

Title:	<b>TC 37/SC 2 Liaison Report to Unicode Consortium</b>
Author:	Peter Constable
Status:	Liaison Contribution
Action:	For consideration by Unicode Technical Committee
Date:	October 4, 2020

During the TC37/SC2 meeting in June of this year, a resolution was passed to assess potential need for revision of the ISO 12199:2000 standard:

**TC 37/SC2 Resolution 2020-02 (virtual Plenary on 2020, June 29)**

ISO/TC 37/SC 2 resolves to call for a group of experts to create an ad-hoc group for *ISO/12199:2000 - Alphabetical ordering of multilingual terminological and lexicographical data represented in the Latin alphabet* in order to assess the need for a revision or amendment, by September 30, 2020 in a view to be ready for the systematic review in April 2021.

This was prompted by a request by the Serbian national body (ISS) to revise ISO 12199 replacing references to the “Serbo-Croatian” language by “Serbian”, and to clarify the use of Latin script for that language.

It is noted that the Serbian national body also submitted a document to JTC1/SC2 (= JTC1/SC2 N4716) earlier this year requesting similar changes to various JTC1/SC2 standards, such as ISO/IEC 8859-1:1998. At their August meeting, JTC1/SC2 passed the following resolution:

**JTC1/SC2 RESOLUTION M 25-06: Contribution from Serbia**

SC 2 thanks ISS (Serbian national body) for the information provided in Document SC 2 N 4716 regarding simultaneous use of Serbian language in both Cyrillic and Latin scripts, as detailed in the Serbian Ad Hoc report in Document SC 2 N 4734. SC 2 recommends National Bodies and implementers of ISO/IEC 10646, stabilized standards for character coding, and ISO/IEC 14651 make reference to the Unicode CLDR data, available at the URLs below, that provides details regarding the Cyrillic and Latin orthographies for the Serbian language.

<https://unicode-org.github.io/cldr-staging/charts/37/summary/sr.html>

[https://unicode-org.github.io/cldr-staging/charts/37/summary/sr\\_Latn.html](https://unicode-org.github.io/cldr-staging/charts/37/summary/sr_Latn.html)

The TC37/SC2 meeting was held at an earlier date than the JTC1/SC2 meeting, and so the TC37/SC2 action was independent of and not informed by the JTC1/SC2 action.

In follow-up to the TC37/SC2 resolution, the TC37/SC2 secretariat has now issued a call for experts to assess ISO 12199:2000.

The ISO 12199 standard defines an ordering of Latin script characters in multilingual contexts. It’s statement of scope begins:

This International Standard specifies the sequence of characters to be used in the alphabetical ordering of multilingual terminological and lexicographical data (terms, term elements, or words) represented in the Latin alphabet. Character sets of languages represented in the Latin alphabet are taken into account insofar as terminological or lexicographical data have been recorded. Character sets used in internationally standardized transliteration into Latin script are also taken into account.

The sequence of alphabetical characters given is intended for multilingual purposes only and is not intended to affect the alphabetical order of any specific language.

It makes normative reference to the first edition of ISO/IEC 14651, which was not yet completed when ISO 12199:2000 was published. The normative clauses specify a four-level ordering of letters, digits and combining marks. Annex G, which is normative, provides a formal description of the rules as an application of ISO/IEC 14651.

The standard also includes informative information regarding languages written with Latin script and the character repertoires used for the orthographies of those languages.

Altogether, then, ISO 12199 has much in common with UTS #10 and supplementary, language-specific data in CLDR.

Here are some questions for UTC to consider:

- *Does UTC want to have participation in TC37/SC2's ad hoc group?*
- *Does UTC want to submit any comments or recommendations to TC37/SC2 (e.g., against revision of ISO 12199)*