



ISO/IEC JTC 1/SC 2
CODED CHARACTER SETS
SECRETARIAT: JAPAN (JISC)

DOC TYPE: Summary of Voting/Table of Replies

TITLE: Summary of Voting on SC 2 N 3389, ISO/IEC FCD 8859-16, Information Technology -- 8-bit single-byte coded graphic character sets, Part 15: Latin alphabet No. 10

SOURCE: Secretariat, ISO/IEC JTC 1/SC 2

PROJECT: JTC 1.02.20.16

STATUS: This voting summary is forwarded to WG 3 for consideration. WG 3 is requested to prepare a disposition of comments report, revised text and a recommendation for further processing.

ACTION ID: ACT

DUE DATE:

DISTRIBUTION: P, O and L Members of ISO/IEC JTC 1/SC 2
WG Conveners and Secretariats
Secretariat, ISO/IEC JTC 1
ISO/IEC ITTF

NO. OF PAGES: 7

ACCESS LEVEL: Defined

WEB ISSUE #: 081

Summary of Voting on SC 2 N 3389

	Approve	Approve with comments	Disapprove	Abstain	No response	Comments
P-member						
Armenia					X	
Austria					X	
Belgium	X					
Brazil					X	
Canada			X			Attachment 1
China	X					
Denmark	X					
Egypt					X	
Finland					X	
France					X	
Germany	X					
Greece	X					
Iceland					X	
India					X	
Iran, Islamic Republic of					X	
Ireland	X					
Israel					X	
Italy					X	
Japan	X					
Korea Rep. of	X					
Mongolia					X	
Morocco					X	
Netherlands				X		Abstain because NNI opposed to the endorsement of this work item by JTC 1
Norway	X					
Poland			X			Attachment 2
Romania	X					
Russian Federation					X	
Singapore	X					
Slovenia					X	
Sweden			X			Attachment 3
Thailand					X	
Tunisia					X	
Turkey	X					
UK		X				Attachment 4
USA			X			Attachment 5
Yugoslavia					X	
	12	1	4	1	18	
Total (36)	13		4	1	18	
O-member (if responding)						
Australia				X		Lack of expertise
Portugal				X		

Attachment 1: Canada

Due date 2000-03-15

Document number FCD 8859-16 (JTC1/SC2N3389)

Canada DISAPPROVES for attached reasons:

The current draft of CD-88591-16 has some serious deficiencies and should be rectified.

Comment 1:

The sixth paragraph in the Scope section has been copied from the template used from the previous parts of 8859 and is outdated. Several of the newer parts of 8859 are NOT reflected in 10367, and there are NO plans to revise 10367 at this time, nor is it considered wise to revise it.

Canada suggests the following modification to the above paragraph:

Change the second sentence to read:

"... 10367, or their corresponding G1 sets from the (in italics) 'ISO International Register of Coded Character Sets to be Used With Escape Sequences', should be used "

The reference to ISO-IR is already in the Bibliography.

The above change should be used in all parts of 8859 that are currently under revision, and for any future parts of 8859. Even if 10367 gets revised some day, the above modification will still be valid into the future.

Comment 2:

Note no. 2 under Table A.1 -- "Use of Latin Alphabet no. 2 for Romania is deprecated." MUST be removed.

8859-16 is being introduced into Romania the first time, and there are a large number of implementations and user data in Latin Alphabet no. 2 in Romania. Till such time 8859-16 becomes popular and all the migration issues are all sorted out by the implementers and users alike, it is inappropriate and premature to deprecate Latin-2 for Romania.

Attachment 3- Poland

Reasons for disapproval of
ISO/IEC FDC 8859-16
8-bit single-byte coded graphic character sets
Part 16: Latin alphabet No. 10

The clause on page 1 (inherited from the earlier parts of the standard)

This part of ISO/IEC 8859 may not be used in conjunction with any other of ISO/IEC 8859. If coded characters from more than one part are to be used together, by means of code extension techniques, the equivalent coded character sets from ISO/IEC 10367 should be used instead within a version of ISO/IEC 4873 at level 2 or level 3.

no longer makes sense, because the ISO/IEC 8859 standard family contains now characters not included ISO/IEC 10367 (e. g. Thai).

Attachment - Sweden

Comments on FCD 8859-16

As stated in the Swedish comments to JTC 1 N 5987 on the proposed inclusion of an ISO/IEC 8859 part 16 ("Latin-10") in the SC 2 Programme of Work, the Swedish NB is of the opinion that such a part shall not be developed at present. Consequently Sweden votes NO on N 3389, for the reasons given in the comments referred to above.

Sweden however submits the following editorial remarks on the English version of the FCD as circulated. (It could be noted that several of the comments apply to part 14 and/or 15 also, and were given by Sweden during the processing of those parts, but not responded to; Sweden however, although maintaining its comments on those parts, did not think it justified to later submit defect reports on the published standards.)

- a. The Foreword in its third paragraph refers to 8859-15; shall be 8859-16.
- b. The reference to part 13 in the Foreword shall not contain the expression "Baltic Rim"; that expression occurs neither in the title of part 13 nor anywhere in its text (although the corresponding ISO-IR 179 is named "Baltic Rim Supplementary Set" in the registration document).
- c. The cross-referencing to the French version of the standard in clause 4 "Definitions" seems to serve no real purpose; and also introduces an internal referencing between documents which may be difficult to maintain. If it is for some reason considered absolutely necessary, for this specific part, to align English and French definitions it should be done by explicit translations, not by references into the French document.
- d. The following error corrections shall be applied to Table 1:

10/05 A5 U+201E
10/11 AB U+00AB
11/03 B3 U+0142 LATIN SMALL ...
- e. The correct glyph shall be inserted in pos 15/00 of Table 2.
- f. The expression "Baltic Rim" shall be removed from the standard parts enumeration in Annex A.
- g. The names of languages in Table A.1 have in two cases been changed from what is used in Table A.1 for parts 1-13. Even if the new names may be more common than the earlier ones there seems to be no real reason to have a change from parts 1-13, since it was not stated during their processing that the names were in error. It appears that either the same names should be used in Table A.1 for this part as for the earlier ones; or that an SC2 decision should be taken to fully align the SC2 names for languages with those in ISO 639-2. Such an alignment, however, would impact on several of the names of table A.1, not only those changed in the FCD.
- h. Table A.1 introduces (as compared to parts 1-13) the languages "Cornish" and "Manx Gaelic". According to authoritative sources, both are extinct. Since the standard addresses "general purpose applications in typical office environments", Table A.1 should not include extinct languages; if so, a large number of languages could be added.
- i. The last sentence in Table A.1 note 4 should read: "These are included in ISO-IR 158, 197 and 209."
- j. The FCD has an A4-only layout. SC 2/WG 3 in its 1998 Seattle meeting decided that WG 3 standards should use a combined A4/letter size format (WG 3 resolution 13.13).

Attachment 4 - UK

UK COMMENTS ACCOMPANYING VOTE OF APPROVAL ON FCD 8859-16 (LATIN ALPHABET NO. 10)

Editorial comment

In "1 Scope", paragraph 6, line 6, after "from ISO/IEC 10367" insert the following words:
"or from the ISO International Register of Coded Character Sets".

Justification: The G1 set specified in this standard is in the ISO Register but is not in ISO/IEC 10367. This editorial change clarifies the use of this G1 set in a version of ISO/IEC 4873 at level 2 or 3.

BSI
24/02/2000

Attachment 5 - USA

January 5, 2000

Ballot: Letter ballot for FCD 8859-16 – Latin alphabet #10

Document: SC2 N3389

U.S. Vote: **Disapprove**

Comments:

We see no technical advantage in continuing to proliferate the 8859 series of 8-bit character encodings. By the time anyone would be able to roll out an 8-bit implementation of 8859-16 for a Romanian-specific application, Windows 2000 will have been distributed, making that the platform of choice for most desktop systems. That system will have support for **all** of the characters in 8859-16 (plus thousands, more, of course). Microsoft has no plans, as far as we can tell, to adjust its 8-bit code pages to adapt to this unending stream of new 8859 character encodings. And in the Unix and database worlds, UTF-8 solutions are now becoming more and more widely available. Once again, with UTF-8, all of the characters encoded in 8859-16 are **already** available.

8859-16 seems to be nothing more than a political response by a standards body to a perceived lack of sufficient stature for Romanian in international character standards. It is inappropriate as a **technical** response to the need to provide correct and culturally appropriate IT infrastructure for text processing in Romania.

Especially pernicious and objectionable in the FCD is the statement on page 7 that “Use of Latin alphabet No. 2 for Romanian is deprecated.” This deprecation is inappropriate since most Romanian data to date in encoded either in 8859-2 or its Windows extension, CP 1250. Correct conversion of that data to and from the UCS is complex enough now, with s-cedilla and s-comma-below separately encoded in the UCS, without the JTC1/SC2/WG3 further stirring the pot and muddling the character distinctions by promulgating an 8-bit character encoding for Romanian that only contains the s-comma-below and declaring, effectively, that Romanian data encoding using the s-cedilla of 8859-2 is deprecated. This is just going to lead to the situation where all Romanian processing is going to have to assume that U+0219 LATIN SMALL LETTER S WITH COMMA BELOW or U+015F LATIN SMALL LETTER S WITH CEDILLA could be meant for each other—they’ll be normalized together for spell-checking, searching, whatever, and rendering will be complicated by having to do context checking to see when both should be rendered with the comma below or both be rendered with the cedilla, or when a distinction is really meant. The very thing that Romania seemed to want—clear national identification of their letter with the comma below form— is going to end up being further muddled for decades. And 8859-16 is just contributing to that mess.

If 8859-16 passes and becomes an International Standard, as seems already a foregone conclusion, our best hope is that it will be widely ignored in implementation, to limit the damage for Romanian implementations based on the UCS.