AGENDA - Malvern
ISO/IEC JTC 1 SC22/WG20
Meeting #19 - Internationalization

November 1, 2000

DATE: October 30 – November 1, 2000

TIME: 9:00 a.m. to 5:00 p.m.

LOCATION: Unisys Malvern
2476 Swedesford Road
Conference rooms A and B
Unisys Corporation
2476 Swedesford Road
Malvern, PA 19355

1. Introduction and announcements by Convenor

Welcome, logistics, etc..

2. Introduction of national delegations, liaisons, and cooperation partners [N787]

787	Participants – WG20 meeting in Malvern	Arnold Winkler	2000-10-30

Participants: LaBonté (Canada), Takata (Japan), Simonsen (Norway), Clews (UK), Whistler (USA), Winkler (convenor).

Germany has given a proxy to Alain LaBonté for this meeting.

3. Appointment of chairperson, secretary, and drafting committee

Chair and secretary: Arnold

Drafting: Keld, help is accepted.

4. Approval of prior meeting's minutes [N749]

749	Québec, May 2000, minutes	Winkler	2000-05-25

Whistler: Number of math symbols is 591 instead of 951 in 8.7, report about WG2.

Approved by unanimous consent.

5. Future Meeting Schedule and Plans

```
2001-05-08/10, Tübingen, Germany changed to May 9-11, 2001, confirmed, option
of CLAUI meeting preceding the WG20 meeting.
2001-10-1/3, Malvern, USA, other suggestions invited. (CEN co-location can not
be confirmed)
2002-05-20/22, Tromsø, Norway.
2002-10, tentatively Japan
```

6. Recognition of new documents and assignment to agenda items

794	Draft for ISO/IEC 15435	Keld Simonsen	2000-10-29	16.2
795	Known problems in case mapping tables for 14652	Ken Whistler	2000-10-23	14.2

7. Approval of Agenda [N784]

784	Agenda for the SC22/WG20 meeting in Malvern,	Arnold Winkler	2000-

	October 30 – November 1, 2000		

Agenda approved.

Priorities:

18.1 – future of WG20 so that a document can be drafted during the meeting.

14.2 – case folding for identifiers in COBOL

8. Report from the SC22 plenary in Nara, Japan [N766]

766	Participation in JTC 1 Development Activities	JTC1 N6206	2000-06-28
776	Resolutions from the SC22 plenary in Nara	SC22 N3178	2000-09-15

N766: SC22 is against the trial – FYI.

N767: AI Alain, send e-mail to Hill and Lisa to indicate that resolutions 16 and 7 are incorrect – it speaks of 14651-2 instead of 14651-1:2000, the published standard.

Documents were discussed, document issue particularly, see resolution 00-05. WG20 will follow the recommendations.

Make the "fair use policy" part of the liaison report from WG20 to Unicode.

9. Report from the CLAUI meeting in France [N788]

788	Resolutions from CLAUI meeting in France	CLAUI meeting	2000-10-19
798	Proposals regarding recommendations Sophia 7-11 from the CLAUI report (JTC1 N6336)	WG20 meeting	2000-10-30

Sophia 4: review of SC35 work. Not a new project for WG20.

Sophia 7: WG20 considers SC2 as the appropriate group to develop such a document, WG20 would provide input and feedback.

SC22 should instruct its working groups to regularly check the resulting document. We don't know, how SC2 is going to react to it. The suggestion came from Jim Melton, Yves, and Melagrakis. Audience needs to be clearly defined (developers of standards). Such reports must be answered quickly, not by standards or TRs from WG20.

Sophia 8: Dangerous proposition to "make" TRs that have not consensus" normative. Keld will contact Lisa Rajchel to have an authoritative reading on that issue.

Sophia 9: The second sentence about transliteration needs to be stricken (use of 14652). The TR 14652 has no consensus and should not be made normative in any standard. TC46/SC2 does not refer to 14652 in their standards. No consensus to do anything.

Keld: it could be done in another way ! (talking about the 14652 dependent actions).

Sophia 10: new work item for SC35 for a compliance check list – the usefulness depends on the scope of the work as defined by SC35. The danger is that procurement uses this list indiscriminately for purchasing products. This might result in work for WG20.

WG20 does not have any desire to produce a list of tests for compliance with I18N. Unicode is developing such a test, it can also be used for 14651. Ken had suggested such a work in WG20 years ago, no interest was expressed. If WG20 would define culturally correct tailoring for specific instances, WG20 should also develop test for these languages.

WG20 has presently no intentions to develop a list of compliance requirements, but is prepared to review and comments on such a list, if I18N issues are included.

Sophia 11: we need a paper to propose actions for 7-12 (where 12 is done and 11 is the action itself).

Sophia 12: done, Tübingen in May 2001, in connection with WG20 meeting.

Proposed document is N798 (Ken, Arnold). We will decide what to do after having seen the paper.

The (over night) produced document N798 was reviewed and accepted.

Resolution in acceptance of N798.

AI: Winkler send N798 to JTC1 secretariat (for forwarding to CLAUI).

10. Liaison Reports

10.1 Additions/deletions/changes to liaisons

SD-4	List of liaison and co-operations to and from SC22/WG20	Winkler	2000-05-22

Add IETF (Keld's action)

Make dormant ITU-T for the time being.

AI.: Arnold will inform Stefan Fuchs.

10.2 SC22/WG4, COBOL

793	Liaison report from COBOL (SC22/WG4) to WG20	Ann Wallace SC22/WG4 N0146	2000-10-26

See discussion and solution of the liaison document in 14.2 below.

10.3 SC22/WG5, Fortran

No report.

10.4 SC22/WG14, C

Keld reports about the meeting in Toronto. 9899 is published, first corrigendum in the works. Keld presented the API standard (C version only), and asked if they would be interested in using it – the result showed no preference for or against. C group would review the API standard.

10.5 SC22/WG15, POSIX

Keld reports about the latest meeting, sorting problem and regular expression was discussed and left "implementation defined". Ulli Dreper suggests a new API for case sensitive ranges. Currently the POSIX utilities are supposed to be locale dependent, but not all are. 14652 does not address regular expressions, a note could be added. Regular expressions belong into the API standard.

Keld suggests to POSIX that ranges can be specified using the collation algorithm (with all levels). Needs correct architecture, sort tables, and will to do it.

14652 should be aligned with POSIX work, especially regarding character width. Use of the repertoire map was rejected, defect report is filed against the solution of using charmaps.

10.6 SC22/WG21, C++

Nothing to report.

10.7 SC2/WG2

See Ken's report is in N791. Many characters are in the pipeline, 1000+ for the BMP, 40000+ in plane 2. Impact on the sort table is minimal for the BMP (mainly symbols, some small Philippine scripts). In part 2 the problem is also limited. Table is going to be a lot bigger.

Keld recommends to use the accelerated JTC1 process for amendment 2 – no subdivision, do registration and PDAM in one ballot.

10.8 GUIDE/SHARE Europe

Dormant.

10.9 SC35

See CLAUI, agenda item 9.

10.10 CEN TC 304

TC304 is a bit in disarray, lots of work is waiting for the independent market relevance report. Next meeting is delayed to February 2001 – no co-location with WG20 is possible in this case.

Guide to character sets and migration might be of interest

AI Clews – get a draft of the CEN "Guide to the use of character set standards in Europe" to WG20.

10.11 TC37

See discussion of language codes in 18.2

10.12 Unicode

Ken reports: UTC met in Boston, Unicode 3.0.1 update is released (now new characters, some properties changed).

Case-folding and normalization-test files were released for conformance testing.

U+5 digits and U+6digits is approved (now inn ballot for WG2). UAX, UTS, and UTR are officially announced.

"32 In-process" codes have been defined – NOT-A-CHARACTER. Not for data interchange.

New TRs 21-case mapping, equivalences, case folding clarifications.

TR 24 is script names, referencing ISO/IEC 19154 in a table. The names in UTR and in 19154 have different objectives. Table is in the report. Aimed at regular expressions. Feedback desired.

Working with IETF for international Domain Names (IDN) specifications. DNS is not case significant, syntax must be maintained. Matching and browsing is vaguely related to IDN discussion.

Major discussion about UTF-8 perceived security issues (non shortest form). This issue will be discussed in detail at the UTC/L2 meeting in San Diego, November 7-10, 2000)

10.13 W3C

Nothing to report.

10.14 IETF

Keld is the designated liaison, once the liaison is established.

11. Review of prior meetings action items [SD-5]

SD-5	Action item list	Winkler	2000-10-23

Done. Result and new items will be distributed with date November 1, 2000

12. NP for a TR about the dangers of not observing I18N

No action at this meeting, Alain invites contributions for SC35.

13. International string ordering ISO/IEC FCD 14651

13.1 Publication of IS 14651 in electronic form [N766, resolution 00-12]

776	Resolutions from the SC22 plenary in Nara	SC22 N3178	2000-09-15

Congratulations to Alain for the approval of this major standard !

SC22 approved the publication in electronic form as requested in convenor's report.

Takata: Japan has voted NO on the DIS. Some technical inconsistencies have been identified. Japan will send a defect report, the editor can use this in the new draft.

AI – Takata: send the defect report for ISO/IEC 14651

13.2 Amendment #1 [N776, resolution 00-07]

776	Resolutions from the SC22 plenary in Nara	SC22 N3178	2000-09-15
792	14651 table for Unicode 3.0	Ken Whistler	2000-10-28

SC22 approved the work item for the amendment #1. Alain is approved by SC22 as the editor.

Ken: there were a few issues for the approved standard which could be addressed now: sorting of numbers (Kent), XML format for the table, etc...

Main issue is the extension to the table, normative part of the standard don't need to be affected by the amendment.

Normative references to the table, to ISO/IEC 10646, etc. need updates and must be considered.

N792 is the document that describes current status of the table with call for help in this area. Ken makes proposals about his possible solutions. Synchronization with UTC next week in San Diego.

Any input from native speakers would be appreciated. Radicals as characters or symbols or both – sort them together or not?

Real scripts problems: vowels for Semitic scripts (Hebrew, Arabic, Syriac, Thaana).

Ethiopic needs input (Daniel Jacob).

How does Sinhala sort ? According to the local character set ?

Tibetan extensions: variable form letters

Thai, Lao, Tibetan, Myanmar, Khmer, in this order. Input needed for Myanmar and Khmer.

Mongolian is a mess – variants make it complex.

Cherokee is OK, also Canadian Syllabics due to good input from experts. Ogham is in order, also Runic, but variants might change in ballot. All before Hangul, Katakana, Bopomofo, Yi,

Input appreciated from the experts.

Ken: the sort standard is a good example for extensibility – amendment can be done fast and correctly.

14. Cultural convention specifications ISO/IEC PDTR 14652

14.1 Progression of PDTR 14652

783	Comments on ISO PDTR 14652	Markus Kuhn	2000-10-12

Discussion of Markus' paper in the meeting (mainly Ken and Keld).

Progression:

New document (N801)	November 15, 2000	Simonsen
Final check by WG20	December 1, 2000	WG20 editorial committee
Possible changes due to review	December 8, 2000	Keld Simonsen
Send to SC22 for DTR ballot	December 10, 2000	Arnold Winkler

Keld will write the document in WP with an ISO approved template, Arnold does not have to do any formatting.

14.2 Identifiers for COBOL (case mapping)

793	Liaison report from COBOL (SC22/WG4) to WG20	Ann Wallace SC22/WG4 N0146	2000-10-26
795	Known problems in case mapping tables for 14652	Ken Whistler	2000-10-23
799	Liaison report to COBOL , responding to N793	WG20	2000-10-31

Ann asks for "tolower" to be added to TR 10176 and to make the TR 14652 case mapping table compatible to the Unicode table.

WG20 does not think that adding the table to TR 10176 is a not a good idea, it would introduce additional synchronization problems.

Ken and Keld suggest to align the 14652 and the Unicode case mapping table.

Cobol shows some interest in API functionality for a future version of COBOL.

Conference call to Ann Wallace. N799 is the document.

In a telephone conversation with Ann Bennett she accepted gladly the draft copy of N799 and will not push for including "tolower" into TR 10176. Ann will reference TR 14652 in COBOL. N799 will explicitly spell out the need to synchronize the "tolower" table in 14652 with the Unicode data.txt table.

15. Registration of cultural elements ISO/IEC 15897

782	Updated issues list for ISO/IEC 15897	Keld Simonsen	2000-10-11
755	Draft ISO 15897 – Registry for cultural elements	Keld Simonsen	2000-10-25

Keld indicates that all comments have been disposed in the new draft. Joint advisory committee was added. This will increase the quality control for submissions. Discussion about sponsoring authorities: preference for NBs, other sources are accepted with the source named, WG20 should not be chartered with this job.

Ken points out that informative material need to be removed from normative annexes. Keld agrees to do that for the next draft.

Alignment with the new RFC 1766 – by reference ?

More details are pointed out, especially the normative references to ISO/IEC 14652. The US will object to any of these clauses.

Discussion about the format in which the specifications are presented (XML, 14652, etc..). If XML, the specification does not need to be in ISO 646, XML needs to be well formed. Unicode tables don't describe all necessary specifications, delete the references.

Progression of 15897:

New document (N803)	January 15, 2001	Keld Simonsen
Final check by WG20	January 31, 2001	WG20
Possible changes due to review	February 10, 2001	Keld Simonsen
Send to SC22 for registration ballot	February 15, 2000	Arnold Winkler

Keld will write the document in WP with an ISO approved template, Arnold does not have to do any formatting.

16. Internationalization API standard ISO/IEC 15435

16.1 Retention of project 22.15435 for 1 year [N766, resolution 00-21]

776	Resolutions from the SC22 plenary in Nara	SC22 N3178	2000-09-15

Convenor objects to SC22 telling him, how to do his job. If WG20 has a high quality document, it will be submitted for ballot.

16.2 Project 15435

752	Issues list #4 for ISO 15435	Keld Simonsen	2000-05-19
794	Draft for ISO/IEC 15435	Keld Simonsen	2000-10-25

Due to fundamental changes in the document, a new issues list will be prepared.

UK (Clews) asks, who is interested in working on this standard. Alain explains that Canada is interested, but they can not participate in meetings. WG20 would appreciate input from Canada, best in form of a document. UK is not interested, Clews considers the current document not a mature document.

Keld points out, that other organizations are interested (CEN through its rather old user requirement document).

Ken: US is against it.

Kent (Sweden) has communicated that the design is lousy and the project should not be progressed.

In Canada there is interest by Francois Yergeau and Glen Seeds. But not in C.

Germany has no big interest, and also no expertise.

Japan has strong objection to the design, both from Kido for the LINUX I18N group as well as the SC22 mirror committee.

Keld: supporters for the API project:

Markus Kuhn ?

Ireland ?

Leca in France?

Christoffersen in Denmark. Promised to provide feedback, but we have not seen any.

Keld explains that these are new commitments.

Clews recommendations to Keld: Rationale for the standard in a WG20 document. Make a much more mature standard for the next meeting.

Arnold: can this be done in WG14 – C?

Keld promises to prepare the rational document, also considering the availability of alternate sources such as Open Source.

Progression:

Produce rationale paper (N805) and new draft for 15435 (N804)	March 1, 2001	Keld Simonsen
Review by the editing committee	March 15, 2001	Editing group
Final document	March 22, 2001	Keld Simonsen
Distribution of rationale and draft standard with a cover letter requesting contributions and commitment to SC22.	April 2, 2001	Arnold Winkler

Contributions and commitments are needed before the next meeting of WG20 in Tübingen in May of 2001.

17. ISO/IEC 10646 Issues

789	Resolutions from the SC2/WG2 meeting in Athens September 2000	SC2 N3489R SC2/WG2 N2254R	2000-10-11
791	Ken Whistlers report about the SC2/WG2 meeting in Athens, September 2000	Ken Whistler L2/00-369	2000-10-06

No issues were identified.

18. Other business

18.1 The future of SC22/WG20 [N776, resolution 00-08]

773	Personal thoughts about the future of WG20 for consideration at the SC22 plenary in Nara	Arnold Winkler SC22 N3164	2000-08-30
774	Collection of reactions to "The future of WG20" (August 30 – September 6, 2000)	E-mails	2000-09-07
775	More reactions to "The future of WG20" (September 7 – 13, 2000)	E-mails	2000-09-13
797	WG20 response to resolution 00-08 from the SC22 plenary in Nara	WG20 meeting	2000-10-31

Alain: the reason for the creation of WG20 still exist, the direction might have to be changed from programming languages and environments to more user oriented subjects. Satisfy user expectations.

Sort should remain in WG20.

Keld: CLAUI meeting has discussed the overall view of I18N, some recommendations to SC35 and call for cooperation.

Clews: what to do with developments outside ISO?

Alain: this outside development has always be the case, it is often brought into ISO after industry acceptance.

Keld: 11017 does not cover keyboards and other stuff outside the functionality of I18N.

Possible additional countries: Thailand, Ireland (sort only), Buthan, ???

Winkler: Draft paper in this meeting, request input from countries that do not participate, finalize report to SC22 in Tübingen?

Ken: US supports the convenor's recommendation about the future in N773.

Discussion about procedural issues – would SC2 have to request to take over the sort standard.

Alain: sort standard could also be in SC2 due to the availability of the knowledge.

Clews: UK has no strong opinion, where sort is being handled.

Ken: some of WG20 standards are not requested by the implementation community, it are requests such as the one from Ann Wallace, that make it necessary – problem lies with JTC1's inability to reference work outside JTC1 work normatively in their standards.

Alain: Large companies do not like standards.

Keld: just don't use standards …

Ken: the problem is that some standards are referenced in other standards and thus create problems for implementations.

Keld: has spoken to Microsoft people who are using 14652 information.

Ken: in 14652 it is the format that is an issue. Interoperability problems are introduced.

Generic discussion about advantages of open source (ICU) over formal standards process for the implementers' community. Even more generic discussion about ISO process to develop International Standards in the face of other standards developing organizations such as IETF, W3C, OpenGroup, IEEE, and Unicode. Some of the work is well coordinated, other work is not.

Ken raises the scope of work in WG20: if no changes are done, 11017 is unchanged, 10176 might remain stable or have table changes, 14651 needs table changes, 15897 needs consideration of the registration process, leaving 15652 and 15435. New work as outlined by Alain should have made a difference in participation.

Clews suggests to maintain WG20 and slowly get rid if the projects, if so desired.

Ken points out that WG20 standards should not screw up situations where consensus already exist. Keld indicates that the API standard is based on requirements.

14652 – move to POSIX might be a possibility: Keld indicates that WG15 could take it, but has some problems with it. Keld wants to look at format (XML, SGML). Ken is absolutely against that. Specification of a format is the question of the effected committee.

Consensus statements:

The US would not consider it the end of the world, if WG20 would continue to exist. Japan supports the US position.

If WG20 continues to exist, no re-assignment of existing projects should happen.

Continue with 14651 amendment 1; continue with 14652 publication as a TR and there is no consensus about future work after publication; continue with the amendment of 15897 to make it a good registration standard; review annex A of 10176 in the context of 10646; and leave 11017 alone. No consensus exists about the project 22.15435, and about the expansion of the scope of work of WG20.

AI – Winkler: send the resulting document N797 to SC22.

Resolution to approve the document N797.

18.2 Language codes

759	Language codes report to WG20 from CEN/TC304	John Clews	2000-05-23
760	Electronic commerce and CLAUI – examples for the use of language and country codes	Jake Knoppers JTC1 N5626	1998-12-07
777	Tags for the Identification of Languages	H. Alvestrand RFC 1766bis5	October 2000
780	Language codes – report to WG20	John Clews	2000-09-21
781	Language identification and IT – addressing linguistic diversity on a global scale	Peter Constable Gary Simons (SIL)	2000-09-22

John explains a chaotic situation in ISO. 2 letter codes in ISO 639. Part 1 is being updated, presently in FDIS state. TC37 is responsible for part 1.

TC46 is responsible for 639-2 and 639-3, for 2 letter codes and 3 letter codes respectively. TC46 has no secretariat.

IETF has RFC 1766 for 2 letter codes.

There are not enough languages covered by ISO 639 codes. (US MARC is equivalent to 639-2).

SIL has many more codes in the Ethnologue – about 7000 language codes.

N777 describes the plan of IETF for a new RFC.

HTML 4.0 in W3C also considers additional language tags.

Private names are also in use. The new internet draft will allow for standardized tags for Ethnologue name with private sub-tags. Cross mapping will be complex or impossible.

WG20 should NOT get involved in the definition of these codes, we should reference the RFC when it is approved.

US supports the SIL proposal.

18.3 Checking E-mail distribution list sc22wg20@dkuug.dk

790	E-mail distribution list <sc22wg20@dkuug.dk>	Simonsen	2000-10-25

Arnold asks the participants to check the accuracy and completeness of their NB addresses and send possible updates to Keld.

Consideration of implementing a closed list.

19. Review of Priorities and Target Dates

Done

20. Review of Actions Items from this meeting

Done

21. Approval of Resolutions

785	Resolutions from the SC22/WG20 meeting in Malvern, October 30 – November 1, 2000	Arnold Winkler	2000-11-01

Anonymous consent on all resolutions with the exception of R0011-05, where the USA abstains, and R0011-08 which was approved by acclamation.

22. Adjournment

Adjourned.,