

ISO/IEC JTC 1/SC 2 N 3970

DATE: 2007-09-25

ISO/IEC JTC 1/SC 2
Coded Character Sets
 Secretariat: **Japan (JISC)**

DOC. TYPE	Meeting Report							
TITLE	Recommendations and Notes taken at the 6th Meeting of ISO/IEC JTC 1/SC 2/OWG-SORT held in Hangzhou, China, 2007-09-19/20							
SOURCE	OWG-SORT							
PROJECT								
STATUS	This document is circulated to the SC 2 members for information.							
ACTION ID	FYI							
DUE DATE								
DISTRIBUTION	P, O and L Members of ISO/IEC JTC 1/SC 2 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF							
ACCESS LEVEL	Open							
ISSUE NO.	287							
FILE	<table border="1"> <tr> <td>NAME</td> <td>02n3970.pdf</td> </tr> <tr> <td>SIZE (KB)</td> <td></td> </tr> <tr> <td>PAGES</td> <td style="text-align: center;">8</td> </tr> </table>	NAME	02n3970.pdf	SIZE (KB)		PAGES	8	
NAME	02n3970.pdf							
SIZE (KB)								
PAGES	8							

Secretariat ISO/IEC JTC 1/SC 2 - IPSJ/ITSCJ (Information Processing Society of Japan/Information Technology Standards Commission of Japan)* Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8, Shiba-Koen, Minato-ku, Tokyo 105-0011 Japan *Standard Organization Accredited by JISC
 Telephone: +81-3-3431-2808; Facsimile: +81-3-3431-6493; E-mail: [kimura @ itscj.ipsj.or.jp](mailto:kimura@itscj.ipsj.or.jp)



DATE: 2007-09-21

ISO/IEC JTC 1/SC 2/OWG-SORT International String Ordering (ISO) - ISO/IEC 14651

DOC TYPE:	Report to SC2
TITLE:	Recommendations and notes taken at the sixth meeting of OWG-SORT, Hangzhou, held 19-20 September 2007
SOURCE:	Alain LaBonté, OWG-SORT Convener, and Dr Umamahesqaran, for the notes
PROJECT:	JTC 1.02.14651.00 - ISO/IEC 14651 (International String Ordering)
STATUS:	For action by SC2
ACTION ID:	Act
DUE DATE:	2007-09-21
DISTRIBUTION:	SC2
MEDIUM:	Electronic
NO. OF PAGES:	7

Part 1 : Recommendations to SC2 as a result of the meeting

Recommendation 1: Subdivision proposal for 14651: 2007/Amd. 1

OWG-SORT requests the SC2 Secretariat to send document **SC2 N 3969** for letter ballot as soon as possible after September 21.

Recommendation 2: Extending OWG-SORT beyond the next SC2 Plenary

Given the discussion held at the Hangzhou meeting and taking into consideration that the SC2 Plenary has been postponed to next April, it is recommended that SC2 issue a resolution, as duly requested by JTC1 Directives, extending the mandate of OWG-SORT until at least the SC2 Plenary following the April 2008 Plenary.

The mandate of the OWG would be to produce amendment 1 to the new edition of ISO/IEC 14651, as presented in **SC2 N 3969**, and to take any necessary action related to the progress of this International Standard, should any new issue occur during that mandate. It is expected that that mandate will need to be renewed at the 2009 SC2 Plenary as, according to the target dates, the FDAM production and ballot is expected for 2009, but after the SC2 Plenary.

Recommendation 3: Action against proliferation of informative annex matters for tailoring different languages' sort in future amendments

OWG-SORT recommends that after amendment 1, any future request for adding new informative annex topics dealing with tailoring different languages and scripts in future amendments (beyond or including what is already allowed in deltas) be questioned to avoid undue proliferation, and that other alternatives be sought (such as publishing a TR, placing the information into an international registry or equivalent, or any other means judged appropriate).

Recommendation 4: Solving Egyptian hieroglyphs sorting issue

Considering that work on Egyptian hieroglyphs sort is immature at this point and could not enter into the production of amendment 1 to the new edition of ISO/IEC 14651, OWG-SORT kindly encourages Mr. Michael Everson and Mr. Ken Whistler to come up with a way to arrive at a revised proposal taking into account actual and realistic needs of Egyptologists, ideally in requesting them to compare the actual results of this proposal versus the results they already achieve with transliteration and see if the potential proposal is acceptable to them. Just approval of a theoretical and untested method by Egyptologists in a letter should not be considered sufficient evidence for inclusion of permanent data in the CTT.

Recommendation 5: Encouraging National Bodies to review both language versions of ISO/IEC 14651

OWG-SORT requests the SC2 Chairman to prepare an SC2 Plenary agenda item appreciating the Canadian national body's support via the project editor in creating simultaneous English and French versions of ISO/IEC 14651 and encouraging NBs to review both language versions of the related ballot documents on this International Standard.

Recommendation 6: Equivalence between defined terms in different language versions

As a follow-up of recommendation 2 of the Frankfurt OWG-SORT meeting, this recommendation is amended as follows, to take into account that the language-version issue won't be brought up at the 2007 JTC1 Plenary:

SC2 chair is requested to bring to the attention of JTC1-SWG-D the following contradiction encountered during French and English ballot processing:

Contradiction between the 3 following mandatory rules to edit different language versions of an International standard :

- All language version of an International shall be equivalent, paragraph by paragraph
- Definitions shall be presented in alphabetical order in each language
- Each definition shall be sequentially numbered in each language version

It is typically impossible to achieve these 3 mandatory rules at once, because the equivalent definitions typically can not share the same order. ISO and IEC have chosen to ignore the first requirement, making in sort that the last two requirements currently have precedence in practice in published IS.

SC2 requests that a new rule be drafted to solve this issue by making sure that there is at least a cross-reference between terms, if possible in both English and French language versions, in particular when both versions are produced simultaneously during the early standard development, so that equivalences be possibly used by translators exploiting the involved international standard. In this way the three current rules would be respected in spirit.

Recommendation 7: Next meeting

OWG-SORT recommends to hold its next meeting co-located with the SC2 Plenary and SC2/WG2 meetings to be held in Redmond (Seattle), the week of April 21, 2008. Schedule will be adjusted after agreement between OWG-SORT and WG2 convenors.

For information, future meetings of SC2 are planned as follows :

WG2 - Fall 2008, seeking host; backup USA

WG2 - Spring 2009, Republic of Korea; backup USA

Plenary and WG2 - Fall 2009; Tokushima, Japan (near Osaka); backup USA

Part 2 : Raw notes taken at the meeting by Dr Umamaheswaran and slightly edited by the OWG-SORT convenor

Attendees:

Convenor:

Alain LaBonté **alabon@gmail.com**

Members:

Tatsuo Kobayashi	tlk@kobysh.com
Masahiro Sekiguchi	seki@jp.fujitsu.com
Kyongsok Kim	gim2007@asadal.pnu.edu
Michael Everson	everson@evertype.com
Deborah Anderson	dwanders@sonic.net
Peter Constable	petercon@microsoft.com
V.S. Umamaheswaran	umavs@ca.ibm.com

1. Approval of the agenda

Adopted as presented.

2. Information on ISO/IEC 14651(second edition) FDIS ballot and action to be taken **Document: SC 2 N 3964**

The FDIS was sent at the last meeting. ITTF just initiated the ballot a couple of weeks ago. The results will be available in 4 months. Only editorial changes are permitted in the ballot comments. We can assume that the FDIS will be published as 2nd edition after the ballot closes.

3. Potential Contribution on Korean preprocessing to be included at the end of informative Annex C at a later amendment stage **Document: SC 2 N xxxx (to be provided by Kim Kyongsok, from Korea)**

Prof. Kim: In Frankfurt, Kent Karlsson suggested that a new contribution was not needed. A simple change to the template table can accommodate the Korean needs. Dr Kyongsok requested feedback from Kent to know if the conditions in Kent's document are sufficient to satisfy all Korean needs. It is not clear whether the proposed Annex from Korea is needed or not.

Alain: What is in the table .. can it satisfy your needs?

Prof. Kim: Most .. as individual syllables it is OK. In other conditions it does not work. Current proposal is to provide an informative annex on how to preprocessing to make this happen. The document presented at the last meeting should be fine for this purpose.

A discussion took place regarding the concern about 'piling up the standard with too many language-specific informative annexes'.

Options are:

- a. to include the current proposed Annexes on Hangeul pre-processing.
- b. To produce Technical Reports on the pre-processing aspects.
- c. place the information into an international registry or equivalent
- d. place the information as a technical note on the Unicode web site

For now, we can continue with current direction of including the informative annex in the standard. If proliferation of such informative annexes becomes an issue in the future we have to reconsider what to do with all such informative annexes in the future.

Action item:

Prof. Kim / Alain: To circulate proposed material on pre-processing of Hangul that is to be included in the Informative Annex in 14651 before October 30, for comments by SC2 member bodies, by 15 December 2007.

Can revisit the question of proliferation at a later date if needed.

4. Contribution (Ireland) on sorting ancient Egyptian hieroglyphs (SC2 N 3966)

Michael Everson presented the proposal.

The code point order in 10646 is based on a classification system of the various Hieroglyphs. That is not the way the data containing these to be sorted. Gardiner's description is shown in the document. As an example, if they name files with Hieroglyphs, how should an operating system sort these to be friendly to the users. A binary order vs phonetic value comparison was done and found that the phonetic value ordering was closer to what the users wanted.

There are a number of characters having the same sound value. Some characters have two phonetic values. There are also some with three phonetic values. Some have no phonetic values -- these are logographic determinatives (?) and appear at the end of words.

The collations require some contractions that need to be done to get the expected sorting results. The most difficult part is to deal with intermixing of phonetic and ideographic values associated with a character. The Egyptologists tend to assign a default sound value to these.

Ken Whistler's feedback were looked at.

The email first:

Looks like these are beyond the scope of CTT. Special processing is required to deal with the phonetic order complications. These should be treated as code point order of CJK or Cuneiforms. Everson - did not agree with most of the points raised by Ken.

In terms of work the Egyptologist group will do most of the work.

Document SC2 N3966 has more detailed analysis of the contributions.

The proposed solution to include in CTT will expand the size and complexity of the CTT considerably.

The argument that Egyptologist does not have access to the mechanisms to invoke the right thing to do in an OS is to be accounted for. It has to be by mechanisms similar to the current locale selection.

Japanese collation has similar problems as well .. mapping of phonetics to CJK ideographs are also complex similar to what Michael has presented for Egyptian Hieroglyphs. For JP, there are mechanisms to deal with phonetics by look up tables etc.

Lots of feedback to Michael.

Action item: ME to work with KW to arrive at a revised proposal; taking into account other concerns that were raised at the meeting during the discussion.

6. Recommendation to SC2 Secretariat on issuing subdivision ballot for 1st amendment to 2nd edition

Document SC2 N 3969

Scope of subdivision for the next amendment was read out.

JP wanted to have the scope of the subdivision to only English version of Amd. and to remove the French version update.
(postponed discussion till item 8 was discussed).

7. Recommendation to SC2 Plenary on extending the mandate of OWG/SORT to take into account developments of the UCS (characters added by its Amendments 3 and 4)

Administrative. OWG-SORT Convenor to take care of it.

**8. Issue about language versions of ISO/IEC 14651, as mentioned in the first SC2 Business Plan (removed after a contribution of Canadian NB, but discussion requested by SC2 Chair)
Document: SC 2 N 3949 (Item 3.2.1, Risk 6), SC 2 N 3961**

Draft business plan from SC2 chair had identified Risk 6. Canada had submitted that the risks are not risks. We could review the Canadian contribution or discuss the risk.

A review of different clauses of the latest JTC1 Directives is done by the OWG convenor and SC2 chairman (in particular clauses 4.1.4, 7.9.1, 8.3.5, 8.3.6 [excerpts found by Alain], and annex E [excerpts found by Kobayashi San] of the latest version of the JTC1 Directives [not to be confused with ISO Directives, IEC Directives, and ISO/IEC Directives which probably have linguistic clauses in the same sense too]).

ITTF has to check the French version equivalence with English version (and according to Alain, it does so; in certain cases the discrepancies they find are not discrepancies but just due because as non experts they are not familiar with the specialized jargon, a fact that goes against a non-expert doing the translation after the fact; case in point: the use of the verb "localize" which does not mean "find the location of something" in our work, and whose translation may be corrupted if the idea is not rendered the way it is meant in the I.T. internationalization process).

If English and French versions of any document in SCs or WGs are available, both versions SHALL be distributed per JTC1 procedures.

By asking ITTF to translate later there is more risk – translation done by non technical personnel. Canada believes that this is a greater risk than progresing two versions at once. Furthermore there is more added value when experts concurrently progress the two versions at once and more people are exposed to being able to review at an earlier stage.

JP - has concern about process issue. Concern is not about the risk. JP agrees that En and Fr version be available as soon as possible. Review of CD, DIS etc. cycles .. have to be reviewed. We have disposition meetings etc. for each cycle. In case of this standard we were doing these for both versions. JP spent most of the time on En version and spent very small effort on Fr version. Did some time on Fr version in the internal review process. JP spent a lot of time at the last OWG SORT discussing the JP comments. Considering this experience we are proposing dropping the Fr version. Directives have several options in JTC1 directives .. we need a better option if available.

Ireland – There are two languages in Ireland. We review the standards in the language we are most comfortable in .. En. Sometimes the Fr. If there is a Fr editor he will update the technical changes to be

made to En version. If JTC1 directives say if both languages are available they SHALL be distributed, then SC2, WG2 or SC2/OWG-SORT should not question that. We have to trust the editor to make the Fr version.

Alain: Both versions are already available, so the directives must be applied as furthermore it does not harm anybody but is added value susceptible to enhance the quality of the standard. No national body has to look for discrepancies between two language versions but if they do, this will better the quality of the standard, just as no national body probably scrutinizes an unilingual English version at every sentence of a standard to see if it has ambiguities, but that scrutinizing, if done, can better the quality of the standard too, in particular if it is done by non-English-speaking natives.

At the last meeting, the JP input was welcome .. and it improved the standard in both En and Fr versions. Canada believes that it is in the better interest of the standard to have both versions available in each cycle. Again, JP is not obligated to review both versions but when it did, it brought better quality to both the English and French versions.

JP: The procedure issue .. if available shall be distributed .. it is not clear. If the OWG has a consensus to hide the Fr version we can do so.

Alain: The editor will not hide the French version. It is clear. It is available. Therefore it shall be distributed.

Uma: Canada believes that if the Fr version is available it is more beneficial to have both reviewed.

Ireland: we don't see any need to change what we are doing here. In the Script standard that was also done in simultaneous En Fr version we also saw advantages.

Kobayashi: Read the annex E out .. which supports the current bilingual version being continued if resources are available.

Would like an SC2 recommendation ... SC2 appreciates the Canadian national body's support via the project editor in creating the bilingual version .. Also SC2 encourages NBs to review both language versions of 14651 related ballot documents.

No changes to the way we do business.

9. Other business

No other business is brought on the floor.

10. Future meetings

Future meetings will be aligned with SC2 Plenary and WG2 meetings.

11. Closing

***** End of this Report *****