025-09-09. Until publication of Unicode 17.0, see planned changes in the Proposed Update.

The ICU Technical Committee has been designing modern C++ APIs for iteration over the code points of a string whose representation may be in UTF-8, UTF-16, or UTF-32 (in C++ terms, these are iterator adaptors and range adaptors). Programming language designers looking to add code point iteration to their standard libraries may find it interesting to look at these APIs, which, in addition to following a modern design, are informed by existing usage in ICU tools and library code itself, as well as in a large production codebase. See utfiterator.h in ICU4C.